On total claim amount for marked Poisson cluster models

Bojan Basrak, Olivier Wintenberger, Petra Zugec

On total claim amount for marked Poisson cluster models

Bojan Basrak[a,*], Olivier Wintenberger[c], Petra Žugec[c]

[a]Department of Mathematics, University of Zagreb, Bijenička 30, Zagreb, Croatia
[b]LPSM, Sorbonne university, Jussieu, F-75005, Paris, France
[c]Faculty of Organization and Informatics, Pavlinska 2, Varaždin, University of Zagreb, Croatia

Abstract

We study the asymptotic distribution of the total claim amount for marked Poisson cluster models. The marks determine the size and other characteristics of the individual claims and potentially influence arrival rate of the future claims. We find sufficient conditions under which the total claim amount satisfies the central limit theorem or alternatively tends in distribution to an infinite variance stable random variable. We discuss several Poisson cluster models in detail, paying special attention to the marked Hawkes processes as our key example.

Keywords: Poisson cluster processes, limit theorems, Hawkes process, total claim amount, central limit theorem, stable random variables

1. Introduction

Elegant mathematical analysis of the classical Cramér–Lundberg risk model has a prominent place in nonlife insurance theory. The theory yields precise or approximate computations of the ruin probabilities, appropriate reserves, distribution of the total claim amount and other properties of an idealized insurance portfolio, see for instance Asmussen and Albrecher [2000] or Mikosch [2009]. In recent years, some special models have been proposed

*Corresponding author

Email addresses: bbasrak@math.hr (Bojan Basrak), olivier.wintenberger@upmc.fr (Olivier Wintenberger), petra.zugec@foi.hr (Petra Žugec)

to account for the possibility of clustering of insurance events. For instance, in the context of Hawkes processes, some results on ruin probabilities can be found in Stabile and Torrisi [2010] and Zhu [2013]. General cluster point processes and Poisson cluster processes in particular, have been proved useful in a variety of fields when modeling events that cluster either in space or time. This includes seismology, telecommunications, forensic science, molecular biology or finance, we refer to section 6.4 in Daley and Vere Jones [2003] for some examples.

The main goal of this article is to study asymptotic distribution of the total claim amount in the setting where Cramér–Lundberg risk model is augmented with a Poisson cluster structure. To make this more precise, we model arrival of claims in an insurance portfolio by a marked point process, say

$$N = \sum_{k=1}^{\infty} \delta_{\tau_k, A^k} \,,$$

where τ_k's are nonnegative random variables representing arrival times with some degree of clustering and A_k's represent corresponding marks in a rather general metric space $\mathbb{S}$. Observe that we do allow for the possibility that marks influence arrival rate of the future claims. In the language of point processes theory, we assume that the marks are merely unpredictable and not independent of the arrival times [Daley and Vere Jones , 2003]. For each marked event, the claim size can be calculated using a measurable mapping of marks to nonnegative real numbers, $f(A^k)$ say. So that the total claim amount in the time interval $[0, t]$ can be calculated as

$$S(t) = \sum_{\tau_k \le t} f(A^k) = \int_{[0,t]\times\mathbb{S}} f(a) N(ds, da) \,. \tag{1}$$

In the sequel, we aim to determine the effect of the clustering on the quantity $S(t)$, as $t \to \infty$ even in the case when the distribution of the individual claims does not satisfy assumptions of the classical central limit theorem. The paper is organized as follows — in the following section we rigorously introduce marked Poisson cluster model and present some specific cluster modes which have attracted attention in the related literature, see Faÿ et al. [2006], Stabile and Torrisi [2010], Karabash and Zhu [2015]. As a proposition in Section 3 we present the central limit theorem for the total claim amount $S(t)$ in our setting under appropriate second moment conditions. In Section 4, we prove a functional limit theorem concerning the sums of regularly varying nonnegative random variables when suboordinated to an independent renewal process. Based on this, we prove the limit theorem for the total claim amount $S(t)$ in cases when individual claims have infinite variance. Finally in Section 5 we apply our results to the models we introduced in Section 2. In particular, we give a detailed analysis of the asymptotic behavior of $S(t)$ for marked Hawkes process which have been extensively studied in recent years.

2. The general model

Consider an independently marked homogeneous Poisson point process with intensity ν on $[0,\infty)$ with marks in a completely metrizable separable space $\mathbb{S}$,

$$N^0 = \sum_{i\geq 1} \delta_{\Gamma_i, A_i}\,.$$

Marks A_i are assumed to follow a common distribution Q on a measurable space $(\mathbb{S}, \mathcal{S})$ where $\mathcal{S}$ denotes a corresponding Borel σ–algebra. In other words, N^0 is a Poisson point process with intensity $\nu \times Q$ on the space $[0,\infty)\times\mathbb{S}$. For non–life insurance modeling purposes, the marks can take values in $\mathbb{R}^d$ with coordinates representing the size of claim, type of claim, severity of accident, etc.

Denote the space of locally finite point measures on this space by $M_p = M_p([0,\infty)\times\mathbb{S})$ and assume that at each time Γ_i with mark A_i another point process in M_p is generated independently, we denote it by G^{A_i}. Intuitively, point process G^{A_i} represents a cluster of points that is superimposed on N^0 after time Γ_i. Formally, there exists a probability kernel K, from $\mathbb{S}$ to M_p, such that, conditionally on N^0, point processes G^{A_i} are independent, a.s. finite and with the distribution equal to $K(A_i,\cdot)$. Based on N^0 and clusters G^{A_i} we define a cluster Poisson process.

In order to keep the track of the cluster structure, we can alternatively consider the process G^{A_i} as a part of the mark attached to N^0 at time Γ_i. Indeed,

$$\sum_{i\geq 1} \delta_{\Gamma_i, A_i, G^{A_i}}$$

can be viewed as a marked Poisson process on $[0,\infty)$ with marks in the space

$\mathbb{S} \times M_p$. We can write

$$G^{A_i} = \sum_{j=1}^{K_i} \delta_{T_{ij}, A_{ij}} ,$$

for some $\mathbb{N}_0$ valued random variable K_i. If we count the original point arriving at time Γ_i, the actual cluster size is $K_i + 1$. Further, for any original arrival point Γ_i and corresponding random cluster G^{A_i}, we introduce a point process

$$C_i = \delta_{0,A_i} + G^{A_i} .$$

Note that K_i may possibly depend on A_i, but we do assume throughout that

$$\mathbb{E}K_i < \infty . \tag{2}$$

It is useful to introduce a time shift operator θ_t, by denoting

$$\theta_t m = \sum_j \delta_{t_j + t, a_j} ,$$

for an arbitrary point measure $m = \sum_j \delta_{t_j, a_j} \in M_p$ and $t \geq 0$. Finally, to describe the size and other characteristics of the claims together with their arrival times, we use a marked point process of the form

$$N = \sum_{i \geq 1} \theta_{\Gamma_i} C_i .$$

Hence, by the assumptions above, point process N can be represented as a random element in M_p of the form

$$N = \sum_{i=1}^{\infty} \sum_{j=0}^{K_i} \delta_{\Gamma_i + T_{ij}, A_{ij}} , \tag{3}$$

where we set $T_{i0} = 0$ and $A_{i0} = A_i$. In this representation, the claims arriving at time Γ_i and corresponding to the index $j = 0$ are called ancestral or immigrant claims, while the claims arriving at times $\Gamma_i + T_{ij},\ j \geq 1$, are referred to as progeny or offspring. Moreover, since N is locally finite, one could also write

$$N = \sum_{k=1}^{\infty} \delta_{\tau_k, A^k} ,$$

with $\tau_k \le \tau_{k+1}$ for all $k \ge 1$. Note that in this representation we ignore the information regarding the clusters of the point process. Clearly, if the cluster processes G^{A_i} are independently marked with the same mark distribution Q independent of A_i, then all the marks A^k are i.i.d.

The size of claims is produced by an application of a measurable function, say $f : \mathbb{S} \to \mathbb{R}_+$, on the marks. In particular, sum of all the claims due to the arrival of an immigrant claim at time Γ_i equals

$$D_i = \int_{[0,\infty)\times\mathbb{S}} f(a) C_i(dt, da)\,, \tag{4}$$

while the total claim size in the period $[0,t]$ can be calculated as

$$S(t) = \sum_{\tau_k \le t} f(A^k) = \int_{[0,t]\times\mathbb{S}} f(a) N(ds, da)\,.$$

Remark 2.1. *In all our considerations, we take into account (without any real loss of generality) the original immigrant claims arriving at times Γ_i as well. In principle, one could ignore these claims and treat Γ_i as times of incidents that trigger, with a possible delay, a cluster of subsequent payments. Such a choice seems particularly useful if one aims to model the so called incurred but not reported (IBNR) claims, when estimating appropriate reserves in an insurance portfolio [Mikosch , 2009]. In such a case, in the definition of the process N, one would omit the points of the original Poisson process N^0 and consider*

$$N = \sum_{i\ge 1} \theta_{\Gamma_i} G^{A_i} = \sum_{i=1}^{\infty}\sum_{j=1}^{K_i} \delta_{\Gamma_i + T_{ij}, A_{ij}}\,,$$

instead.

2.1. Some special models

Several examples of Poisson cluster processes have been studied in the monograph of Daley and Vere Jones [2003], see Example 6.3 therein for instance. Here we study marked adaptation of the first three examples 6.3 (a)-(b) and (c) of Daley and Vere Jones [2003].

2.1.1. Mixed binomial Poisson cluster process

Assume that the clusters have the following form

$$G^{A_i} = \sum_{j=1}^{K_i} \delta_{W_{ij},A_{ij}} ,$$

with $(K_i, (W_{ij})_{j\geq 1}, (A_{ij})_{j\geq 0})_{i\geq 0}$ being an i.i.d. sequence. Assume moreover that $(A_{ij})_{j\geq 0}$ are i.i.d. for any fixed $i = 1, 2, \ldots$ and that $(A_{ij})_{j\geq 1}$ is independent of $K_i, (W_{ij})_{j\geq 1}$ for all $i \geq 0$. We allow for possible dependence between $K_i, (W_{ij})_{j\geq 1}$ and the ancestral mark A_{i0}, however, we assume that K_i and $(W_{ij})_{j\geq 1}$ are conditionally independent given A_{i0}. As before we assume $\mathbb{E}[K] < \infty$. Such a process N is a version of the so–called Neyman–Scott process, e.g. see Example 6.3 (a) of Daley and Vere Jones [2003].

2.1.2. Renewal Poisson cluster process

Assume next that the clusters G^{A_i} are i.i.d. and independent of N^0, with the following distribution

$$G^{A_i} = \sum_{j=1}^{K_i} \delta_{T_{ij},A_{ij}} ,$$

where K_i's are $\mathbb{N}_0$ valued random variables with finite mean and $(T_{ij})_j$ represents a renewal sequence

$$T_{ij} = W_{i1} + \cdots + W_{ij} ,$$

for some array of i.i.d. nonnegative random variables (W_{ij}) independent of the sequence (K_i) and an array of i.i.d. nonnegative random variables (A_{ij}) representing the marks. A general unmarked model of this type is called Bartlett-Lewis model and analyzed in Daley and Vere Jones [2003], see Example 6.3 (b). See also Faÿ et al. [2006] for an application of such a point process to modeling of teletraffic data.

