



HAL
open science

Extremal index for a class of heavy-tailed stochastic processes in risk theory

Charles Tillier

► **To cite this version:**

Charles Tillier. Extremal index for a class of heavy-tailed stochastic processes in risk theory. 2018. hal-01741190

HAL Id: hal-01741190

<https://hal.science/hal-01741190>

Preprint submitted on 22 Mar 2018

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.

Extremal index for a class of heavy-tailed stochastic processes in risk theory

Charles Tillier

Abstract Extreme values for dependent data corresponding to high threshold exceedences may occur in clusters, *i.e.*, in groups of observations of different sizes. In the context of stationary sequences, the so-called extremal index measures the strength of the dependence and may be useful to estimate the average length of such clusters. This is of particular interest in risk theory where public institutions would like to predict the replications of rare events, in other words, to understand the dependence structure of extreme values. In this paper, we characterise the extremal index for a class of stochastic processes that naturally appear in risk theory under the assumption of heavy-tailed jumps. We focus on Shot Noise type-processes and we weaken the usual assumptions required on the Shot functions. Precisely, they may be possibly random with not necessarily compact support and we do not make any assumption regarding the monotonicity. We bring to the fore the applicability of the result on the Kinetic Dietary Exposure Model introduced in [6] used in modeling pharmacokinetics of contaminants.

1 Motivations and framework

The assessment of major risks in our technological society has become vital because of the economic, environmental and human impacts of recent industrial disasters. Hence, risk analysis has received an increasing attention the last past years in the scientific literature in various areas, *e.g.*, in dietary risk, hydrology, finance and insurance; see [7], [1], [14] for instance. By nature, risk theory concerns the probability of occurrence of rare events which are functions - sums or products - of heavy-tailed random variables. Hence, stochastic processes provide an appropriate framework for modeling such phenomena through time. For instance, non-life insurance mathematics deal with particular types of *Shot Noise Processes* (SNP) defined

Charles Tillier, University of Hamburg, SPST, Bundestrasse 55, 20146 Hamburg, Germany, e-mail: charles.tillier@gmail.com

as

$$S_1(t) = \sum_{i=0}^{N(t)} W_i h(t - T_i), \quad t \geq 0, \quad (1)$$

where usually $(W_i)_{i \geq 0}$ are independent and identically distributed (i.i.d.) random variables (r.v.'s), h is a nonincreasing measurable function and N is a homogeneous Poisson process. In this insurance context, S_1 may be used to represent the amount of aggregate claims that an insurer has to cope with; see [21] for a complete review of non-life insurance mathematics. More generally, this kind of jump processes are useful in many applications to model time series for which sudden jumps occur such as in dietary risk assessment, finance, hydrology or as reference models for intermittent fluctuation in physical systems; see [10] and [27] for instance. The study of the extremal behavior of these stochastic processes leads to risk indicators such as the expected time over a threshold or the expected shortfall, which supply information about the exceedences that give rise to hazardous situations; see [26] for instance.

Besides, since extremal events may occur in clusters, the study of the dependence structure of rare events is a major issue, for example to predict potential replications of earthquakes in environmental sciences. This dependence structure may be captured by the extremal index defined in the seminal contribution [18]. Recall that a stationary sequence $(Z_i)_{i \in \mathbb{Z}}$ has an extremal index $\theta \in [0, 1]$ if for all $\tau > 0$ and all sequence $u_n(\tau)$ such that $\lim_{n \rightarrow \infty} \mathbb{P}(Z_1 \geq u_n(\tau)) = \tau$, it holds that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\max_{i=1, \dots, n} Z_i \leq u_n(\tau) \right) = e^{-\theta \tau}. \quad (2)$$

Less formally, the extremal index indicates somehow, how many times in average an extremal event will reproduce. The case $\theta = 1$ (respectively $\theta = 0$) corresponds to independent data, i.e., to extreme values occurring in an isolated fashion (respectively to potentially infinite size clusters).

Authors in [11], [16] and [19] characterise the extremal index in several particular configurations of (1) and study the extremal properties of the process. More recently, [20] compute the extremal index when the jumps $(W_i)_{i \geq 0}$ form a chain-dependent sequence (the cumulative distribution function (c.d.f.) is linked to a secondary Markov chain) and they assume that h is a bounded positive strictly decreasing function supported on a finite interval.

In this paper, we continue the investigation of the extremal index for such stochastic processes relaxing the conditions required on h . We focus on an extension of such SNP on the form

$$S(t) = \sum_{i=0}^{\infty} W_i h_i(t - T_i), \quad t \geq 0, \quad (3)$$

where $t \rightarrow h_i(t)$ is a random function. Throughout the paper, we work under the following three conditions (C1)-(C3).

