DYNAMICAL COUNTEREXAMPLES REGARDING THE EXTREMAL INDEX AND THE MEAN OF THE LIMITING CLUSTER SIZE DISTRIBUTION

MIGUEL ABADI, ANA CRISTINA MOREIRA FREITAS, AND JORGE MILHAZES FREITAS

Abstract. The Extremal Index is a parameter that measures the intensity of clustering of rare events and is usually equal to the reciprocal of the mean of the limiting cluster size distribution. We show how to build dynamically generated stochastic processes with an Extremal Index for which that equality does not hold. The mechanism used to build such counterexamples is based on considering observable functions maximised at at least two points of the phase space, where one of them is an indifferent periodic point and another one is either a repelling periodic point or a non periodic point. The occurrence of extreme events is then tied to the entrance and recurrence to the vicinities of those points. This enables to mix the behaviour of an Extremal Index equal to 0 with that of an Extremal Index larger than 0. Using bi-dimensional point processes we explain how mass escapes in order to destroy the usual relation. We also perform a study about the formulae to compute the limiting cluster size distribution introduced in [14, 4] and prove that ergodicity is enough to establish that the reciprocal of the Extremal Index is equal to the limit of the mean of the finite time cluster size distribution, which, in the case of the counterexamples given, does not coincide with the mean of the limit of the cluster size distribution.

Contents

Date: August 10, 2018.

2010 *Mathematics Subject Classification.* 37A50, 60G70, 60G55, 37B20, 37A25.

All authors were partially supported by the joint project FAPESP (SP-Brazil) and FCT (Portugal) with reference FAPESP/19805/2014. ACMF and JMF were partially supported by FCT projects PTDC/MAT-CAL/3884/2014 and PTDC/MAT-PUR/28177/2017, with national funds, and by CMUP (UID/MAT/00144/2013), which is funded by FCT with national (MCTES) and European structural funds through the programs FEDER, under the partnership agreement PT2020.

1. Introduction

In Extreme Value Theory, there is a parameter that quantifies the intensity of clustering of extreme events. This parameter, which we will denote by θ, was called Extremal Index (EI) by Leadbetter in [21], takes values in $[0, 1]$ and is such that $\theta = 1$ means absence of clustering while θ close to 0 means intensive clustering.

In order to keep track of the occurrence of extreme events, one can consider point processes that count the number of such occurrences on a normalised time frame. In [17], these point processes were proved to converge to a compound Poisson process where the Poisson events are charged by a multiplicity corresponding to the cluster size. Under some regularity conditions, the EI can be identified as the inverse of the mean cluster size, *i.e.*, , θ^{-1} is the average of multiplicity distribution of the limiting compound Poisson process. When $\theta = 1$, the cluster size is 1 a.s. and the limiting process is a simple Poisson process.

However, in [28] a counterexample was given where the EI does not coincide with the inverse of the mean cluster size of the limiting compound Poisson process. This example is based on a regenerative sequence with EI equal to 1/2 but with a simple Poisson process limit, which means a mean cluster size equal to 1. The regenerative property of the sequence is the key to prove the existence of an EI equal to 1/2, which is guaranteed by [27, Theorem 3.1].

More recently, a theory of extreme values for dynamical systems has been developed (see [23] and references therin). The idea is to consider stochastic processes arising from dynamical systems by evaluating a given observable along the orbits of that system. This observable is typically maximised at a single point ζ chosen in the phase space and then, as observed in [11], the study of the occurrence of extreme events is related to problems of entrance and recurrence times. In [13], the authors have shown that periodicity of ζ implies the appearance of clustering and, consequently, an EI less than 1, which is given by the rate of expansion of the system at the maximal point ζ. Later, in [14], the authors showed that at periodic points the point processes of extreme events (or Rare Events Point Process (REPP)) converge to a compound Poisson process with a geometric multiplicity distribution of average θ^{-1}. Moreover, for sufficiently regular systems, a full dichotomy exists (see [20, 4]), *i.e.*, , either ζ is periodic and we have clustering or ζ is non-periodic and we have the absence of clustering with $\theta = 1$ and standard Poisson process as a limit for the REPP. In [5], the authors introduced a new device to create clustering: instead of considering observables maximised at a single point, they consider multiple maximising points and show that if these points are related by belonging to the same orbit then a fake periodic behaviour emerges, which is responsible for the appearance of clustering of extreme observations. This approach yielded examples of different clustering patterns corresponding to different multiplicity distributions pertaining to the cluster size. However, in all such examples the EI coincides with inverse of the mean cluster size.