2.1.3. Marked Hawkes process

Key motivating example in our analysis is the so called (linear) marked Hawkes process. It can be introduced as a point process $N = \sum_k \delta_{\tau_k, A^k}$, where random marks (A^k) have identical distribution Q on the space $\mathbb{S}$ and arrivals (τ_k) have the conditional intensity of the form

$$\lambda_t = \lambda(t) = \nu + \sum_{\tau_i < t} h(t - \tau_i, A^i) , \tag{5}$$

where $\nu > 0$ is a constant and $h : [0,\infty)\times\mathbb{S} \to \mathbb{R}_+$ is assumed to be integrable in the sense that $\int_0^\infty \mathbb{E}h(s,A)ds < \infty$. Denote by $N|_I$, restriction of the point process N to the set $I \times \mathbb{S}$. Observe, λ is $\mathcal{F}_t$—predictable, where $\mathcal{F}_t$ stands for its internal history, $\mathcal{F}_t = \sigma\{N|_I, I \in \mathcal{B}(\mathbb{R}), I \subset (-\infty, t]\}$. Moreover, A^n's are assumed to be independent of the past arrival times τ_i, $i < n$, see also Bremaud [1981]. Writing $N_t = N((0,t]\times\mathbb{S})$, one can observe that (N_t) is an integer valued process with nondecreasing paths. The role of intensity can be described heuristically by the relation

$$\mathbb{P}(dN_t = 1 \mid \mathcal{F}_{t-}) \approx \lambda_t dt\,.$$

It turns out that Hawkes processes of this form have a neat Poisson cluster representation due to Hawkes and Oakes [1974]. For this model, the clusters G^A are recursive aggregation of Cox processes, i.e. Poisson processes with random mean measure $\mu_A \times Q$ where μ_A has the following form

$$\mu_A((0,t]) = \int_0^t h(s,A)ds\,,$$

for some fertility (or self–exciting) function h, cf. Example 6.4 (c) of Daley and Vere Jones [2003]. More precisely, if $N^A = \sum_{l=1}^{L_A} \delta_{\tau_l^1, A_l^1}$ is a Poisson processes with random mean measure $\mu_A \times Q$, the cluster process corresponding to a point (τ, A) satisfies the following recursive relation

$$G^A = \sum_{l=1}^{L_A} \left(\delta_{\tau_l^1, A_l^1} + \theta_{\tau_l^1} G^{A_l^1}\right)\,,$$

where the sequence $(G^{A_l^1})_l$ on the r.h.s. is i.i.d., distributed as G^A and independent of N^A. Thus, to each ancestral point, (Γ_i, A_i) we add a cluster of points denoted by C_i, which contains this point and in which this and any other point generates another independent marked point process to the right in time; all these newly generated point processes are again Poisson conditionally on the corresponding mark cf. Example 6.3 (c) of Daley and Vere Jones [2003].

Under the assumption

$$\kappa = \mathbb{E}\int h(s,A)ds < 1\,, \tag{6}$$

the total number of points in a cluster is generated by a subcritical branching process. Therefore, the clusters are finite almost surely, and we denote their size by $K_i+1 = C_i[0,\infty)$. It is known and not difficult to show that under (6), the clusters always satisfy

$$\mathbb{E}K_i+1 = \frac{1}{1-\kappa}\,.$$

Observe that the clusters C_i are independent by construction and can be represented as

$$C_i = \sum_{j=0}^{K_i} \delta_{\Gamma_i+T_{ij},A_{ij}}\,, \tag{7}$$

with A_{ij} being i.i.d. and $T_{i0} = 0$. We note that in the case when marks do not influence conditional density, i.e. when $h(s,a) = h(s)$, random variable K_i+1 has a so-called Borel distribution with parameter κ, see Haight and Breuer [1960]. Observe also that in general, marks and arrival times of the final Hawkes process N are not independent of each other, rather, in the terminology of Daley and Vere Jones [2003], the marks in the process N are only unpredictable.

2.1.4. Stationary version

In any of the three examples above, the point process N can be clearly made stationary if we start the construction in (3) on the state space $\mathbb{R}\times\mathbb{S}$ with a Poisson process $\sum_i \delta_{\Gamma_i}$ on the whole real line. The resulting stationary cluster process is denoted by N^*. Still, from applied perspective, it seems more interesting to study the nonstationary version where both the ground process N^0 and the cluster process itself have arrivals only from some point onwards, e.g. in the interval $[0,\infty)$ as for instance in Karabash and Zhu [2015].

Stability of various cluster models, i.e. convergence towards a stationary distribution in appropriate sense has been extensively studied for various point processes. For instance, it is known that the unmarked Hawkes process on $[0,\infty)$ converges to the stationary version on any compact set and on the positive line under the condition that

$$\int_0^\infty sh(s)ds < \infty\,, \tag{8}$$

see Daley and Vere Jones [2003], p. 232. Using the method of Poisson embedding, originally due to Kerstan [1964], Bremaud and Massoulie [1996]

(Section 3) obtained general results on stability of Hawkes processes, even in the non–linear case.

3. Central limit theorem

As explained above, the total claim amount for claims, arriving before time t, can be written as

$$S(t)=\sum_{\tau_k\le t} f(A_k)=\int_0^t\int_{\mathbb{S}} f(u)N(ds,du)\,.$$

The long term behavior of $S(t)$ for general marked Poisson cluster processes is the main goal of our study. As before, by Q we denote the probability distribution of marks on the space $\mathbb{S}$. Moreover, unless stated otherwise, we assume that the process starts from 0 at time $t=0$, that is $N(-\infty,0]=0$.

In the case of the Hawkes process, the process $N_t=N([0,t]\times\mathbb{S}),\ t\ge 0$ which only counts the arrival of claims until time t has been studied in the literature before. It was shown recently under appropriate moment conditions, that in the unmarked case multitype Hawkes processes satisfy (functional) central limit theorem, see Bacry et al. [2013]. Karabash and Zhu [2015] showed that N_t satisfies central limit theorem even in the more general case of nonlinear Hawkes process and that linear but marked Hawkes have the same property. In the present section we describe the asymptotic behavior of the total claim amount process $(S(t))$ for a wide class of marked Poisson cluster processes, even in the case when the total claim process has heavy tails, and potentially infinite variance or infinite mean.

It is useful in the sequel to introduce random variables

$$\tau(t)=\inf\{n:\Gamma_n>t\}\,,\ t\ge 0\,,$$

and

$$D_i=\int_{[0,\infty)\times\mathbb{S}} f(u)C_i(ds,du)=\sum_{j=0}^{K_i} f(A_{ij})=\sum_{j=0}^{K_i} X_{ij}\,,$$

where $K_i+1=C_i[0,\infty)$ denotes the size of the ith cluster and where we denote $X_{ij}=f(A_{ij})$. As before, D_i has an interpretation as the total claim amount coming from the ith immigrant and its progeny. Observe that in the nonstationary case we can write

$$S(t)=\sum_{i=1}^{\tau(t)} D_i-D_{\tau(t)}-\varepsilon_t\,,\ t\ge 0\,, \tag{9}$$

where the last error term represents the leftover or the residue at time t, i.e. the sum of all the claims arriving after t which belong to the progeny of immigrants arriving before time t, that is

$$\varepsilon_t = \sum_{0 \le \Gamma_i \le t, t < \Gamma_i + T_{ij}} f(A_{ij}) \quad t \ge 0\,.$$

Clearly, in order to characterize limiting behavior of $S(t)$, it is useful to determine moments and the tail behavior of random variables D_i for each individual cluster model. To simplify the notation, for a generic member of an identically distributed sequence or an array, say (D_n), (A_{ij}), we write D, A etc. Under the conditions of existence of second order moments and the behavior of the residue term ε_t, it is not difficult to derive the following proposition.

Proposition 3.1. *Assume that* $\mathbb{E}D^2 < \infty$ *and that* $\varepsilon_t = o_P(\sqrt{t})$ *then, for* $t \to \infty$*,*

$$\frac{S(t) - t\nu\mu_D}{\sqrt{t\nu\mathbb{E}D^2}} \xrightarrow{d} N(0,1)\,, \tag{10}$$

where $\mu_D = \mathbb{E}D$*.*

Proof. Denote the first term on the r.h.s. of (9) by

$$S^D(t) = \sum_{i=1}^{\tau(t)} D_i \;\; t \ge 0\,.$$

An application of the central limit theorem for two-dimensional random walks, see [Gut , 2009, Section 4.2, Theorem 2.3] yields

$$\frac{S^D(t) - t\nu\mu_D}{\sqrt{t\nu\mathbb{E}D^2}} \xrightarrow{d} N(0,1)\,,$$

as $t \to \infty$. Since we assumed $\varepsilon_t/\sqrt{t} \xrightarrow{P} 0$, it remains to show that

$$\frac{D_{\tau(t)}}{\sqrt{t}} \xrightarrow{P} 0 \quad t \to \infty.$$

However, this follows at once from [Gut , 2009, Theorem 1.2.3] for instance, or from the fact that in this setting sequences (Γ_n) and (D_n) are independent. $\square$

In the special case $f \equiv 1$, one obtains the central limit theorem for the number of arrivals in time interval $[0, t]$. Related results have appeared in the literature before, see for instance Daley [1972] or Karabash and Zhu [2015]. The short proof above stems from the classical Anscombe's theorem, as presented in [Gut , 2009, Chapter IV] (cf. [Daley , 1972, Theorem 3 ii]) unlike the argument in Karabash and Zhu [2015] which relies on martingale central limit theorem and seems not easily extendable esp. for heavy tailed claims we consider next.

4. Infinite variance stable limit

It is known that if the claims are sufficiently heavy tailed, properly scaled and centered sums $S(t)$ may converge to an infinite variance stable random variable. In the case of random sums $S_n = X_1 + \cdots + X_n$ of i.i.d. random variables, the corresponding statement is true if and only if the claims are regularly varying with index $\alpha \in (0, 2)$. For the Cramér–Lundberg model, i.e. when $N = N_0$, with i.i.d. regularly varying claims of index $\alpha \in (1, 2)$, corresponding limit theorem follows from Theorem 4.4.3 in Gut [2009]. A crucial step in the investigation of the heavy tailed case is to determine the tail behaviour of the random variables of (4)

For regularly varying D_i with index $\alpha \in (1, 2)$, limit theory for two-dimensional random walks in Section 4.2 of Gut [2009] still applies. Note, if one can show that D_i's have regularly varying distribution, then there exists a sequence (a_n), $a_n \to \infty$, such that

$$nP(D > a_n) \to 1 , \qquad n \to \infty,$$

and an α–stable random variable G_α such that $S_n^D = D_1 + \cdots + D_n$, $\qquad n \to \infty$, satisfies

$$\frac{S_n^D - n\mu_D}{a_n} \xrightarrow{d} G_\alpha , \tag{11}$$

where $\mu_D = \mathbb{E}D_i$. It is also known that the sequence (a_n) is regularly varying itself with index $1/\alpha$, see Resnick [1987]. In the sequel ,we also set $a_t = a_{\lfloor t \rfloor}$ for any $t \geq 1$.

4.1. Case $\alpha \in (1, 2)$

In this case, the arguments of the previous section can be adopted to show.

Proposition 4.1. *Assume that D_i's are regularly varying with index $\alpha \in (1,2)$ and that $\varepsilon_t = o_P(a_t)$, then there exists an α–stable random variable G_α such that for $\mu_D = \mathbb{E}D_i$*

$$\frac{S(t) - t\nu\mu_D}{a_{\nu t}} \xrightarrow{d} G_\alpha\,, \tag{12}$$

as $t \to \infty$.

Proof. The proof again follows from the representation (9), by an application of Theorem 4.2.6 from Gut [2009] on random walks (Γ_n) and (S_n^D) together with relation (11). By assumption $\varepsilon_t/a_{\nu t} \sim \nu^{-1/\alpha}\varepsilon_t/a_t \xrightarrow{P} 0$, to finish the proof, observe that the sequences (Γ_n) and (D_n) are independent, hence

$$\frac{D_{\tau(t)}}{a_{\nu t}} \xrightarrow{P} 0, \qquad t \to \infty.$$

□

4.2. Case $\alpha \in (0,1)$

In this case, we were not able to find a theorem for two–dimensional random walks of the type used above. Therefore, as our initial step, we prove a theorem which we believe is new and of independent interest. It concerns partial sums of i.i.d. nonnegative regularly varying random variables, say (Y_n), subordinated to an independent renewal process. More precisely, set $V_n = Y_1 + \cdots + Y_n\,, n \geq 1$. Suppose that the sequence (Y_n) is independent of another i.i.d. sequence of nonnegative and nontrivial random variables (W_n). Denote by

$$\sigma(t) = \sup\{k : W_1 + \cdots + W_k \leq t\}$$

the corresponding renewal process, where we set $\sup\emptyset = 0$. Recall that for regularly varying random variables Y_i's there exists a sequence (a_n) such that $n\mathbb{P}(Y_i > a_n) \to 1$, as $n \to \infty$. The following functional limit theorem describes the asymptotic behavior of $V_{\sigma(t)}$ in this case.