(C1) *Jumps* $(W_i)_{i \in \mathbb{N}}$ are nonnegative r.v.'s with c.d.f. H such that $\bar{H} = 1 - H$ is regularly varying at infinity with index $-\alpha$, $\alpha > 0$, that is

$$\frac{\bar{H}(wx)}{\bar{H}(x)} \xrightarrow{x \rightarrow \infty} w^{-\alpha}, \quad \forall w > 0.$$

(C2) *Jumps instants* $(T_i)_{i \in \mathbb{N}}$ are defined for $i \geq 1$ by $T_i = \sum_{k=1}^i \Delta T_k$ and $T_0 = 0$ while the *inter-jumps* $(\Delta T_i)_{i \in \mathbb{N}^*}$ are i.i.d. positive r.v.'s with finite expectation.

(C3) For all $i \geq 0$, the random functions h_i are positive, stationary and independent of T_i .

The condition (C1) is the heavy-tailed distribution assumption on the jumps. We refer to [24] for an exhaustive review of the univariate regular variation theory. The condition (C2) means that (T_0, T_1, \dots) forms a renewal sequence so that one may define the associated renewal process $\{N(t)\}_{t \geq 0}$ by $N(t) := \#\{i \geq 0 : T_i \leq t\}$ for $t \geq 0$. The remaining part of the manuscript is organized as follows. In Section 2, we present the main result regarding the extremal index of the process (3) while an illustrative application is given in Section 3. In the Appendix, we recall the main notions involved in the proof of the theorem.

2 The extremal index

The extremal index defined in Equation (2) holds for discrete-time series. The purpose of this work is to investigate the dependence structure of the extreme values of the continuous-time stochastic processes S defined in Equation (3). Depending on the context, it means that we are either interested in the dependence structure of its maxima or of its minima. In dietary risk assessment, S aims at representing the evolution of a contaminant in the human body through time; see Section 3 for more details. Toxicologists determine thresholds from which the exceedence may have some adverse effect for the health of an individual and we are therefore interested in the maxima of S . Similarly, in hydrology, (3) may be used to describe the flow of a river and a hazardous situation - seen as a rare event - arises when the flow exceeds a critical threshold; see [9]. On the other hand, in most of the applications, the random functions h_i are in essence monotonic for each $i \geq 0$ when t grows. To be convinced, let us go back to the dietary risk assessment. In this context, h_i models the elimination of the contamination and is thus a decreasing function for each $i \geq 0$. For instance, $h_i(t) \equiv e^{-t}$, $t \geq 0$ has been proposed in [6] and is also used in non-life insurance mathematics; see [21].

Assuming that the random function h_i are monotonic for each $i \geq 0$, it is straightforward to see that the extreme values - the maxima or the minima - occur on the embedded chain, i.e., the process $\{S(t)\}_{t \geq 0}$ sampled at the jump arrivals T_1, T_2, \dots . As a consequence, the dependence structure of the continuous-time stochastic process S may be deduced from the analysis of the dependence structure of the under-

lying sequence $(S(T_1), S(T_2), \dots)$.

This is the purpose of this section: to compute the extremal index of the embedded chain of the jump process (3). This means that we focus on the following discrete-time risk process

$$S(T_k) = \sum_{i=1}^k W_i h_i(T_k - T_i), \quad k > 0. \quad (4)$$

Hereinafter, for $i \geq 1$, define $-T_{-i}$ (respectively W_{-i} and h_{-i}) as an independent copie of T_i (respectively of W_i and h_i) so that under (C1) and (C3), $(W_i)_{i=-\infty}^{\infty}$ and $(h_i)_{i=-\infty}^{\infty}$ form respectively an i.i.d. and a stationary sequence of positive r.v.'s. Facing with the issue of non stationarity of the embedded chain (4) - required for the computation of the extremal index - we study a stationary version/modification denoted $(S_k)_{k \in \mathbb{Z}}$ and defined by

$$S_k = \sum_{i=-\infty}^k W_i h_i(T_k - T_i), \quad k > 0. \quad (5)$$

We now introduce the condition (D1), under which the stationary sequence (5) is well defined.

(D1) The random function h_i satisfy

- $\sum_{i=0}^{\infty} \mathbb{E}[h_1(T_i)] < \infty, \quad \alpha < 1.$
- There exists $\varepsilon > 0$ such that $\sum_{i=0}^{\infty} \mathbb{E}[h_1^{\alpha-\varepsilon}(T_i)] < \infty, \quad \alpha \leq 2.$
- $\sum_{i=0}^{\infty} \mathbb{E}[h_1^2(T_i)] < \infty, \quad \alpha > 2.$

Theorem 1. *Assume Model (4) holds. Under Conditions (C1)-(C3) and (D1), the extremal index θ is given by*

$$\theta = \frac{\mathbb{E}[\max_{i \geq 0} h_i^{\alpha}(T_i)]}{\sum_{i=0}^{\infty} \mathbb{E}[h_i^{\alpha}(T_i)]}. \quad (6)$$

We do not raise the question of the estimation of the extremal index in this work. In many cases, the jump process (3) is a PDMP (Piecewise-Deterministic Markov Process) and [5] propose a robust estimator for the extremal index; see the application in Section 3 for an illustrative example.