In this paper, we use the same mechanism to produce new counterexamples of stochastic processes with an EI that cannot be interpreted as the inverse of the mean cluster size of the corresponding limiting process. The idea is to consider an observable maximised at (at least) two points, where one of them is an indifferent periodic point while the other is either a non-periodic point or a repelling periodic point. We recall that when an observable is maximised at a single indifferent fixed point, we obtain an EI $\theta = 0$ (see [16]). Hence, we are mixing a degenerate behaviour corresponding to an EI equal to 0 with an EI strictly larger than 0 to obtain a stochastic process with an EI, which somehow corresponds to an average of these two types of behaviour, but whose finite time multiplicity distributions are not uniformly integrable and, therefore, the mean of the respective limit does not coincide with the inverse of the EI. To prove these statements we will use the formulas for EI given in [13], the formulas for the multiplicity distributions given in [14, 4], the dynamics of the Manneville-Pomeau map and also some tools from [16].

We remark that, in the counterexample built by Smith in [28], one can show that the regenerative process is also mixing the behaviour of an EI equal to 0 with that of an EI equal to 1, which we defer to [3]. Moreover, in all counterexamples, the EI still coincides with reciprocal of the limit of the mean finite time cluster size distribution. Hence, the problem is that the limit of the mean finite time cluster size distribution does not coincide with the mean of the limiting cluster size distribution. This happens because there exists an escape of mass, which can be detected by looking at bi-dimensional point processes of rare events, which can be projected to obtain the one dimensional REPP mentioned earlier. In Section 4 , we describe how the behaviour corresponding to an EI equal to 0 is responsible for the escape of mass observed in the counterexamples, which ultimately explains why the usual interpretation for the EI fails in these situations.

Another highlight of this paper is the fact that we provide a nice interpretation of the formula to compute the cluster size distribution of the limiting process that was introduced in [14, 4] and relate it to the one used by Robert in [26], for example. Moreover, we prove that ergodicity is sufficient to show that the EI still coincides with reciprocal of the limit of the mean finite time cluster size distribution.

2. Extremal analysis of stationary stochastic processes

In this section we let $X_0, X_1, \ldots$ denote a general stationary stochastic process, which we identify with the respective coordinate-variable process on $(\mathcal{R}^{\mathbb{N}_0}, \mathcal{B}^{\mathbb{N}_0}, \mathbb{P})$, where $\mathcal{R} = \mathbb{R}^d$ and $\mathcal{B}^{\mathbb{N}_0}$ is the σ-field generated by the coordinate functions $V_n : \mathcal{R}^{\mathbb{N}_0} \to \mathcal{R}$, with $V_n(x_0, x_1, \ldots) = x_n$, for $n \in \mathbb{N}_0$, so that there is a natural measurable map, the shift operator $\mathcal{T} : \mathcal{R}^{\mathbb{N}_0} \to \mathcal{R}^{\mathbb{N}_0}$, given by $\mathcal{T}(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$, which when applied later in the dynamical systems context can be identified with T_α.

Observe that:

$$V_{i-1} \circ \mathcal{T} = V_i, \quad \text{for all } i \in \mathbb{N}.$$

Since, we are assuming that the process is stationary, then $\mathbb{P}$ is $\mathcal{T}$-invariant. Note that $V_i = V_0 \circ \mathcal{T}^i$, for all $i \in \mathbb{N}_0$, where $\mathcal{T}^i$ denotes the i-fold composition of $\mathcal{T}$, with the convention that $\mathcal{T}^0$ denotes the identity map on $\mathcal{R}^{\mathbb{N}_0}$.

In what follows, for every $A \in \mathcal{B}$, we denote the complement of A as $A^c := \mathcal{X} \setminus A$.

2.1. **Clustering of rare events.** Consider an extreme or rare event $A \in \mathcal{B}$ whose occurrence we want to study. For independent and identically distributed (iid) stochastic processes we expect the occurrences of A to appear scattered along the time line. When the random variables are not independent then there may be a tendency for the observations of A to appear concentrated in groups (clusters). This is sometimes referred as the law of series, see [8]. Identifying the clusters becomes a problem because sometimes is not clear if a certain observation of A is sufficiently close (in time) to others in order to be classified as belonging to the same cluster. We are going to assume that there exists a fixed $q \in \mathbb{N}$, which will be the maximum waiting time between the occurrence of two extreme events on the same cluster. We define the sequence of nested sets $\left(U^{(\kappa)}(A)\right)_{\kappa \geq 0}$ of $\mathcal{B}^{\mathbb{N}_0}$ given by:

$$U^{(0)}(A) = V_0^{-1}(A)$$

$$Q_q^{(0)}(A) = U^{(0)}(A) \cap \bigcap_{i=1}^{q} \mathcal{T}^{-i}((U^{(0)}(A))^c),$$

and for $\kappa \in \mathbb{N}$,

$$U^{(\kappa)}(A) = U^{(\kappa-1)}(A) \setminus Q_q^{(\kappa-1)}(A) \tag{2.1}$$

$$Q_q^{(\kappa)}(A) := U^{(\kappa)}(A) \cap \bigcap_{i=1}^{q} \mathcal{T}^{-i}\left((U^{(\kappa)}(A))^c\right) \tag{2.2}$$

$$U^{(\infty)}(A) = \bigcap_{\kappa \geq 0} U^{(\kappa)}(A). \tag{2.3}$$

Note that $U^{(\kappa-1)}(A)$ corresponds to observing A at time 0 and then observing A for at least κ times so that the waiting time between two observations of A is at most q. Namely, briefly, once A is observed, the size of the cluster is at least κ. Observe that $Q_q^{(\kappa-1)}(A) = U^{(\kappa-1)}(A) \setminus U^{(\kappa)}(A)$ corresponds to the observing A exactly κ times within no more than q units of time between one and the next observation of A. This means that the $\kappa+1$-th observation of A occurs at least $q+1$ iterations after the κ-th observation of A. Again, briefly, in this case, we are saying that the size of the cluster is exactly κ. The event $U^{(\infty)}(A)$ corresponds to the occurrence of an observation of A, which is followed by an infinite number of observations of A which are at most q units of time apart from each other. To put it in a different way, if we define

$$\begin{array}{rll} h: & \mathcal{R}^{\mathbb{N}_0} & \to \{0,1\}^{\mathbb{N}_0} \\ & \underline{x} = (x_0, x_1, \ldots) & \mapsto \underline{\omega} = \omega_0\omega_1\ldots \end{array} \tag{2.4}$$

by setting for each $n \in \mathbb{N}_0$ that $\omega_n = V_n(h(\underline{x})) = 1$ if $x_n \in A$ and $\omega_n = V_n(h(\underline{x})) = 0$ if $x_n \notin A$, then if $\underline{x} = (x_0, x_1, \ldots) \in U^{(\kappa)}(A)$ then $h(\underline{x})$ is a binary sequence, which starts with a 1 and has no block of more than $q-1$ consecutive 0's. Let J be an interval contained in $[0,\infty)$. We define

$$\mathscr{W}_J(A) := \bigcap_{i\in J\cap\mathbb{N}_0} \mathcal{T}^{-i}(Z_0^{-1}(A^c)). \tag{2.5}$$

Note that if $\underline{x} \in \mathscr{W}_J(A)$ means that $h(\underline{x})$ has a block of consecutive 0's that correspond to the observations in $J\cap\mathbb{N}_0$. We can now write a formula to determine the cluster sizer distribution of observations of A. We define the mass probability function π_A supported on the positive integers by

$$\pi_A(\kappa) = \frac{\mathbb{P}(Q_q^{(\kappa-1)}(A)) - \mathbb{P}(Q_q^{(\kappa)}(A))}{\mathbb{P}(Q_q^{(0)}(A))}, \quad \text{for each } \kappa\in\mathbb{N}. \tag{2.6}$$

This formula for the finite time cluster size distribution was used first in [14] and explicitly written for the first time in [4]. It appeared subsequently in [5, 6]. This formula was derived during the proof of the convergence of REPP, which was based on a blocking type of argument. Although very useful it lacked a clear intuitive interpretation, which we mean to provide next.

In order to establish the convergence of the REPP, we will describe a condition $Д'_q(u_n)^*$ inspired in condition $D'_p(u_n)^*$ from [14], which is also very similar to the condition $D^{(k)}(u_n)$ introduced by Chernick et al. in [7]. This condition implies that the maximum waiting time before another observation of A within the same cluster is q units of time. Hence, if $(x_0,x_1,\ldots)\in\mathcal{R}^{\mathbb{N}_0}$ is a realisation of $X_0,X_1,\ldots$ then the beginning of cluster and the ending of cluster can be easily identified in $h(x_0,x_1,\ldots)$ by the appearance of a block of at least q consecutive 0's. Let $q,\kappa\in\mathbb{N}$ be fixed and consider the set of finite strings of 0's and 1's such that each string starts and ends with a 1, has exactly κ 1's, which are separated by at most $q-1$ 0's, *i.e.*, there is no block of q or more consecutive 0's in the string. Namely, let