Theorem 4.1. *Suppose that (Y_n) and (W_n) are independent nonnegative i.i.d. sequences of random variables such that Y_i's are regularly varying with index $\alpha \in (0,1)$, and such that $0 < 1/\nu = \mathbb{E}W_i < \infty$. Then in the space $D[0,\infty)$ endowed with Skorohod's J_1 topology*

$$\frac{V_{\sigma(t\cdot)}}{a_{\nu t}} \xrightarrow{d} G_\alpha(\cdot)\,, \qquad t \to \infty, \tag{13}$$

where $(G_\alpha(s))_{s\geq 0}$ is an α–stable subordinator.

Proof. Since Y_i's are regularly varying, it is known, Resnick [1987, 2007], that the following point process convergence holds as $t \to \infty$

$$M'_n = \sum_i \delta_{\frac{i}{n}, \frac{Y_i}{a_n}} \xrightarrow{d} M_\alpha \sim \mathrm{PRM}(Leb \times d(-y^{-\alpha}))\,, \tag{14}$$

with respect to the vague topology on the space of Radon point measures on $[0,\infty) \times (0,\infty]$. Abbreviation PRM stands for Poisson random measure indicating that the limit is a Poisson process. Starting from (14), it was shown in [Resnick , 2007, Chapter 7] for instance, that for an α–stable subordinator $G_\alpha(\cdot)$ as in the statement of the theorem

$$V'_n(\cdot) = \frac{V_{\lfloor n\cdot \rfloor}}{a_n} \xrightarrow{d} G_\alpha(\cdot)\,, \qquad t \to \infty, \tag{15}$$

in Skorohod's J_1 topology on the space $D[0,\infty)$. Observe that since $\alpha \in (0,1)$, no centering is needed, and that one can substitute the integer index n by a continuous index $t \to \infty$. Note further that we have the joint convergence

$$(M'_t, V'_t) \xrightarrow{d} (M_\alpha, G_\alpha), \qquad t \to \infty, \tag{16}$$

in the product topology on the space of point measures and càdlàg functions. Moreover, it is known that the jump times and sizes of the α–stable subordinator G_α correspond to the points of the limiting point process M_α. The space of point measures and the space of càdlàg functions $D[0,\infty)$ are both Polish, in respective topologies, therefore, Skorohod's representation theorem applies. Thus, we can assume that convergence in (16) holds a.s. on a certain probability space $(\Omega, \mathcal{F}, P)$, and in particular there exists $\Omega' \subseteq \Omega$, such that $P(\Omega') = 1$ and for all $\omega \in \Omega'$, $V'_t \to G_\alpha$ in J_1 and $M'_t \to M_\alpha$ in vague topology. By Chapter VI, Theorem 2.15 in Jacod and Shiryaev [2003], for any such ω there exists a dense set $B = B(\omega)$ of points in $[0,\infty)$ such that

$$V'_t(s) \to G_\alpha(s)\,, \qquad t \to \infty,$$

for every $s \in B$, where actually B is simply the set of all nonjump times in the path of the process G_α. On the other hand, it is known that in J_1 topology, on some set Ω'' such that $P(\Omega'') = 1$,

$$\frac{\sigma(t\cdot)}{t\nu} \to id(\cdot)\,, \qquad t \to \infty, \tag{17}$$

where *id* stands for the identity map. This follows directly by an application of Theorem 2.15 in Chapter VI of Jacod and Shiryaev [2003]. Moreover, by Proposition VI.1.17 in Jacod and Shiryaev [2003], the convergence in (17) holds locally uniformly on $D[0,\infty)$.

Consider now for fixed $t>0$ and $\omega\in\Omega'\cap\Omega''$

$$V_t(s)=\frac{V_{\sigma(ts)}}{a_{\nu t}}\,,\quad s\geq 0\,.$$

From (17) we may expect that $V_t(s)\approx V'_{t\nu}(s)$. Indeed, for any fixed $0<\delta<1$ and all large t, we know that $\lfloor tc\nu(1-\delta)\rfloor\leq\sigma(tc)\leq\lfloor tc\nu(1+\delta)\rfloor$, which by monotonicity of the sums implies

$$\frac{V_{\lfloor tc\nu(1-\delta)\rfloor}}{a_{\nu t}}\leq\frac{V_{\sigma(tc)}}{a_{\nu t}}\leq\frac{V_{\lfloor tc\nu(1+\delta)\rfloor}}{a_{\nu t}}\,.$$

Now, for $c(1-\delta)$ and $c(1+\delta)$ in B, the left hand side and the right hand side above converge to $G_\alpha(c(1-\delta))$ and $G_\alpha(c(1+\delta))$. Thus, if we consider $c\in B$ and let $\delta\to 0$, then

$$\frac{V_{\sigma(tc)}}{a_{\nu t}}\to G_\alpha(c)\,,\qquad t\to\infty, \tag{18}$$

for all $\omega\in\Omega'\cap\Omega''$ and thus with probability 1.

By Theorem 2.15 in Chapter VI in Jacod and Shiryaev [2003], to prove (13), it remains to show that for all $\omega\in\Omega'\cap\Omega''$ and $c\in B$, as $t\to\infty$

$$\sum_{0<s\leq c}|\Delta V_t(s)|^2=\sum_{i<\sigma(tc)}\left(\frac{Y_i}{a_{t\nu}}\right)^2\to\sum_{0<s\leq c}|\Delta G_\alpha(s)|^2\,, \tag{19}$$

where, for an arbitrary càdlàg process $X(t)$ at time $t\geq 0$, we denote $\Delta X(t)=X_t-X_{t-}$. Observe that

$$G_{\alpha/2}(c):=\sum_{0<s\leq c}|\Delta G_\alpha(s)|^2$$

defines an $\alpha/2$–stable subordinator and that the squared random variables Y_i^2 are again regularly varying with index $\alpha/2$ with the property that $n\mathbb{P}(Y_i^2>a_n^2)\to 1$. A similar approximation argument as for (18) shows that (19) indeed holds, which concludes the proof. $\square$

Assume now that $P(D > x) = x^{-\alpha}\ell(x)$ for some slowly varying function ℓ and $\alpha \in (0,1)$. Select a sequence $a_n \to \infty$ such that $nP(D > a_n) \to 1$, as $n \to \infty$. Under suitable conditions on the residue term ε_t we obtain the following.

Proposition 4.2. *Assume that D_i's are regularly varying with index $\alpha \in (0,1)$ and that $\varepsilon_t = o_P(a_t)$. Then, there exists an α–stable random variable G_α such that*

$$\frac{S(t)}{a_{\nu t}} \xrightarrow{d} G_\alpha \,, \tag{20}$$

as $t \to \infty$.

Proof. The proof follows roughly the same lines as the proof of Proposition 4.1, but here we rely on an application of the previous theorem to the random walks (Γ_n) and (S_n^D). Just, instead of Y_i's and W_i's we have D_i's and an independent sequence of i.i.d. exponential random variables with parameter ν. □

Remark 4.1. *One can consider total claim amount in the period $[0,t]$ for the stationary model of subsection 2.1.4, i.e.*

$$S^*(t) = \int_0^t \int_{\mathbb{S}} f(u) N^*(ds, du) \,, \qquad t \geq 0 \,.$$

Here again, $S^*(t)$ has a similar representation as in (9) but with an additional term on the right hand side, i.e.

$$S^*(t) = \sum_{i=1}^{\tau(t)} D_i - D_{\tau(t)} - \varepsilon_t + \varepsilon^*_{0,t} \,, \qquad t \geq 0 \,, \tag{21}$$

where

$$\varepsilon^*_{0,t} = \sum_{\Gamma_i \leq 0,\ 0 < \Gamma_i + T_{ij} < t} X_{ij}.$$

Clearly, by stationarity

$$\varepsilon_t = \sum_{0 \leq \Gamma_i \leq t, t < \Gamma_i + T_{ij}} X_{ij} \stackrel{d}{=} \varepsilon_t^- = \sum_{-t \leq \Gamma_i \leq 0, 0 < \Gamma_i + T_{ij}} X_{ij}. \tag{22}$$

Hence, $\varepsilon_t = o_P(a_t)$ yields $\varepsilon_t^- = o_P(a_t)$ for any sequence (a_t) and therefore

$$\varepsilon^*_{0,t} \leq \varepsilon_t^- + \sum_{\Gamma_i < -t,\ 0 < \Gamma_i + T_{ij} < t} X_{ij} = \sum_{\Gamma_i < -t,\ 0 < \Gamma_i + T_{ij} < t} X_{ij} + o_P(a_t) \,.$$

In particular, conclusions of propositions 3.1, 4.1 and 4.2 hold for random variables $S^*(t)$ too under the additional assumption that

$$\tilde{\varepsilon}_t := \sum_{\Gamma_i \leq -t,\ 0<\Gamma_i+T_{ij}<t} X_{ij} = o_P(a_t). \tag{23}$$

5. Total claim amount for special models

As we have seen in the previous two sections, it is relatively easy to describe asymptotic behavior of the total claim amount $S(t)$ as long as we are able to determine the moments and tail properties of the random variables D_i and the residue random variable ε_t in (9) (and also $\tilde{\varepsilon}_t$ in (23) for the stationary version). However, this is typically a rather technical task, highly dependent on an individual Poisson cluster model. In this section we revisit three models introduced in Subsection 2.1, characterizing for each of them the limiting distribution of the total claim amount under appropriate conditions. Note that the cluster sum D for all three models admits the following representation

$$D \stackrel{d}{=} \sum_{j=0}^{K} X_j,$$

for $(X_j)_{j\geq 0}$ i.i.d. copies of $f(A)$ and some integer valued K such that $\mathbb{E}[K_1] < \infty$. Throughout, we assume that the random variables K and $(X_j)_{j\geq 1}$ are independent. The sum $\sum_{j=1}^{K} X_j$ has a so called compound distribution. Its first two moments exist under the following conditions

- if $\mathbb{E}[X] < \infty$ and $\mathbb{E}[K] < \infty$, then $\mu_D = \mathbb{E}D = (1 + \mathbb{E}[K])\mathbb{E}[X] < \infty$,
- if $\mathbb{E}[X^2] < \infty$ and $\mathbb{E}[K^2] < \infty$, then $\mathbb{E}D^2 = (\mathbb{E}[K]+1)\mathbb{E}[X^2]+(\mathbb{E}[K^2]+\mathbb{E}[K])\mathbb{E}[X]^2 < \infty$.

The tail behavior of compound sums was often studied under various conditions (see Robert and Segers [2008], Faÿ et al. [2006], Hult and Samorodnitsky [2008], Denisov et al. [2010]). We list below some of these conditions, which are applicable to our setting.

(RV1) If X is regularly varying with index $\alpha > 0$ and $\mathbb{P}(K > x) = o(\mathbb{P}(X > x))$, then $\mathbb{P}(D > x) \sim (\mathbb{E}[K] + 1)\mathbb{P}(X > x)$ as $x \to \infty$, see [Faÿ et al. , 2006, Proposition 4.1],

(RV2) If K is regularly varying with index $\alpha \in (1,2)$ and $\mathbb{P}(X > x) = o(\mathbb{P}(K > x))$, then $\mathbb{P}(D > x) \sim \mathbb{P}(K > x/\mathbb{E}[X])$ as $x \to \infty$, see [Robert and Segers , 2008, Theorem 3.2] or [Faÿ et al. , 2006, Proposition 4.3],

(RV3) If X and K are both regularly varying with index $\alpha \in (1,2)$ and tail equivalent, see [Embrechts et al. , 1997, Definition 3.3.3], then $\mathbb{P}(D > x) \sim (\mathbb{E}[K]+1)\mathbb{P}(X > x) + \mathbb{P}(K > x/\mathbb{E}[X])$ as $x \to \infty$, [Denisov et al. , 2010, Theorem 7].