2.1 Proof of Theorem 1

For reader's convenience, the definitions of the main notions, namely the *tail index*, the *anti-clustering* and the *strong mixing* conditions are postponed in the Appendix. Let us first define the intermediate stationary sequences $\{S_k^{(m)}, k \geq 0\}_{m \geq 0}$ such that

$$S_k^{(m)} = \sum_{i=k-m}^k W_i h_i (T_k - T_i). \quad (7)$$

The extremal index θ of the stationary sequence $\{S_k\}_{k \geq 0}$ will be deduced from the extremal index θ_m of $\{S_k^{(m)}, k \geq 0\}_{m \geq 0}$ in the following way. In [3], Theorem 4.5, the authors show that if a jointly regularly varying (see the definition of "jointly regularly varying" in Definition 1 in the Appendix) stationary sequence $(Z_i)_{i \in \mathbb{Z}}$ is strongly mixing and satisfies the anti-clustering condition, then $(Z_i)_{i \in \mathbb{Z}}$ admits an extremal index $\tilde{\theta}$ given by

$$\tilde{\theta} = \mathbb{P} \left(\max_{k \geq 1} Y_k \leq 1 \right), \quad (8)$$

where $(Y_i)_{i \in \mathbb{N}}$ is the tail process of $(Z_i)_{i \in \mathbb{Z}}$. Using this result, we obtain the extremal index θ_m for each $m \geq 1$. Next, we show that the assumptions of Proposition 1.4 in [11] hold to conclude that $\lim_{m \rightarrow \infty} \theta_m = \theta$. We start by characterizing the tail process of sequence $\{S_k^{(m)}\}_{k \in \mathbb{N}}$ in the following lemma.

Lemma 1. *Assume that Conditions (C1)–(C3) and (D1) hold. For each $m \geq 1$, the tail process of $\{S_k^{(m)}\}_{k \in \mathbb{N}}$ denoted by $\{Y_k^{(m)}\}_{k \in \mathbb{N}}$ is defined by*

$$Y_n^{(m)} = \begin{cases} \frac{h_{N_m(T_n+N_m)}}{h_{N_m(T_{N_m})}} Y_0^{(m)}, & 0 \leq n \leq m, \\ 0 & \text{for } n > m, \end{cases}$$

with $P(Y_0^{(m)} > y) = y^{-\alpha}$ and N_m is an integer-valued random variable such that

$$\mathbb{P}(N_m = n) = \frac{\mathbb{E}[h_n^\alpha(T_n)]}{\sum_{i=1}^m \mathbb{E}[h_i^\alpha(T_i)]}, \quad 0 \leq n \leq m.$$

Besides, for any random variable U measurable with respect to $(h_j, T_j)_{j \in \mathbb{Z}}$, we have

$$\mathbb{E}[U \mid N_m = i] = \frac{\mathbb{E}[h_i^\alpha(T_i)U]}{\mathbb{E}[h_i^\alpha(T_i)]}, \quad 0 \leq i \leq m.$$

Proof (Proof of Lemma 1). For clarity of notation, we omit the superscript (m) and we assume that $h_k \equiv 0$ if $k > m$ and we denote $X_1 \stackrel{d}{=} X_2$ when two random variables X_1, X_2 share the same distribution. Then, for a fixed n , we have

$$\begin{aligned} \mathbb{P} \left(\max_{i=0, \dots, n} S_i / y_i \leq x \mid S_0 > x \right) &= 1 - \mathbb{P} \left(\max_{i=0, \dots, n} S_i / y_i > x \mid S_0 > x \right) \\ &= 1 - \frac{\mathbb{P}((\max_{i=0, \dots, n} S_i / y_i) \wedge S_0 > x)}{\mathbb{P}(S_0 > x)}. \end{aligned}$$