We will refer to the last three conditions as the sufficient conditions **(RV)**.

5.1. Mixed binomial cluster model

Assume that $(K_i, (W_{ij})_{j\geq 1}, (A_{ij})_{j\geq 0})_{i\geq 0}$ constitutes an i.i.d. sequence with the following properties

- $(A_{ij})_{j\geq 0}$ are i.i.d. for any fixed i ,
- $(A_{ij})_{j\geq 1}$ is independent of $K_i, (W_{ij})_{j\geq 1}$ for all $i \geq 0$,
- $(W_{ij})_{j\geq 1}$ are conditionally i.i.d. and independent of K_i given A_{i0}.

Thus we do not exclude the possibility of dependence between $K_i, (W_{ij})_{j\geq 1}$ and the ancestral mark A_{i0}. For any $\gamma > 0$, we denote by

$$A,\ X_j,\ K, W_j,\ m_A\,, m_A(\gamma)\,,$$

generic random variables with the same distribution as A_{ij}, $X_{ij} = f(A_{ij})$, K_i, W_{ij}, $\mathbb{E}[K_i \mid A_{i0}]$ and $\mathbb{E}[K_i^\gamma \mid A_{i0}]$ respectively. Using the cluster representation, one can derive the asymptotic properties of $S(t)$. Let us first consider the Gaussian CLT under appropriate 2nd moment assumptions. Denote by $\mathbb{P}(W \in \cdot \mid A)$ the distribution of W_{ij}'s given A_{i0}.

Corollary 5.1. *Assume that* $\mathbb{E}[X^2] < \infty$ *and* $\mathbb{E}[K^2] < \infty$*. If*

$$\sqrt{t}\mathbb{E}[m_A \mathbb{P}(W > t \mid A)] \to 0, \qquad t \to \infty, \tag{24}$$

then the relation (10) *holds.*

Observe that (24) is slightly weaker than the existence of the moment $\mathbb{E}[K\sqrt{W}] < \infty$.

Proof. By Proposition 3.1, it follows that $\mathbb{E}D^2 < \infty$ as soon as $\mathbb{E}[X^2] < \infty$ and $\mathbb{E}[K^2] < \infty$. It remains to show that $\varepsilon_t = o_P(\sqrt{t})$. In order to do so, we use the Markov inequality

$$\mathbb{P}(\varepsilon_t > \sqrt{t}) \le \frac{\mathbb{E}[\varepsilon_t]}{\sqrt{t}} = \frac{\mathbb{E}\left[\sum_{0\le\Gamma_i\le t}\sum_{j=1}^{K_i}\mathbb{I}_{t<\Gamma_i+W_{ij}}f(A_{ij})\right]}{\sqrt{t}}.$$

We use Lemma 7.2.12 of Mikosch [2009] in order to compute the r.h.s. term as

$$\begin{aligned}\frac{\int_0^t \mathbb{E}\left[\sum_{j=1}^{K_i}\mathbb{I}_{W_{ij}>t-s}f(A_{ij})\right]\nu ds}{\sqrt{t}} &= \frac{\nu\mathbb{E}[X]\int_0^t \mathbb{E}\big[\mathbb{E}\big[\sum_{j=1}^{K_i}\mathbb{I}_{W_{ij}>t-s} \mid A_{i0}\big]\big]ds}{\sqrt{t}} \\ &= \frac{\nu\mathbb{E}[X]\int_0^t \mathbb{E}[m_A\mathbb{P}(W > x \mid A)]dx}{\sqrt{t}}.\end{aligned}$$

Notice that the last identity is obtained thanks to the independence of K_i and $(W_{ij})_{j\ge0}$ conditionally on A_{i0}. We conclude by the L'Hôpital's rule that this converges to 0 under (24). □

For the regularly varying D of order $1 < \alpha < 2$, we obtain the corresponding limit theorem under weaker assumptions on the tail of the waiting time W.

Corollary 5.2. *Assume that one of the conditions* **(RV)** *holds for* $1 < \alpha < 2$, *so that* D *is regularly varying. When*

$$t^{1+\delta-1/\alpha}\mathbb{E}[m_A\mathbb{P}(W > t \mid A)] \to 0, \qquad t \to \infty, \tag{25}$$

for some $\delta > 0$ *the relation* (12) *holds.*

The condition (25) is slightly weaker than assuming $\mathbb{E}[m_A W^{1+\delta-1/\alpha}] < \infty$. Notice that when $\alpha \to 1^+$ and K is independent of W, this condition boils down to the existence of an r'th moment of W for any strictly positive r.

Proof. By definition, (a_t) satisfies $t\mathbb{P}(D > a_t) \to 1$ as $t \to \infty$ and (a_t) is regularly varying with index $1/\alpha$. Applying the Markov inequality as in the proof of Corollary 5.1, we obtain

$$\mathbb{P}(\varepsilon_t > a_t) \le \frac{\mathbb{E}[\varepsilon_t]}{a_t} = \frac{\nu\mathbb{E}[X]\int_0^t \mathbb{E}[m_A\mathbb{P}(W > s \mid A)]ds}{a_t}.$$

The claim follows now by the L'Hôpital's rule and the relation $t^{1/\alpha-\delta} = o(a_t)$ for any $\delta > 0$. □

Remark 5.1. *In the context of the mixed binomial model, consider the total claim amount of the stationary process denoted by $S^*(t)$ which takes into account also the arrivals in the interval $(-\infty, 0)$, see Remark 4.1. Assume for simplicity that K_i's and (W_{ij})'s are unconditionally independent. Then $\tilde{\varepsilon}_t$ from (23) is $o_P(a_t)$ under the same conditions as in Corollaries 5.1 and 5.2, where we set $a_t = \sqrt{t}$ in the former case. Indeed, we will show that*

$$\mathbb{E}\tilde{\varepsilon}_t = \mathbb{E}\left(\sum_{\Gamma_i \le -t,\ 0<\Gamma_i+W_{ij}<t} X_{ij}\right) = \mathbb{E}\left(\sum_{-t\le\Gamma_i\le 0, t<\Gamma_i+W_{ij}} X_{ij}\right),$$

so that $\mathbb{E}\tilde{\varepsilon}_t = o_P(a_t)$ as well since the r.h.s. is dominated by $\mathbb{E}\varepsilon_t^- = o(a_t)$, cf. (22).

Note first that under assumption of the last two corollaries, individual claims have finite expectation, i.e. $\mathbb{E}X < \infty$. So it suffices to show that

$$I_1 := \mathbb{E}\tilde{\varepsilon}_t/\mathbb{E}X = \mathbb{E}\sum_{\Gamma_i<-t}\sum_{j=1}^{K_i}\mathbb{I}_{0<\Gamma_i+W_{ij}<t} = \mathbb{E}\sum_{-t<\Gamma_i<0}\sum_{j=1}^{K_i}\mathbb{I}_{\Gamma_i+W_{ij}>t} =: I_2\,.$$

From I_1, I_2 we subtract respectively l.h.s. and r.h.s. of the equality

$$\mathbb{E}\sum_{-2t<\Gamma_i<-t}\sum_{j=1}^{K_i}\mathbb{I}_{\Gamma_i+W_{ij}\in(0,t)}\mathbb{I}_{W_{ij}\in(t,2t]} = \mathbb{E}\sum_{-t<\Gamma_i<0}\sum_{j=1}^{K_i}\mathbb{I}_{\Gamma_i+W_{ij}\in(t,2t)}\mathbb{I}_{W_{ij}\in(t,2t]}\,,$$

where the equality follows by the stationarity of the underlying Poisson process, to obtain

$$J_1 = \mathbb{E}\sum_{-\infty<\Gamma_i<-t}\sum_{j=1}^{K_i}\mathbb{I}_{0<\Gamma_i+W_{ij}<t}\mathbb{I}_{W_{ij}>2t} = \mathbb{E}K\int_{-\infty}^{-t}\nu ds\int_{-s\vee 2t}^{t-s} dF_W(u)$$

and

$$J_2 = \mathbb{E}\sum_{-t<\Gamma_i<0}\sum_{j=1}^{K_i}\mathbb{I}_{\Gamma_i+W_{ij}>t}\mathbb{I}_{W_{ij}>2t} = \mathbb{E}K\int_{-t}^{0}\nu ds\int_{2t}^{\infty} dF_W(u)$$

where F_W denotes the distribution function of delays (W_{ij}). Finally, note that

$$J_1 = \mathbb{E}K\int_{2t}^{\infty} dF_W(u)\int_{-u}^{t-u}\nu ds = J_2\,.$$

Since we assumed that $\mathbb{E}[K] < \infty$, the regular variation property of D with index $\alpha \in (0,1)$ can arise only through the claim size distribution, see Proposition 4.8 in Faÿ et al. [2006]. It turns out that in such a heavy tailed case, no additional assumption on the waiting time W is needed.

Corollary 5.3. *Assume that X is regularly varying of order $0 < \alpha < 1$, then the relation (20) holds.*

Proof. Observe that one cannot apply Markov inequality anymore because $\mathbb{E}D = \infty$. Instead, we use the fact that $\sum_{j=1}^{t} X_j/a_t$ converges because X and D have equivalent regular varying tails. Recall from (22) that

$$\varepsilon_t \stackrel{d}{=} \varepsilon_t^- = \sum_{-t \le \Gamma_i \le 0, 0 < \Gamma_i + T_{ij}} X_{ij}.$$

We denote the (increasing) number of summands in the r.h.s. term by $M_t = \#\{i,j : -t \le \Gamma_i \le 0, -\Gamma_i < T_{ij}\}$. We can apply Proposition 4.2 after observing that $\sum_{j=1}^{M_t} X_j/a_{M_t}$ is a tight family of random variables, because M_t is independent of the array (X_{ij}). Writing

$$\frac{\varepsilon_t}{a_t} \stackrel{d}{=} \frac{\sum_{j=1}^{M_t} X_j}{a_{M_t}} \frac{a_{M_t}}{a_t}, \tag{26}$$

and observing that a_t is regularly varying with index $1/\alpha$, we obtain the desired result provided that $M_t = o_P(t)$. It is sufficient to show the convergence to 0 of the ratio

$$\frac{\mathbb{E}[M_t]}{t} = \frac{\mathbb{E}[\#\{i,j : -t \le \Gamma_i \le 0, -\Gamma_i < T_{ij}\}]}{t} = \frac{\mathbb{E}\left[\sum_{0 \le \Gamma_i \le t} \sum_{j=1}^{K_i} \mathbb{I}_{t \le \Gamma_i + W_{ij}}\right]}{t}.$$

Using similar calculation as in the proof of Corollary 5.1 (setting $X = 1$), we obtain an explicit formula for the r.h.s. term as

$$\frac{\nu \int_0^t \mathbb{E}[m_A \mathbb{P}(W > x \mid A)] dx}{t} \to 0, \qquad t \to \infty,$$

the convergence to 0 following from a Cesarò argument. $\square$

5.2. Renewal cluster model

Starting with the same assumption and notation as for the binomial cluster model of the previous subsection, the total claim amount coming from the ith immigrant and its progeny is again

$$D \stackrel{d}{=} \sum_{j=0}^{K} X_j,$$

for (X_j) i.i.d. copies of $f(A)$. Dealing with the waiting times

$$T_{ij} = W_{i1} + \cdots + W_{ij}$$

requires additional care here. We obtain first

Corollary 5.4. *Suppose* $\mathbb{E}X^2 < \infty$ *and* $\mathbb{E}K^2 < \infty$ *then, when* $\mathbb{E}[m_A(2)W^\delta] < \infty$ *for some* $\delta > 1/2$*, the relation* (10) *holds.*

Proof. The proof follows from Theorem 5.1. Second moment of D_i's is finite by the moment assumptions on X and K. It remains to show that the residue term satisfies $\varepsilon_t = o_P(\sqrt{t})$. Using Lemma 7.2.12 of Mikosch [2009] as in the proof of Corollary 5.1 we obtain

$$\begin{aligned}
\mathbb{E}[\varepsilon_t] &= \int_0^t \mathbb{E}\left[\sum_{j=1}^{K_i} \mathbb{I}_{W_{i1}+\cdots+W_{ij}>x} f(A_{ij})\right] \nu dx \\
&= \nu \int_0^t \mathbb{E}\left[\mathbb{E}\left[\sum_{j=1}^{K_i} \mathbb{I}_{W_{i1}+\cdots+W_{ij}>x} f(A_{ij}) \mid K_i, (W_{ij})_{j\geq 1}\right]\right] dx \\
&= \nu \mathbb{E}[X] \int_0^t \mathbb{E}\left[\sum_{j=1}^{K_i} \mathbb{I}_{W_{i1}+\cdots+W_{ij}>x}\right] dx
\end{aligned}$$

by independence between $K_i, (W_{ij})_{j\geq 1}$ and $(f(A_{ij}))_{j\geq 1}$. The key argument in dealing with the renewal cluster model is the following upper bound

$$\sum_{j=1}^{K_i} \mathbb{I}_{W_{i1}+\cdots+W_{ij}>x} \leq \mathbb{I}_{W_{i1}+\cdots+W_{iK_i}>x} K_i \,. \tag{27}$$

Assume with no loss of generality that $\delta \leq 1$. By the Markov inequality and the conditional independence of K_i and $(W_{ij})_{j\geq 0}$ conditionally on A_{i0}, we obtain

$$\mathbb{E}\left[\mathbb{I}_{W_{i1}+\cdots+W_{iK_i}>x}K_i \mid A_{i0}\right] \leq \frac{\mathbb{E}[K_i(W_{i1}+\cdots+W_{iK_i})^\delta \mid A_{i0}]}{x^\delta}$$
$$\leq m_{A_{i0}}(2)\frac{\mathbb{E}[W^\delta \mid A_{i0}]}{x^\delta}. \qquad (28)$$

The last inequality follows from the sub-linearity of the mapping $x \mapsto x^\delta$ for $\delta \leq 1$. Thus, we obtain for some constant $C > 0$

$$\mathbb{E}[\varepsilon_t] \leq \nu\mathbb{E}[X]\mathbb{E}[m_A(2)W^\delta]\int_1^t x^{-\delta}dx + C = O(\mathbb{E}[m_A(2)W^\delta]t^{1-\delta}) = o(\sqrt{t})$$

as $\delta > 1/2$ by assumption. □

Regularly varying claims can be handled with additional care as K may not be square integrable.

Corollary 5.5. *Assume that one of the conditions* **(RV)** *holds so that D is regularly varying of order $1 < \alpha < 2$ and that $\mathbb{E}K^{1+\gamma} < \infty$ for some $\gamma > 0$. If $\mathbb{E}[m_A(1+\gamma)W^\delta] < \infty$ for some $\delta > (\alpha-\gamma)/\alpha$, then the relation* (12) *holds.*

Observe that we obtain somewhat stronger conditions than in the mixed binomial case, see Corollary 5.2 and remark following it.

Proof. With no loss of generality we assume that $\gamma \leq 1$. We use the Markov inequality of order γ

$$\mathbb{P}(\varepsilon_t > a_t) \leq \frac{\mathbb{E}[\varepsilon_t^\gamma]}{a_t^\gamma}.$$

Thanks to the the sub-additivity of the function $x \mapsto x^\gamma$ we have

$$\mathbb{E}[\varepsilon_t^\gamma] = \mathbb{E}\left[\left(\sum_{-t\leq\Gamma_i\leq 0}\sum_{j=1}^{K_i}\mathbb{I}_{W_{i1}+\cdots+W_{ij}>0}f(A_{ij})\right)^\gamma\right]$$
$$\leq \mathbb{E}\left[\sum_{-t\leq\Gamma_i\leq 0}\left(\sum_{j=1}^{K_i}\mathbb{I}_{W_{i1}+\cdots+W_{ij}>0}f(A_{ij})\right)^\gamma\right]$$

Using Lemma 7.2.12 of Mikosch [2009] as in the proof of Corollary 5.1 we obtain

$$\mathbb{E}[\varepsilon_t^\gamma] \leq \int_0^t \mathbb{E}\left[\left(\sum_{j=1}^{K_i}\mathbb{I}_{W_{i1}+\cdots+W_{ij}>x}f(A_{ij})\right)^\gamma\right]\nu dx. \qquad (29)$$

We use Jensen's inequality as follows

$$\mathbb{E}\left[\mathbb{E}\left[\left(\sum_{j=1}^{K_i}\mathbb{I}_{W_{i1}+\cdots+W_{ij}>x}f(A_{ij})\right)^{\gamma}\mid K_i,(W_{ij})_{j\geq 1}\right]\right]$$
$$\leq \mathbb{E}\left[\left(\mathbb{E}\left[\sum_{j=1}^{K_i}\mathbb{I}_{W_{i1}+\cdots+W_{ij}>x}f(A_{ij})\mid K_i,(W_{ij})_{j\geq 1}\right]\right)^{\gamma}\right]$$

so that, using the independence between $K_i,(W_{ij})_{j\geq 1}$ and $(f(A_{ij}))_{j\geq 1}$, one gets

$$\mathbb{E}[\varepsilon_t^{\gamma}]\leq \nu\mathbb{E}[X]^{\gamma}\int_0^t\mathbb{E}\left[\left(\sum_{j=1}^{K_i}\mathbb{I}_{W_{i1}+\cdots+W_{ij}>x}\right)^{\gamma}\right]dx$$

Using the stochastic domination (27), we obtain

$$\mathbb{E}[\varepsilon_t^{\gamma}]\leq \nu\mathbb{E}[X]^{\gamma}\int_0^t\mathbb{E}\left[\mathbb{I}_{W_{i1}+\cdots+W_{iK_i}>x}K_i^{\gamma}\right].$$

With no loss of generality we assume $0<\delta<1$. Applying the Markov inequality of order δ conditionally on A_{i0} as in (28), we have

$$\mathbb{E}[\mathbb{I}_{W_{i1}+\cdots+W_{iK_i}>x}K_i^{\gamma}\mid A_{i0}]\leq m_{A_{i0}}(1+\gamma)\frac{\mathbb{E}[W^{\delta}\mid A_{i0}]}{x^{\delta}}.$$

Plugging in this bound in the previous inequality, we obtain for some $C>0$,

$$\mathbb{E}[\varepsilon_t^{\gamma}]\leq \nu\mathbb{E}[X]^{\gamma}\mathbb{E}\big[m_A(1+\gamma)W^{\delta}\big]t^{1-\delta}+C=o(a_t^{\gamma})$$

as $1-\delta<\gamma/\alpha$ by assumption. □

Corollary 5.6. *If X is regularly varying of order $\alpha\in(0,1)$ and $\mathbb{E}[W^{\delta}]<\infty$ for some $\delta>(\alpha-\gamma)_+/\alpha$, then the relation* (20) *holds.*

Proof. We use the same arguments as in the proof of Corollary 5.3 in order to obtain (26). The desired result follows if one can show that $M_t=o_P(t)$. Using the Markov's inequality, it is enough to check that $\mathbb{E}[M_t]/t=o(1)$. Following the same reasoning than in the proof of Corollary 5.5, we estimate the moment of M_t similarly as the one of ε_t in (29):

$$\mathbb{E}[M_t]\leq\int_0^t\mathbb{E}\left[\sum_{j=1}^{K_i}\mathbb{I}_{W_{i1}+\cdots+W_{ij}>x}\right]\nu dx\leq\int_0^t\mathbb{E}\left[K_i\mathbb{I}_{W_{i1}+\cdots+W_{iK_i}>x}\right]\nu dx.$$

We used the again the stochastic domination (27) to obtain the last upper bound. From a Cesaró argument, the result will follow if

$$\mathbb{E}\left[K_i \mathbb{I}_{W_{i1}+\cdots+W_{iK_i}>x}\right] \to 0, \qquad x \to \infty.$$

One can actually check this negligibility property because the random sequence $K_i \mathbb{I}_{W_{i1}+\cdots+W_{iK_i}>x} \to 0$ a.s. by finiteness of $W_{i1} + \cdots + W_{iK_i}$ and because the sequence is dominated by K_i that is integrable. $\square$

5.3. Marked Hawkes process

In general, it is not entirely straightforward to see when the moments of D are finite. However, note that D_i's are i.i.d. and satisfy distributional equation

$$D \stackrel{d}{=} f(A) + \sum_{j=1}^{L_A} D_j \,, \tag{30}$$

where L_A has the Poisson distribution conditionally on A, with mean $m_A = \int_0^\infty h(s, A)ds$ and D_j's on the right hand side are independent of m_A and i.i.d. with the same distribution as D. Conditionally on A, the waiting times are i.i.d. with common density $h(t, A)/m_A$, $t \geq 0$. Thus, one can relate the clusters of the Hawkes process with those of a mixed binomial process from Section 5.1 with $K = L_A$. In order to obtain the asymptotic properties of $S(t)$ one still needs to characterize the moment and tail properties of D.

Consider the Laplace transform of D, i.e. $\varphi(s) = \mathbb{E}e^{-sD}$, for $s \geq 0$. Note, φ is an infinitely differentiable function for $s > 0$. To simplify the notation, denote by

$$X = f(A) \,,$$

a generic claim size and observe that by (30), φ satisfies the following

$$\begin{aligned} \varphi(s) &= \mathbb{E}\left[\mathbb{E}\left(e^{-s(X+\sum_{j=1}^{L_A} D_j)} \,\middle|\, A\right)\right] = \mathbb{E}\left[e^{-sX}\mathbb{E}\left(e^{-s\sum_{j=1}^{L_A} D_j} \,\middle|\, A\right)\right] \\ &= \mathbb{E}\left[e^{-sX} e^{m_A(\mathbb{E}e^{-sD}-1)}\right] = \mathbb{E}\left[e^{-sX} e^{m_A(\varphi(s)-1)}\right] . \end{aligned} \tag{31}$$

When $\mathbb{E}[m_A] = \kappa < 1$, it is known that this functional equation has a unique solution φ which further uniquely determines the distribution of D. By studying the behavior of the derivatives of φ for $s \to 0+$, we get the following result.

Lemma 5.1. *If* $\mathbb{E}X^2 < \infty$ *and* $\mathbb{E}m_A^2 < \infty$ *then*

$$\mathbb{E}D^2 = \frac{\mathbb{E}X^2}{1-\kappa} + \frac{(\mathbb{E}X)^2}{(1-\kappa)^3}\mathbb{E}m_A^2 + 2\frac{\mathbb{E}X}{(1-\kappa)^2}\mathbb{E}(f(A)m_A) < \infty\,.$$

Notice that this expression coincides with the expression in Karabash and Zhu [2015], when $X = f(A) \equiv 1$, i.e. in the case when one simply counts the number of claims.