Under the two conditions (C1) and (C2), $(W_k)_{k \in \mathbb{Z}}$ is a i.i.d. sequence and $T_i - T_k \stackrel{d}{=} T_{i-k}$ for $i \geq 0$ and $k \leq 0$. It follows that

$$\begin{aligned} \lim_{x \rightarrow \infty} \mathbb{P} \left(\max_{i=0, \dots, n} S_i / y_i > x, S_0 > x \right) &\underset{x \rightarrow \infty}{\sim} \sum_{k=0}^{\infty} \mathbb{P} \left(W_k \left(\max_{i=0, \dots, n} h_k(T_i - T_k) / y_i \right) \wedge h_k(-T_k) > x \right) \\ &\underset{x \rightarrow \infty}{\sim} \sum_{k=0}^{\infty} \mathbb{E} \left[\bigvee_{i=0}^n \frac{h_k^\alpha(T_{k+i})}{y_i} \wedge h_k^\alpha(T_k) \right] \mathbb{P}(W_1 > x), \end{aligned}$$

and

$$S_0 = \sum_{i=-\infty}^0 W_i h_i(T_0 - T_i) \stackrel{d}{=} \sum_{i=0}^{\infty} W_i h_i(T_i),$$

under (C3). Moreover, since the two sequences $(h_i)_{i \in \mathbb{Z}}$ and $(T_i)_{i \in \mathbb{Z}}$ are mutually independent, the results in Section 3 of [17] imply that the series S_0 is almost surely convergent under Condition (D1) and we have

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(S_0 > x)}{\bar{H}(x)} = \sum_{i=0}^{\infty} \mathbb{E}[h_1^\alpha(T_i)] < \infty. \quad (9)$$

From Equation (9), we have proved that S_0 is regularly varying at infinity with the same index α than the jumps $(W_i)_{i \geq 0}$. It also follows that

$$\lim_{x \rightarrow \infty} \mathbb{P} \left(\max_{i=0, \dots, n} S_i / y_i \leq x \mid S_0 > x \right) = 1 - \frac{\mathbb{E} \left[\bigvee_{i=0}^n \frac{h_k^\alpha(T_{k+i})}{y_i} \wedge h_k^\alpha(T_k) \right]}{\sum_{k=0}^{\infty} \mathbb{E}[h_k^\alpha(T_k)]}.$$

Setting

$$p_k = \frac{\mathbb{E}[h_k^\alpha(T_k)]}{\sum_{j=0}^{\infty} \mathbb{E}[h_j^\alpha(T_j)]},$$

we obtain

$$\begin{aligned} \lim_{x \rightarrow \infty} \mathbb{P} \left(\max_{i=0, \dots, n} S_i / y_i \leq x \mid S_0 > x \right) &= 1 - \sum_{k=0}^{\infty} p_k \frac{\mathbb{E} \left[\bigvee_{i=0}^n h_k^\alpha(T_{k+i}) / y_i \wedge h_k^\alpha(T_k) \right]}{\mathbb{E}[h_k^\alpha(T_k)]} \\ &= 1 - \sum_{k=0}^{\infty} p_k \mathbb{E} \left[\bigvee_{i=0}^n \frac{h_N^\alpha(T_{N+i})}{y_i h_N^\alpha(T_N)} \wedge 1 \mid N = k \right] \\ &= 1 - \mathbb{E} \left[\bigvee_{i=0}^n \frac{h_N^\alpha(T_{N+i})}{y_i h_N^\alpha(T_N)} \wedge 1 \right] \\ &= 1 - \mathbb{P} \left(Y_0 \bigvee_{i=0}^n \frac{h_N^\alpha(T_{N+i})}{y_i h_N^\alpha(T_N)} > 1 \right), \end{aligned}$$

where Y_0 is a Pareto random variable independent of $\{h_i, T_i\}_{i \in \mathbb{Z}}$. This proves our claim.

For each $m \geq 1$, the strong mixing condition holds for each sequence $\{S_k^{(m)}, k \geq 0\}$ since it is m -dependent. Indeed, by independence $\alpha_h = 0$ for all $h \geq m + 1$. Likewise, since $\{S_k^{(m)}, k \geq 0\}$ is m -dependent, the anti-clustering condition is satisfied with $r_n \mathbb{P}(Z_1 > a_n) = o(1)$; see Section 4.1 in [2]. As a first consequence, we obtain in the following lemma the expression of the extremal index θ_m of the intermediate sequence $\{S_k^{(m)}\}_{k \geq 0}$.

Lemma 2. *Assume that Conditions (C1)-(C3) and (D1) hold. For each $m \geq 1$, the extremal index θ_m of the intermediate sequence $\{S_k^{(m)}, k \geq 0\}$ defined in (7) is given by*

$$\theta_m = \frac{\mathbb{E} \left[\bigvee_{j=0}^m h_j^\alpha(T_j) \right]}{\sum_{i=0}^m \mathbb{E}[h_1^\alpha(T_i)]}. \quad (10)$$