Proof. Differentiating the equation (31) with respect to $s > 0$ produces

$$\varphi'(s) = \mathbb{E}\left[e^{-sX}e^{m_A(\varphi(s)-1)}\left(-X + m_A\varphi'(s)\right)\right]\,.$$

As $\mathbb{E}(m_A) = \kappa < 1$ we obtain

$$\varphi'(s) = \frac{-\mathbb{E}\left[e^{-sX}e^{m_A(\varphi(s)-1)}X\right]}{1-\mathbb{E}[e^{-sX}e^{m_A(\varphi(s)-1)}m_A]} \tag{32}$$

As $\varphi(s) \leq 1$, $s \geq 0$, the integrand in the numerator is dominated by X and the one in the denominator by m_A. By the dominated convergence argument, $\lim_{s\to 0+}\varphi'(s)$ exists and is equal to

$$\varphi'(0) = \frac{-\mathbb{E}X}{1-\kappa}\,.$$

In particular $\mathbb{E}D = \mathbb{E}X/(1-\kappa)$. Differentiating (31) again produces second moment of D. Indeed, we have

$$\varphi''(s) = \mathbb{E}\left[e^{-sX}e^{m_A(\varphi(s)-1)}\left(\left(-X + m_A\varphi'(s)\right)^2 + m_A\varphi''(s)\right)\right]\,,$$

so that

$$\varphi''(s) = \frac{\mathbb{E}\left[e^{-sX}e^{m_A(\varphi(s)-1)}\left(-X + m_A\varphi'(s)\right)^2\right]}{1-\mathbb{E}[e^{-sX}e^{m_A(\varphi(s)-1)}m_A]}. \tag{33}$$

Here again, applying the dominated convergence theorem twice, one can let $s \to 0+$ and obtain

$$\varphi''(0) = \frac{\mathbb{E}\left(-X + m_A\varphi'(0)\right)^2}{1-\kappa} = \frac{\mathbb{E}\left(X + m_A\mathbb{E}D\right)^2}{1-\kappa}\,.$$

Which concludes the proof since $X = f(A)$. □

The following theorem describes the behavior of the total claim amount $(S(t))$ for the marked Hawkes process under appropriate 2nd moment assumptions.

Theorem 5.1. *If $\kappa < 1$, $\mathbb{E}X^2 < \infty$ and $\mathbb{E}[m_A^2] < \infty$ then, in either stationary or nonstationary case, if*

$$\sqrt{t}\mathbb{E}[\mu_A(t, \infty)]ds \to 0, \qquad t \to \infty, \tag{34}$$

then the relation (10) *holds.*

Proof. In order to apply Proposition 3.1 one has to check that $\varepsilon_t = o_P(\sqrt{t})$. The proof is based on the following domination argument on ε_t. Recall that one can write

$$N = \sum_i \sum_j \delta_{\Gamma_i + T_{ij}, A_{ij}} = \sum_{k=1}^{\infty} \delta_{\tau_k, A^k}\,,$$

w.l.o.g. assuming that $0 \le \tau_1 \le \tau_2 \le \dots$. At each time τ_j, a claim arrives generated by one of the previous claims or an entirely new (immigrant) claim appears. In the former case, if τ_j is a direct offspring of a claim at time τ_i, we will write $\tau_i \to \tau_j$. Progeny τ_j then creates potentially further claims. We denote by D_{τ_j} the total amount of claims generated by the arrival at τ_j (counting the claim at τ_j itself as well). Clearly, D_{τ_j}'s are identically distributed as D and even independent if we consider claims which are not offspring of one another. They are also independent of everything happening in the past.

The process N is naturally dominated by the stationary marked Hawkes process N^* which is well defined on the whole real line as we assumed $\kappa = \mathbb{E}m_A < 1$, see discussion at the end of Subsection 2.1. For the original and stationary Hawkes processes, N and N^*, by λ and λ^*, we denote corresponding predictable intensities. By the construction of these two point processes, $\lambda \le \lambda^*$. Recall that $\tau_i \to \tau_j$ is equivalent to $\tau_j = \tau_i + W_{ik}$, $k \le L^i = L_{A^i}$, where, by assumption, W_{ik} are i.i.d. with common density $h(t, A^i)/m_{A^i}$, $t \ge 0$, and independent of L^i conditionally on the mark A^i of the claim at τ_i. Moreover, conditionally on A^i, the number of direct progeny of the claim at τ_i, denoted by L^i, has Poisson distribution with parameter

μ_{A^i}. Therefore

$$\begin{aligned}
\mathbb{E}[\varepsilon_t] &= \mathbb{E}\Big[\sum_{\Gamma_i \le t}\sum_j \mathbb{I}_{\Gamma_i + T_{ij} > t} X_{ij}\Big] \\
&= \mathbb{E}\Big[\sum_{\tau_i \le t}\sum_{\tau_j > t} D_{\tau_j}\mathbb{I}_{\tau_i \to \tau_j}\Big] \\
&= \mathbb{E}\Big[\sum_{\tau_i \le t}\mathbb{E}\Big[\sum_{k=1}^{L^i} D_{\tau_i + W_{ik}}\mathbb{I}_{\tau_i + W_{ik} > t} \mid (\tau_i, A^i)_{i\ge 0}; \tau_i \le t\Big]\Big] \\
&= \mu_D \mathbb{E}\left[\int_0^t \int_{\mathbb{S}} \mu_a((t-s,\infty)) N(ds, da)\right],
\end{aligned}$$

where $\mu_a((u,\infty)) = \int_u^\infty h(s,a)ds$. Observe that from projection theorem, see Bremaud [1981], Chapter 8, Theorem 3, the last expression equals to

$$\mu_D \mathbb{E}\left[\int_0^t \int_{\mathbb{S}} \mu_a((t-s,\infty)) Q(da)\lambda(s)ds\right],$$

where we also used that the Ds are conditionally i.i.d. to simplify the notation. One can further bound this estimate by

$$\begin{aligned}
\mathbb{E}\left[\int_0^t \int_{\mathbb{S}} \mu_a((t-s,\infty)) Q(da)\lambda^*(s)ds\right] &= \int_0^t \int_{\mathbb{S}} \mu_a((t-s,\infty)) Q(da)\mathbb{E}[\lambda^*(s)]ds \\
&= \frac{\nu}{1-\kappa}\int_0^t \int_{\mathbb{S}} \mu_a((t-s,\infty)) Q(da)ds
\end{aligned}$$

Here we used Fubini's theorem, and the expression $\mathbb{E}\left[\lambda^*(s)\right] \equiv \nu/(1-\kappa)$. Observe that this expectation is constant since N^* is a stationary point process, to show that it equals $\nu/(1-\kappa)$, note that

$$\begin{aligned}
\mu^* = \mathbb{E}\lambda^*(s) = \mathbb{E}\left[\nu + \int_{-\infty}^s \int_{\mathbb{S}} h(s-u,a) N^*(du,da)\right] \\
= \quad \nu + \int_{-\infty}^s \int_{\mathbb{S}} h(s-u,a)\mathbb{E}(\lambda^*(u))du Q(da) \\
= \quad \nu + \mu^* \int_{-\infty}^s \mathbb{E}h(s-u,A)du \\
= \quad \nu + \mu^* \int_0^\infty \mathbb{E}h(v,A)dv,
\end{aligned}$$

see also Daley and Vere Jones [2003], Example 6.3(c). Hence, $\mu^* = \nu + \mu^* \cdot \kappa$ and $\mu^* = \nu/(1-\kappa)$ as we claimed above. Now, we have

$$\mathbb{E}\varepsilon_t \leq \frac{\nu}{1-\kappa} \int_0^t \int_{\mathbb{S}} \mu_a((t-s,\infty))Q(da)ds = \frac{\nu}{1-\kappa} \int_0^t \mu_D \int_s^\infty \mathbb{E}[h(u,A)]duds\,. \tag{35}$$

Hence the residual term is bounded in expectation by the expression we obtained in the mixed binomial case in Section 5.1. Thus, the result will follow from the proof of Corollary 5.1 under the condition (24) which is further equivalent to (34) thanks to the expression of the density of the waiting times.

Dividing the last expression by $\sqrt{t}$ and applying L'Hôpital's rule, proves the theorem for the nonstationary or pure Hawkes process, see Karabash and Zhu [2015] where the same idea appears in the proof of Theorem 1.3.2.

To show that the the central limit theorem holds in the stationary case, note that $S(t)$ now has a similar representation as in (9) but with an additional term on the right hand side, i.e.

$$S(t) = \sum_{i=1}^{\tau(t)} D_i - D_{\tau(t)} - \varepsilon_t + \varepsilon_{0,t}\,,\ t \geq 0\,, \tag{36}$$

where

$$\varepsilon_{0,t} = \sum_{\Gamma_i \leq 0,\ 0<\Gamma_i+T_{ij}<t} X_{ij}\,.$$

Similar computation provides

$$\begin{aligned} \mathbb{E}\varepsilon_{0,t} = \mathbb{E}\sum_{\Gamma_i\leq 0}\sum_j \mathbb{I}_{0<\Gamma_i+T_{ij}<t}X_{ij} &= \mathbb{E}\sum_{\tau_i\leq 0}\sum_{0<\tau_j<t} D_{\tau_j}\mathbb{I}_{\tau_i\to\tau_j} \\ &= \mathbb{E}\left[\sum_{\tau_i\leq 0}\mu_D\mathbb{E}\left(\sum_{0<\tau_j<t}\mathbb{I}_{\tau_i\to\tau_j}\middle|\,\mathcal{G}_0\right)\right] \\ &= \mu_D\mathbb{E}\left[\sum_{\tau_i\leq 0}\mu_{A^i}((0-\tau_i,t-\tau_i))\right] \\ &= \mu_D\mathbb{E}\left[\int_{-\infty}^0\int_{\mathbb{S}}\mu_a((-s,t-s))N^*(ds,da)\right]. \end{aligned}$$

where we denote $\mu_a(B) = \int_] Bh(s,a)ds$. Again, by the projection theorem, see Bremaud [1981], Chapter 8, Theorem 3, the last expression equals to

$$\mu_D \mathbb{E}\left[\int_{-\infty}^0 \int_{\mathbb{S}} \mu_a((-s,t-s))\lambda^*(s)dsQ(da)\right].$$

Which is further equal to

$$\begin{aligned}
&\mu_D \int_{-\infty}^0 \int_{\mathbb{S}} \mu_a((-s,t-s))\mathbb{E}\left[\lambda^*(s)\right] dsQ(da) \\
&= \mu_D \frac{\nu}{1-\kappa} \int_{-\infty}^0 \int_{\mathbb{S}} \mu_a((-s,t-s))dsQ(da) \\
&= \mu_D \frac{\nu}{1-\kappa} \int_{-\infty}^0 \mathbb{E}\mu_A((-s,t-s))ds \\
&= \mu_D \frac{\nu}{1-\kappa} \int_0^\infty \mathbb{E}\mu_A((s,s+t))ds \\
&= \mu_D \frac{\nu}{1-\kappa} \int_0^\infty \mathbb{E}\int_s^{s+t} h(u,A)duds \\
&= \mu_D \frac{\nu}{1-\kappa} \int_0^\infty \mathbb{E}(t\wedge u)h(u,A)du \\
&= \mu_D \frac{\nu}{1-\kappa} \left(\int_0^t \mathbb{E}[uh(u,A)]du + t\int_t^\infty \mathbb{E}[h(u,A)]du\right).
\end{aligned}$$

Notice that the second term in the last expression divided by $\sqrt{t}$ tends to 0 by (34). Using integration by parts for the first term, we have

$$\int_0^t \mathbb{E}[uh(u,A)]du = t\int_t^\infty \mathbb{E}[h(s,A)]ds + \int_0^t \int_u^\infty \mathbb{E}[h(s,A)]dsdu.$$

The first integral on the r.h.s. divided by $\sqrt{t}$ tends to 0 under (34). The last term divided by $\sqrt{t}$ also tends to 0 by an application of the L'Hôpital rule as in the non-stationary case.

Finally, we observe that $\varepsilon_{0,t}/\sqrt{t} \xrightarrow{P} 0$ and the result in the stationary case is proved. $\square$

Observe that (34) is substantially weaker than (8) in the unmarked case. Namely the former condition only requires that the total residue due to the claims on the compact interval $[0,t]$ is of the order $o(\sqrt{t})$ in probability. In particular, in the unmarked case, the central limit theorem holds for the stationary and the non-stationary case even if (8) is not satisfied, i.e. even when non-stationary process is not convergent. As we mentioned above, there are related limit theorems in the literature concerning only the counting process N_t, see Karabash and Zhu [2015], but in the contrast to their result, our proof does not rely on the martingale central limit theorem, it stems from rather simple relations (9) and (36).

In the following example, we consider some special cases of Hawkes processes for which a closed form expression for the 2nd moment $\mathbb{E}D^2$ can be found.

Example 5.1. *(Marked Hawkes processes with claims independent of the cluster size) Assume that the random measure*

$$\mu_A(B) = \int_B h(s,A)ds\,, \tag{37}$$

on $\mathbb{R}_+$ and the corresponding claim size $X = f(A)$ are independent. In particular, this holds if $\mu_A(B) = \int_B h(s)ds$, for some integrable function h, i.e. when μ_A is a deterministic measure and we actually have standard Hawkes process with independent marks. In this special case $K+1$ is known to have the so–called Borel distribution, see Haight and Breuer [1960].

Using the arguments from the proof of Lemma 5.1, one obtains $\mu_D = \mathbb{E}D_i = \mathbb{E}X/(1-\kappa)$. Similarly the variance of D_i's is finite as the variance of a compound sum, and equals

$$\sigma_D^2 = \frac{\sigma_X^2}{1-\kappa} + \frac{\kappa(\mathbb{E}X)^2}{(1-\kappa)^3}\,,$$

cf. Lemma 2.3.4 in Mikosch [2009]. Hence $\mathbb{E}D^2$ in Theorem 5.1 has the form

$$\mathbb{E}D^2 = \sigma_D^2 + \mu_D^2 = \frac{\sigma_X^2}{1-\kappa} + \frac{(\mathbb{E}X)^2}{(1-\kappa)^3}..$$

In the special case, when the claims are all constant, say $X = f(A) \equiv c > 0$, direct calculation yields $\mathbb{E}D = c/(1-\kappa)$, with $\kappa = \mathbb{E}[m_A]$, and

$$\varphi''(0) = \frac{\mathbb{E}\left(-c + m_A\varphi'(0)\right)^2}{1-\kappa} = c^2\frac{\operatorname{Var} m_A + 1}{(1-\kappa)^3}\,,$$

obtaining

$$\mathbb{E}D^2 = \varphi''(0) = c^2 \frac{\operatorname{Var} m_A + 1}{(1-\kappa)^3}\,,$$

in particular, for $c = 1$ we recover expression in Karabash and Zhu [2015].

In the rest of this subsection, we study marked Hawkes process in the case when D_i's are regularly varying with index $\alpha < 2$. Using the result of Hult and Samorodnitsky [2008], one can show that when the individual claims $X = f(A)$ are regularly varying, this property is frequently passed on to the random variable D under appropriate moment assumptions on m_A. However, using the specific form of the Laplace transform for D given in (31), one can show regular variation of D under weaker conditions. This is the content of the following lemma.

Lemma 5.2. *Assume that $\kappa < 1$ and that $X = f(A)$ is regularly varying with index $\alpha \in (0,1) \cup (1,2)$. When $\alpha \in (1,2)$, assume additionally that $Y = X + m_A\mu_D$ is regularly varying of order α. Then the random variable D is regularly varying with the same index α.*

Proof. We will use Karamata's Tauberian Theorem, as formulated and proved in Theorem 8.1.6 of Bingham et. al [1987]. In particular, the equivalence between (8.1.12) and (8.1.11b) in Bingham et. al [1987] yields the following.

Theorem 5.2. *The nonnegative random variable X is regularly varying with a noninteger tail index $\alpha > 0$, i.e. $\bar{F}(x) \sim x^{-\alpha}\ell(x)$ as $x \to \infty$ if and only if*

$$\varphi^{(\lceil\alpha\rceil)}(s) \sim cs^{\alpha-\lceil\alpha\rceil}\ell(1/s), \qquad s \to 0+,$$

for some slowly varying function ℓ and a constant depending only on α: $c = -\Gamma(\alpha+1)\Gamma(1-\alpha)/\Gamma(\alpha - \lfloor\alpha\rfloor)$.

Consider first the case $0 < \alpha < 1$. By differentiating once the expression for the Laplace transform, we obtain the identity (32)

$$\varphi'(s) = \frac{-\mathbb{E}\left[e^{-sX}e^{m_A(\varphi(s)-1)}X\right]}{1-\mathbb{E}[e^{-sX}e^{m_A(\varphi(s)-1)}m_A]}, \qquad s > 0.$$

We are interested in the behavior of this derivative as $s \to 0+$. Using the inequality $|1-e^{-x}| = 1-e^{-x} \le x$, we have

$$\left|\varphi'(s) - \frac{-\mathbb{E}\left[e^{-sX}X\right]}{1-\mathbb{E}[e^{-sX-m_A(1-\varphi(s))}m_A]}\right| \le \frac{\mathbb{E}\left[m_A(1-\varphi(s))e^{-sX}X\right]}{1-\mathbb{E}[e^{-sX-m_A(1-\varphi(s))}m_A]}$$
$$\le \frac{1-\varphi(s)}{s}\frac{\mathbb{E}\left[m_A e^{-sX}sX\right]}{1-\mathbb{E}[e^{-sX-m_A(1-\varphi(s))}m_A]}.$$

As $e^{-sX}sX \le e^{-1}$, we prove that $\mathbb{E}\left[m_A e^{-sX}sX\right] = o(1)$ as $s \to 0+$ by dominated convergence. Moreover, using again $1-e^{-x} \le x$ and denoting $\varphi_X(s) = \mathbb{E}[e^{-sX}]$ the Laplace transform of X, we have

$$0 \le \varphi_X(s) - \varphi(s) \le \mathbb{E}[e^{-sX}m_A(1-\varphi(s))] \le \kappa(1-\varphi(s))$$

so that

$$1-\varphi(s) \le \frac{1}{1-\kappa}(1-\varphi_X(s)).$$

Collecting all those bounds and using the identity $|\varphi_X'(s)| = \mathbb{E}[e^{-sX}X]$, we obtain

$$\left|\varphi'(s) - \frac{\varphi_X'(s)}{1-\mathbb{E}[e^{-sX}e^{m_A(\varphi(s)-1)}m_A]}\right| = o\left(\frac{1-\varphi_X(s)}{s}\right), \qquad s \to 0^+. \quad (38)$$

The regular variation of the random variable D follows now by two consecutive applications of Theorem 5.2. First, as X is regularly varying of order $0<\alpha<1$, applying the direct part of the equivalence in Theorem 5.2 we obtain

$$\varphi_X'(s) \sim cs^{\alpha-1}\ell(1/s), \qquad s \to 0^+.$$

Applying Karamata's theorem again, i.e. the equivalence between (8.1.9) and (8.1.11b) in [Bingham et. al , 1987, Theorem 8.1.6], one can show that that $(1-\varphi_X(s))/s = O(\varphi_X'(s))$ as $s \to 0+$. Using (38) and the limiting relation $\mathbb{E}[e^{-sX}e^{m_A(\varphi(s)-1)}m_A] \to \kappa$ as $s \to 0^+$, we obtain

$$\varphi'(s) \sim \frac{\varphi_X'(s)}{1-\kappa} \sim \frac{cs^{\alpha-1}\ell(1/s)}{1-\kappa}, \qquad s \to 0+.$$

Finally, applying the reverse part of Theorem 5.2, we obtain

$$\bar{F}_D(x) \sim \frac{\ell(x)x^{-\alpha}}{1-\kappa} = \frac{\bar{F}_X(x)}{1-\kappa}, \qquad x \to \infty.$$

The case $1 < \alpha < 2$ can be treated similarly, under the additional assumption that $Y = X + m_A \mu_D$ is regularly varying. We will again show that $P(D > x) \sim (1-\kappa)^{-1} P(Y > x)$ as $x \to \infty$. To prove this equivalence, recall the identity (33)

$$\varphi''(s) = \frac{\mathbb{E}\left[e^{-sX} e^{m_A(\varphi(s)-1)} \left(-X + m_A \varphi'(s)\right)^2\right]}{1 - \mathbb{E}[e^{-sX} e^{m_A(\varphi(s)-1)} m_A]}.$$

As $\alpha > 1$, we have that $\mathbb{E}[Y] < \infty$ and thus $\mathbb{E}[X] < \infty$ and $E[D] = \mu_D = (1-\kappa)^{-1}\mathbb{E}[X]$ by an application of the tower property. Observe that, for any $s > 0$,

$$\begin{aligned}
&\left| \varphi''(s) - \frac{\mathbb{E}\left[e^{-sY} Y^2\right]}{1 - \mathbb{E}[e^{-sX - m_A(1-\varphi(s))} m_A]} \right| \\
&\qquad = \left| \frac{\mathbb{E}\left[e^{-sX - m_A(1-\varphi(s))} \left(-X + m_A \varphi'(s)\right)^2\right] - \mathbb{E}\left[e^{-sY} Y^2\right]}{1 - \mathbb{E}[e^{-sX - m_A(1-\varphi(s))} m_A]} \right|.
\end{aligned}$$

Let us decompose the numerator into two terms

$$\begin{aligned}
&\underbrace{\mathbb{E}\left| e^{-sX - m_A(1-\varphi(s))} \left(\left(-X + m_A \varphi'(s)\right)^2 - Y^2 \right) \right|}_{I_1} \\
&+ \underbrace{\mathbb{E}\left| \left(e^{-sY} - e^{-sX - m_A(1-\varphi(s))} \right) Y^2 \right|}_{I_2}.
\end{aligned}$$

Using the identity $a^2 - b^2 = (a-b)(a+b)$, I_1 is bounded by

$$\begin{aligned}
I_1 &\le (\mu_D + \varphi'(s)) \mathbb{E}\left[e^{-sX - m_A(1-\varphi(s))} m_A \left(2X + m_A(\mu_D - \varphi'(s))\right)\right] \\
&\le \frac{\mu_D + \varphi'(s)}{s} \left(\mathbb{E}\left[2 m_A e^{-sX} sX\right] + \mathbb{E}\left[m_A e^{-m_A(1-\varphi(s))} s(\mu_D - \varphi'(s))\right] \right).
\end{aligned}$$

As $e^{-sX} sX \ge e^{-1}$ then $\mathbb{E}\left[2 m_A e^{-sX} sX\right] = o(1)$ as $s \to 0^+$ by dominated convergence. By convexity of $\varphi(s)$ we have $1 - \varphi(s) \ge -\varphi'(s)s$ for any $s > 0$. Thus

$$e^{-m_A(1-\varphi(s))}(-\varphi'(s)s) \le e^{-m_A(-\varphi'(s)s)}(-\varphi'(s)s) \le e^{-1}$$

and the dominated convergence argument also applies to the second integrand as $-\varphi'(s)s \le 1 - \varphi(s) = o(1)$. We obtain $I_1 = o((\mu_D + \varphi'(s))/s)$ as $s \to 0^+$.