Proof (Proof of Lemma 2). Fix $m \geq 1$ throughout the proof. By Equation (8), since $\mathbb{P}(Y_0 > x) = x^{-\alpha}$, we have

$$\begin{aligned} \theta_m &= \mathbb{P} \left(\max_{k \geq 1} Y_k^{(m)} \leq 1 \right) \\ &= \mathbb{P} \left(\max_{k \geq 1} Y_0 \Theta_k^{(m)} \leq 1 \right) \\ &= 1 - \mathbb{P} \left(\max_{k \geq 1} Y_0 \Theta_k^{(m)} \geq 1 \right) \\ &= 1 - \mathbb{E} \left[\max_{k \geq 1} \left(\Theta_k^{(m)} \right)^\alpha \wedge 1 \right], \end{aligned}$$

where $(\Theta_k^{(m)})_{k \geq 0}$ refers to the spectral tail process of the intermediate sequence $\{S_k^{(m)}, k \geq 0\}$ defined in [3]. Applying Lemma 1, we obtain

$$\begin{aligned} \theta_m &= \mathbb{P} \left(Y_0 \max_{1 \leq k \leq m} \frac{h_{k+N}(T_{k+N})}{h_N(T_N)} \leq 1 \right) \\ &= 1 - \mathbb{E} \left[\max_{1 \leq k \leq m} \frac{h_{k+N}^\alpha(T_{k+N})}{h_N^\alpha(T_N)} \wedge 1 \right] \\ &= 1 - \sum_{n=0}^m \mathbb{E} \left[\max_{1 \leq k \leq m} \frac{h_{k+N}^\alpha(T_{k+N})}{h_N^\alpha(T_N)} \wedge 1 \mid N = n \right] \mathbb{P}(N = n) \\ &= 1 - \frac{\sum_{n=0}^m \mathbb{E} \left[(\max_{1 \leq k \leq m} h_{n+k}^\alpha(T_{n+k})) \wedge h_n^\alpha(T_n) \right]}{\sum_{j=0}^m \mathbb{E}[h_j^\alpha(T_j)]} \\ &= \frac{\sum_{n=0}^m \mathbb{E} \left[h_n^\alpha(T_n) - (\max_{1 \leq k \leq m} h_{n+k}^\alpha(T_{n+k})) \wedge h_n^\alpha(T_n) \right]}{\sum_{j=0}^m \mathbb{E}[h_j^\alpha(T_j)]}. \quad (11) \end{aligned}$$

Observe that using the identity $\max(a, b) = a + b - \min(a, b)$ for any a, b in \mathbb{R}^+ , one can show that for any sequence $(a_n)_{n \in \mathbb{N}}$ of nonnegative real numbers such that $\sum_{n=0}^{\infty} a_n < \infty$, we have

$$\max_{n \in \mathbb{N}} a_n = \sum_{n \in \mathbb{N}} \left(a_n - \max_{k \geq 1} a_{n+k} \wedge a_n \right). \quad (12)$$

Under (D1), from Equation (9), $\sum_{n=0}^{\infty} \mathbb{E}[h_n^\alpha(T_n)] < \infty$ and we can apply the relation (12) to the last equality (11). This proves Lemma 2.

Up to now, we have characterised the extremal index θ_m from the tail process of $\{S_k^{(m)}\}_{k \in \mathbb{N}}$. To conclude, it remains to prove that the extremal index of $(S_k)_{k \in \mathbb{Z}}$ is given by $\lim_{m \rightarrow \infty} \theta_m = \theta$. For this purpose, we apply Proposition 1.4 of [11]. We must check the following two conditions: for all sequence u_n such that

$$n\mathbb{P}(S_0 > u_n) \rightarrow \beta \in (0, \infty),$$

we have

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} n\mathbb{P}((1 - \varepsilon)u_n < S_0 \leq (1 + \varepsilon)u_n) = 0 \quad (13)$$

and

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} n\mathbb{P}(|S_0 - S_0^{(m)}| > \varepsilon u_n) = 0. \quad (14)$$

Since we have already proved in Equation (9) that S_0 is regularly varying with index $-\alpha$, $\alpha > 0$, we have

$$\limsup_{n \rightarrow \infty} n\mathbb{P}((1 - \varepsilon)u_n < S_0 \leq (1 + \varepsilon)u_n) = \beta((1 - \varepsilon)^{-\alpha} - (1 + \varepsilon)^{-\alpha}).$$

Letting $\varepsilon \rightarrow 0$ proves Equation (13). Moreover, by the same arguments which lead to the expression for the tail behavior of S_0 , we have

$$\lim_{n \rightarrow \infty} n\mathbb{P}(|S_0 - S_0^{(m)}| > u_n \varepsilon) = \varepsilon^{-\alpha} \beta \frac{\sum_{n=m+1}^{\infty} \mathbb{E}[h_1^\alpha(T_n)]}{\sum_{n=0}^{\infty} \mathbb{E}[h_1^\alpha(T_n)]}.$$

Letting $m \rightarrow \infty$ proves Equation (14). We finally have

$$\theta = \lim_{m \rightarrow \infty} \theta_m = \frac{\mathbb{E}[\sum_{i=0}^{\infty} h_i^\alpha(T_i)]}{\sum_{i=0}^{\infty} \mathbb{E}[h_1^\alpha(T_i)]},$$

which concludes the proof of Theorem 1.

3 Application

For the sake of application of Theorem 1, we consider a specific dietary risk assessment model introduced in [6] called KDEM for Kinetic Dietary Exposure Model; see also [7] for the statistical analysis of the model and for more details on dynamic dietary risk processes. For each $i \geq 0$, we assume that the intakes $(W_i)_{i \geq 0}$ are pure Pareto distributed with tail index $\alpha > 0$ and we set $h_i(t) = e^{-\omega_i t} \mathbb{I}_{[0, \infty[}(t)$, where $\mathbb{I}_{[\cdot]}(\cdot)$ is the indicator function. Besides, we consider that $(\omega_i)_{i \in \mathbb{N}}$ is an i.i.d. sequence of

nonnegative r.v.'s with finite expectation. In this context, for each $i \geq 0$, h_i is a non-increasing random elimination function that governs the elimination process of the i -th intake W_i ingested at time T_i up to time t . Then $(\omega_i)_{i \in \mathbb{N}}$ is a random elimination parameter, which permits to take into account fluctuations in the assimilation process. The model may be written as

$$S(t) = \sum_{i=1}^{N(t)} W_i e^{-\omega_i(t-T_i)}, \quad t > 0, \quad (15)$$

where $t \rightarrow N(t) := \#\{i \geq 0 : T_i \leq t\}$ is a renewal process that counts the numbers of intakes that occurred until time $t > 0$. Figure 1 shows how the process (15) evolves through time. To get an explicit result, we consider that the intakes arise regarding a

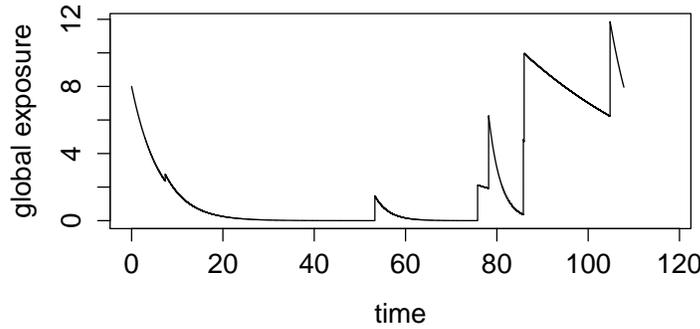


Fig. 1 The elimination driven by the random variable $(\omega_i)_i$ may vary between two intakes. Observe the PDMP-type trajectory. Due to the heavy-tailed distribution of the intakes $(W_i)_i$, some jumps are rather large.

homogeneous Poisson process meaning that the duration between intakes are independent and exponentially distributed. Applying Theorem 1, we get the following explicit formulae of the extremal index.

Proposition 1. *Assume that Model (15) holds with positive i.i.d. $(\omega_i)_{i \in \mathbb{N}}$ satisfying $\mathbb{E}[\omega_1] < \infty$. Assume moreover that $\bar{H}(x) = x^{-\alpha}$, $\alpha > 0$ for all $x > 0$ and N is a Poisson process with intensity $\lambda > 0$. Then we have*

$$\theta = \frac{\alpha}{\alpha + \lambda \mathbb{E}[\omega^{-1}]}. \quad (16)$$

Proof. Note first that in this setup, it is straightforward that for all $\alpha > 0$, Assumptions (C1)-(C3) as well as (D1) hold. Indeed, the Pareto distribution is a particular

case of such regularly varying random variables so that (C1) is satisfied. (C2) holds as a sum of i.i.d. r.v.'s whose distribution is exponential with mean $1/\lambda$, $\lambda > 0$. Finally, (C3) is satisfied since the random variables $(\omega_i)_{i \in \mathbb{N}}$ are i.i.d., then the random functions $(h_i)_{i \in \mathbb{N}}$ are positive i.i.d. r.v.'s with $0 < \mathbb{E}[\omega_1] < \infty$ implying (D1). Now, observe first that for the numerator, we have

$$\mathbb{E} \left[\max_{i \geq 0} \{h_i^\alpha(T_i)\} \right] = \mathbb{E} \left[\max_{i \geq 0} \{e^{-\alpha \omega_i T_i} \mathbb{1}_{[0, \infty[}(T_i)\} \right] \leq 1.$$