In order to control this term, we notice that $\varphi(s)$ is μ_D Lipschitz on $s \geq 0$ so that $|1-\varphi(s)| = 1-\varphi(s) \leq \mu_D s$. Thus

$$\varphi'(s) \leq \frac{\mathbb{E}\left[e^{-sY}X\right]}{1-\kappa} \leq \frac{\varphi_Y'(s)}{1-\kappa} + \frac{\kappa\mu_D}{1-\kappa}$$

where $\varphi_Y(s) = \mathbb{E}[e^{-sY}]$ denotes the Laplace transform of Y. It yields to the estimates $\mu_D + \varphi'(s) = O(\mu_D + \varphi_Y'(s))$ and $I_1 = o((\mu_D + \varphi_Y'(s))/s)$ as $s \to 0^+$. We bound I_2 noticing that $\varphi(s)$ is μ_D Lipschitz on $s \geq 0$ so that $|1-\varphi(s)| = 1-\varphi(s) \leq \mu_D s$. Then

$$sX + m_A(1-\varphi(s)) \leq sX + sm_A\mu_D = sY.$$

Thus, one rewrites I_2 as

$$I_2 = e^{s\mu_D-(1-\varphi(s))}\mathbb{E}\left[e^{-sY}\left(1-e^{sX+m_A(1-\varphi(s))-sY}\right)Y^2\right].$$

Using again the basic inequality $1-e^{-x} \leq x$ for $x \geq 0$ we obtain the new estimate

$$\begin{aligned} I_2 &\leq e^{s\mu_D-(1-\varphi(s))}\mathbb{E}\left[m_A(s\mu_D-(1-\varphi(s))e^{-sY}Y^2\right] \\ &\leq e^{s\mu_D-(1-\varphi(s))}\frac{s\mu_D-(1-\varphi(s))}{s^2}\mathbb{E}\left[m_A e^{-sY}(sY)^2\right]. \end{aligned}$$

As $e^{-sY}(sY)^2 \leq 4e^{-2}$, we prove that $\mathbb{E}\left[m_A e^{-sY}(sY)^2\right] = o(1)$ as $s \to 0^+$ by dominated convergence. Moreover, as $s\mu_D - (1-\varphi(s)) = o(1)$, we obtain

$$I_2 = o\left(\frac{s\mu_D-(1-\varphi(s))}{s^2}\right), \qquad s \to 0^+.$$

Similar computation than above yields

$$\begin{aligned} 0 \leq \varphi(s) - \varphi_Y(s) \leq \mathbb{E}\left[m_A(s\mu_D-(1-\varphi(s))e^{-sX+m_A(\varphi(s)-1))}\right] \\ \leq \kappa(s\mu_D-(1-\varphi(s)). \end{aligned}$$

Thus $(s\mu_D-(1-\varphi(s)) \leq (s\mu_D-(1-\varphi_Y(s))/(1-\kappa)$ and as $\mathbb{E}[Y] = \mathbb{E}[X] + \kappa\mu_D = \mu_D$ we conclude that

$$\left|\varphi''(s) - \frac{\varphi_Y''(s)}{1-\mathbb{E}[e^{-sX-m_A(1-\varphi(s))}m_A]}\right| = o\left(\frac{\mathbb{E}[Y]+\varphi_Y'(s)}{s} + \frac{s\mathbb{E}[Y]-(1-\varphi_Y(s))}{s^2}\right),$$

as $s \to 0^+$. Let us first apply Theorem 5.2 on Y so that $\varphi_Y''(s)$ is $\alpha-2$ regularly varying around 0. Applying Karamata's theorem again, i.e the equivalences between (8.1.11b) and (8.1.9), (8.1.11b) and (8.1.10) in [Bingham et. al , 1987, Theorem 8.1.6] assert respectively that $(s\mathbb{E}[Y] - (1 - \varphi_Y(s)))/s^2 = O(\varphi_Y''(s))$ and $(\mathbb{E}[Y] + \varphi_Y'(s))/s = O(\varphi_Y''(s))$ as $s \to 0^+$. We then obtain

$$\varphi''(s) \sim \frac{\varphi_Y''(s)}{1-\kappa} \sim \frac{cs^{\alpha-2}\ell(1/s)}{1-\kappa}, \qquad s \to 0+,$$

and finally $\bar{F}(x) \sim \bar{F}_Y(x)/(1-\kappa)$, $x \to \infty$, by applying the reverse part of Theorem 5.2. $\square$

We are now ready to characterize the asymptotic behavior of $S(t)$ in the regularly varying case.

Theorem 5.3. *Assume that the assumptions of Lemma 5.2 hold.*
i) If $\alpha \in (0,1)$ and

$$\int_0^\infty u\mathbb{E}h(u,A)du < \infty\,. \tag{39}$$

holds, there exists a sequence (a_n), $a_n \to \infty$, and an α–stable random variable G_α such that

$$\frac{S(t)}{a_{\lfloor \nu t \rfloor}} \xrightarrow{d} G_\alpha\,.$$

ii) If $\alpha \in (1,2)$ and

$$t^{1+\delta-1/\alpha}\mathbb{E}[\mu_A(t,\infty)]ds \to 0\,, \tag{40}$$

as $t \to \infty$ holds for some $\delta > 0$, there exists a sequence (a_n), $a_n \to \infty$, and an α–stable random variable G_α such that

$$\frac{S(t) - t\nu\mu_D}{a_{\lfloor \nu t \rfloor}} \xrightarrow{d} G_\alpha\,.$$

Notice that Condition (39) corresponds to the natural extension of the condition (8) to the marked Hawkes process.

Proof. The proof is based on the representation (9), and application of Propositions 4.1 and 4.2. In either case, it remains to show that

$$\varepsilon_t = o_P(a_t)\,.$$

Consider first the case $\alpha \in (1, 2)$. Since then $\mu_D = \mathbb{E}D < \infty$, the argument in the proof of Theorem 5.1 still yields the bound (35) on $\mathbb{E}\varepsilon_t$. Using L'Hôpital's rule again together with condition (40), shows that $\mathbb{E}\varepsilon_t = o(t^{1/\alpha-\delta})$, where we assume without loss of generality that $\delta < 1/\alpha$. Since, $a_t = t^{1/\alpha}\ell(t)$ for some slowly varying function ℓ, it follows that $\varepsilon_t/a_t \xrightarrow{P} 0$ as $t \to \infty$.

For $\alpha \in (0, 1)$, random variable D has no finite mean. Still, we can again compare the marked Hawkes process N with a stationary version of it, N^* say. The residues at time t of the two point processes are ordered in the sense that ε_t is stochastically dominated by the residue of N^* at the same time, say ε_t^*. However, for that process, by stationarity $\varepsilon_t^* \stackrel{d}{=} \varepsilon_0^*$ for all t. Denote

$$N^* = \sum_k \delta_{\tau_k^*, A^{*k}} \, ,$$

w.l.o.g. assuming that $\tau_i^* \le \tau_{i+1}^*$ for all $i \in \mathbb{Z}$. To denote that τ_j^* is a direct offspring of a claim at time τ_i^*, we write $\tau_i^* \to \tau_j^*$. One can show that under assumption (39), expected number of claims arriving after time 0, triggered by the points before 0 is finite. Indeed, by Fubini's theorem as in the proof of Theorem 5.1, this expectation equals

$$\begin{aligned}
\mathbb{E}\left[\sum_{\tau_i^* \le 0} \sum_{0 < \tau_j^*} \mathbb{I}_{\tau_i^* \to \tau_j^*}\right] &= \mathbb{E}\left[\sum_{\tau_i^* \le 0} \mathbb{E}\left(\sum_{0<\tau_j^*} \mathbb{I}_{\tau_i^* \to \tau_j^*} \middle| \mathcal{G}_0\right)\right] \\
&= \mathbb{E}\left[\sum_{\tau_i^* \le 0} \mu_{A^{*i}}((0 - \tau_i^*, \infty))\right] \\
&= \mathbb{E}\left[\int_{-\infty}^{0} \int_{\mathbb{S}} \mu_a((-s, \infty)) N^*(ds, da)\right].
\end{aligned}$$

Once again, from the projection theorem, Bremaud [1981], Chapter 8, Theorem 3, the last expression equals to

$$
\begin{aligned}
& \mathbb{E}\left[\int_{-\infty}^{0}\int_{\mathbb{S}} \mu_a((-s,\infty))\lambda^*(s)dsQ(da)\right] \\
= \;& \int_{-\infty}^{0}\int_{\mathbb{S}} \mu_a((-s,\infty))\mathbb{E}\left[\lambda^*(s)\right] dsQ(da) \\
= \;& \frac{\nu}{1-\kappa}\int_{0}^{\infty}\int_{\mathbb{S}} \mu_a(s,\infty)dsQ(da) \\
= \;& \frac{\nu}{1-\kappa}\int_{0}^{\infty}\int_{s}^{\infty} \mathbb{E}h(u,A)duds \\
= \;& \frac{\nu}{1-\kappa}\int_{0}^{\infty} \mathbb{E}h(u,A)\int_{0}^{u} ds\, du \\
= \;& \frac{\nu}{1-\kappa}\int_{0}^{\infty} u\mathbb{E}h(u,A)du < \infty,
\end{aligned}
$$

by assumption (39). Therefore, ε_t^* is a proper random variable with the same distribution for any t. Since it dominates ε_t and $a_t \to \infty$, we have $\varepsilon_t/a_t \xrightarrow{P} 0$ as $t \to \infty$. □

Remark 5.2. *Theorem 5.3 i) also holds on the stationary version. It is due to the assumption* (39)*, stronger than conditions* (40) *and* (34)*. We suspect that* (39) *should be weakened, even avoided, thanks to similar arguments than in the proof of Corollary 5.3. However, the dependence between the claim sizes and the arrivals in the marked Hawkes process makes impossible to apply Proposition 4.2 as such.*

Acknowledgements

The work of Bojan Basrak has been supported in part by Croatian Science Foundation under the project 3526. The work of Bojan Basrak and Olivier Wintenberger has been supported in part by the AMERISKA network, project ANR-14-CE20-0006. The authors thank Philippe Soulier for pointing out the reference Hult and Samorodnitsky [2008].

Asmussen, S., Albrecher, H., 2010. Ruin Probabilities. Second Edition. Singapore: World Scientific Publishing Co.Pte. Ltd..

Bacry, E., Delattre, S., Hoffmann, M., Muzy, J.F., 2013. Some limit theorems for Hawkes processes and application to financial statistics. Stochastic Processes and their Applications 123, 2475-2499.

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular variation. Cambrige: Cambrige University Press.

Bremaud, P., 1981. Point processes and queues. New York: Springer Verlag.

Bremaud, P., Massoulie, L., 1996. Stability of nonlinear Hawkes processes. Annals of Probability 24, 1563-1588.

Daley, D.J., 1972. Asymptotic properties of stationary point processes with generalized clusters. Zeitschrift für Wahrscheinlichkeitstheorie und verweite Gebiete 21, 21-65.

Daley, D.J., Vere-Jones, D., 2003. An introduction to the theory of Point processes, Volume I, II., Second Edition. New York: Springer Verlag.

Denisov, D., Foss, S., Korshunov, D., 2010. Asymptotics of randomly stopped sums in the presence of heavy tails. Bernoulli 16, 971-994.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events*. Springer-Verlag, Berlin.

Faÿ, G., Gonzales-Alevaro, B., Mikosch, T., Samorodnitsky, G., 2006. Modeling telegraffic arrivals by a Poisson cluster process. Queueing Systems 54, 121-140.

Gut, A., 2009. Stopped Random Walks, 2nd Edition. New York: Springer Science+Business Media.

Haight, F.A., Breuer, H.A., 1960. The Borel-Tanner Distribution. Biometrika 47, 143-150.

Hawkes, A.G., Oakes, D., 1974. A cluster process representation of a self - exciting process. Journal of Applied Probability 11, 493-503.

Hult, H., Samorodnitsky, G., 2008. Tail probabilities for infinite series of regularly varying random vectors. Bernoulli 14, 838-864.

Jacod, J., Shiryaev, A.N., 2003. Limit Theorems for Stochastic Processes. Berlin Heidelberg: Springer-Verlag.

Karabash, D., Zhu, L., 2015. Limit theorems for marked Hawkes processes with application to a risk model. Stochastic Models 31.

Kerstan, J., 1964. Teilprozesse Poissonscher Prozesse. In Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, Czech. Academy of Science, Prague, 377-403.

Mikosch, T., 2009. Non life insurance mathematics, Second Edition. New York: Springer Verlag.

Resnick, S.I., 1987. Extreme Values, Regular Variation, and Point Processes. New York: Springer-Verlag.

Resnick, S.I., 2007. Heavy tail phenomena. New York: Springer Science+Business Media.

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(1), 85–92.

Stabile, G., Torrisi, G.L., 2010. Risk processes with Non-stationary Hawkes Claims Arrivals. Methodology and Computing in Applied Probability 12, 415-429.

Zhu, L., 2013. Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. Insurance: Mathematics and Economics 53, 544-550.