Besides, since $T_0 = 0$ under (C2), we have

$$\mathbb{E}[h_0^\alpha(T_0)] = \mathbb{E}[e^{-\alpha \omega_0}] = 1,$$

leading to $\mathbb{E}[\max_{i \geq 0} \{h_i^\alpha(T_i)\}] = 1$. It also follows that the denominator may be written as

$$\sum_{i=0}^{\infty} \mathbb{E}[h_i^\alpha(T_i)] = 1 + \sum_{i=1}^{\infty} \mathbb{E}[h_i^\alpha(T_i)]$$

with

$$\begin{aligned} \sum_{i=1}^{\infty} \mathbb{E}[h_i^\alpha(T_i)] &= \sum_{i=1}^{\infty} \int_0^{\infty} \left(\int_0^{\infty} e^{-\alpha \omega t} dF_{\Delta T}^{*i}(t) \right) dF_W(\omega) \\ &= \int_0^{\infty} \left(\sum_{i=1}^{\infty} \mathbb{E}[e^{-\alpha \omega \Delta T}]^i \right) dF_W(\omega) \\ &= \int_0^{\infty} \frac{\mathbb{E}[e^{-\alpha \omega \Delta T}]}{1 - \mathbb{E}[e^{-\alpha \omega \Delta T}]} dF_W(\omega) \\ &= \int_0^{\infty} \frac{\lambda}{\alpha \omega} dF_W(\omega) \\ &= \frac{\lambda}{\alpha} \mathbb{E}[\omega^{-1}], \end{aligned}$$

where "*" refers to the convolution operator. This concludes the proof.

To conclude this part, we briefly discuss the veracity of Proposition 1. In this regard, assume Model (15) holds with the assumptions of Proposition 1. Assume moreover that the elimination parameter $\omega > 0$ is constant. Observe now that its embedded chain, namely

$$S(T_k) = \sum_{i=1}^k W_i e^{-\omega(T_k - T_i)}, \quad k > 0$$

may be expressed as

$$S(T_k) = e^{-\Delta T_k} S(T_{k-1}) + W_k, \quad k > 0. \quad (17)$$

The latest equation is nothing else than a particular case of the so-called SDE for Stochastic Differential Equation. It has been studied for a while. In particular, [22]

showed that its extremal index is given by $\theta = 1 - \mathbb{E}[e^{-\Delta T_1}]$. In the specific setup of Proposition 1 where the $(\Delta T_i)_{i \in \mathbb{N}^*}$ are exponentially distributed, the Laplace transform $\mathbb{E}[e^{-\Delta T_1}]$ is explicit. It follows that the extremal index is given by

$$\theta = \frac{\alpha}{\alpha + \lambda \omega}.$$

We retrieve the result of Proposition 1 with constant $\omega > 0$.

4 Appendix

For reader's convenience, we recall in this part important notions involved in the proof of Theorem 1. We start by the definition of the so-called "tail process" introduced recently by [3].

Definition 1 (The tail process). Let $(Z_i)_{i \in \mathbb{Z}}$ be a stationary process in \mathbb{R}^+ and let $\alpha \in (0, \infty)$. If $(Z_i)_{i \in \mathbb{Z}}$ is jointly regularly varying with index $-\alpha$, that is, all vectors of the form $(X_k, \dots, X_l), k \leq l \in \mathbb{Z}$ are multivariate regularly varying, then there exists a process $(Y_i)_{i \in \mathbb{Z}}$ in \mathbb{R}^+ , called the tail process such that $\mathbb{P}(Y_0 > y) = y^{-\alpha}, y \geq 1$ and for all $(n, m) \in \mathbb{Z}^2, n \geq m$

$$\lim_{z \rightarrow \infty} \mathbb{P}((z^{-1}Z_n, \dots, z^{-1}Z_m) \in \cdot \mid Z_0 > z) = \mathbb{P}((Y_n, \dots, Y_m) \in \cdot).$$

We recall now the strong mixing and anti-clustering conditions.

Definition 2 (Strong mixing condition). A stationary sequence $(Z_k)_{k \in \mathbb{Z}}$ is said to be strongly mixing with rate function α_h if

$$\sup |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| = \alpha_h \rightarrow 0, \quad h \rightarrow \infty, \quad (18)$$

where the supremum is taken over all sets $A \in \sigma(\dots, Z_{-1}, Z_0)$ and $B \in \sigma(Z_h, Z_{h+1}, \dots)$

Definition 3 (Anti-clustering condition). A positive stationary sequence $(Z_k)_{k \in \mathbb{Z}}$ is said to satisfy the anti-clustering condition if for all $u \in (0, \infty)$,

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\max_{k \leq |i| \leq r_n} Z_i > a_n u \mid Z_0 > a_n u \right) = 0. \quad (19)$$

Acknowledgment. Charles Tillier would like to thank Patrice Bertail and Philippe Soulier, both Professors at Paris Nanterre University for insightful comments and discussions which led to an improvement of the paper. Financial supports by the ANR network AMERISKA ANR 14 CE20 0006 01 and the Labex MME-DII are also gratefully acknowledged by the author.

References

1. Asmussen, S. (2003). *Applied Probabilities and Queues*. Springer-Verlag, New York.
2. Bartkiewicz, K., Jakubowski, A., Mikosch T. and Wintenberger, O. (2011). Stable limits for sums of dependent infinite variance random variables. *Probability Theory and Related Fields*, Springer Verlag, **150** (3), pp. 337-372.
3. B. Basrak and Segers (2009). Regularly varying multivariate time series. *Stochastic Processes and their Applications*. Vol **119**, Issue 4, 1055–1080.
4. Bertail, P., Clmenon, S. and Tillier, C. (2016) Extreme Values Statistics for Markov Chains with Applications to Finance and Insurance, in *Extreme Events in Finance: A Handbook of Extreme Value Theory and its Applications* (ed F. Longin), John Wiley and Sons, Inc., Hoboken, New Jersey. doi: 10.1002/9781118650318.ch7.
5. Bertail P., Ciolek G. and Tillier C. (2018). Robust estimation for Markov chains with application to PDMPs. To appear in the book *Statistical Inference for Piecewise-Deterministic Markov Processes*.
6. Bertail, P., Clemencon, S. and Tressou, J. (2008). A storage model with random release rate for modelling exposure to food contaminants. *Mathematical Bioscience and Engineering*, **5**, 35-60.
7. Bertail, P., Clemencon, S. and Tressou, J. (2010). Statistical analysis of a dynamic model for food contaminant exposure with applications to dietary methylmercury contamination. *Journal of Biological Dynamics*, **4**, 212-234.
8. Bertail, P., Clemencon, S. and Tillier, C. (2016). *Extreme Events in Finance: an handbook of extreme value theory and its applications in the handbooks series Financial engineering and econometrics*. Edited by F. Longin, Wiley.
9. Lefebvre, M. and Guibault, J.L. (2008). Using filtered Poisson processes to model a river flow. *Applied Mathematical Modelling*. Vol **32**, Issue 12, pp 2792-2805.
10. Bondesson, L. (2006). Shot-Noise Processes and Distributions. *Encyclopedia of Statistical Sciences*, **12**.
11. Chernick, M. R., Hsing, T. and McCormick, W.P. (1991). Calculating the extremal index for a class of stationary sequences. *Adv. Appl. Prob.*, **25**, 835-850.
12. Cline, D.B.H and Samorodnitsky, G. (1994). Subexponentiality of the product of independent random variables. *Stochastic process and their applications*, **49**, 75-98.
13. Doney, R.A. and O'Brien, G.L. (1991). Loud shot noise. *Annals of applied probabilities*, **1**, 88-103.
14. Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Applications of Mathematics, **33**, Springer-Verlag.
15. Ferreira, H. and Pereira, L. (2010). How to compute the extremal index of stationary random fields. *Statistics and Probability Letters*, Elsevier, **78**.
16. Hsing, T. L. and Tengels, J. L. (1989). Extremal properties of shot noise processes, *Adv. Appl. Prob.* **21**, 513-525.
17. Hult, H. and Samorodnitsky, G. (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli*, **14**, 838864.
18. Leadbetter M. R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrscheinlichkeitsthe.* **65**, 291306.
19. McCormick, W. P. (1997). Extremes for shot noise processes with heavytailed amplitudes. *J. Appl. Prob.*, **34**, 643-656.
20. McCormick, W.P. and Seymour, L. (2001). Extreme values for a class of shot-noise processes. *The Institute of Mathematical Statistics, Lecture Notes–Monograph Series*, **37**, 33-46.
21. Mikosch, T. (2010). *Non life Insurance Mathematics*. Springer-Verlag Berlin Heidelberg.
22. Perfekt, R. (1994). Extremal Behaviour of Stationary Markov Chains with Applications. *The Annals of Applied Probability*. Vol. **4**, No. 2, 529–548.
23. Tressou, J. (2005). *Methodes statistiques pour l' evaluation du risque alimentaire*. Ph.D. thesis, University Paris X, 2005, <http://tel.archives-ouvertes.fr/tel-00139909>.

24. Resnick, S.I. (2007). *Heavy-Tail Phenomena and statistical modeling*. Springer series in operations Research and Financial Engineering, Springer-Verlag, New-York.
25. Hombler, P. and Williams, P. (1995). Weak Limit Results for the Extremes of a Class of Shot Noise Processes. *Journal of Applied Probability*, Vol. **32**, 707-726.
26. Tillier, C. and Wintenberger, O. (2017). Regular variation of a random length sequence of random variables and application to risk assessment. *Extremes* . p 1-30. <https://doi.org/10.1007/s10687-017-0297-1>
27. Weng, C., Zhang, Y. and Tang, K.S. (2013). Tail behavior of Poisson Shot Noise Processes under heavy-tailed shocks and Actuarial applications. *Methodology and Computing in Applied Probability*, **15**, 665-682.