$$W_q(\kappa) = \Bigg\{ \varpi \in \bigcup_{i=\kappa}^{q(\kappa-1)+1} \{0,1\}^i : V_0(\varpi) = V_{|\varpi|-1}(\varpi) = 1, \sum_{i=0}^{|\varpi|-1} V_0(\mathcal{T}^i(\varpi)) = \kappa, \\ \mathcal{T}^i(\varpi) \in \bigcup_{j=0}^{q-1} V_j^{-1}(1), \text{ for all } i=0,\ldots,|\varpi|-1 \Bigg\},$$

where we still use the notation $\mathcal{T}$ and V_j for the shift map and the projection on the j-th coordinate even when leading with finite strings and $|\varpi|$ is the length of the finite string ϖ. Finally we define:

$$\mathcal{H}_q(\kappa) = h^{-1}\left(\left\{ \underline{\omega}\in\{0,1\}^{\mathbb{N}_0} : \underline{\omega} = \underbrace{0\ldots0}_{q\text{ symbols}}\,\varpi\,\underbrace{0\ldots0}_{q\text{ symbols}}\ldots, \text{ for some } \varpi\in W_q(\kappa)\right\}\right)$$

and also set

$$\mathcal{H}_q(0) = h^{-1}\left(\left\{ \underline{\omega}\in\{0,1\}^{\mathbb{N}_0} : \underline{\omega} = \underbrace{0\ldots0}_{q\text{ symbols}}1\ldots\right\}\right).$$

Observe that $\mathcal{H}_q(0)$ determines the beginning of a new cluster, while $\mathcal{H}_q(\kappa)$ corresponds to the appearance of a cluster of size κ. Observe that $\mathcal{H}_q(\kappa)\subset\mathcal{H}_q(0)$ and to illustrate the definition we note that $0011010110100\ldots\in h(\mathcal{H}_2(6))$ and $0001001011000\ldots\in h(\mathcal{H}_3(4))$.

The next result gives an interpretation of π_A defined in (2.6) as the cluster size distribution, *i.e.*, as the probability of having a cluster of size κ conditioned to knowing that we have initiated a cluster.

Theorem 2.1. *Consider the distribution π_A given by* (2.6)*. We can write:*

$$\pi_A(\kappa) = \mathbb{P}(\mathcal{H}_q(\kappa)|\mathcal{H}_q(0)). \tag{2.7}$$

Remark 2.2. We note that the formula on the right hand side of (2.7) can be identified precisely as the distribution of $C_r^{|k}$ considered in [26] for the cluster size distribution.

Proof. By definition of $Q_q^{(\kappa)}(A)$ given in (2.2), we have

$$\mathcal{H}_q(\kappa) = \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A)) \setminus \bigcup_{i=0}^{q-1} \left(\mathcal{T}^{-i}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))\right).$$

Observe that $\left(\mathcal{T}^{-i}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))\right) \bigcap \left(\mathcal{T}^{-j}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))\right) = \emptyset$, for all $0 \leq i \neq j \leq q-1$. To see this, assume w.l.o.g. that $0 \leq i < j \leq q-1$ and take $\underline{x} = (x_0, x_1, \ldots) \in \mathcal{R}^{\mathbb{N}_0}$ such that $\underline{x} \in \mathcal{T}^{-i}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))$ then realise that $V_i(h(\underline{x}))) = 1$ and $V_\ell(h(\underline{x}))) = 0$ for all $\ell = i+1, \ldots, q-1$, while $i < j \leq q-1$ and $\underline{x} \in \mathcal{T}^{-j}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))$ means that in particular that $V_j(h(\underline{x}))) = 1$, which establishes that the two events are definitely incompatible. Hence, by stationarity we have:

$$\begin{aligned}\mathbb{P}(\mathcal{H}_q(\kappa)) &= \mathbb{P}(\mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))) - \sum_{i=0}^{q-1} \mathbb{P}\left(\mathcal{T}^{-i}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))\right) \\ &= \mathbb{P}(Q_q^{(\kappa-1)}(A)) - \sum_{i=0}^{q-1} \mathbb{P}\left(Q_q^{(\kappa)}(A) \cap \mathcal{T}^{-q+i}(Q_q^{(\kappa-1)}(A))\right).\end{aligned}$$

Now, we claim that $Q_q^{(\kappa)}(A) = \dot\bigcup_{i=0}^{q-1} Q_q^{(\kappa)}(A) \cap \mathcal{T}^{-q+i}(Q_q^{(\kappa-1)}(A))$, where $\dot\cup$ stands for disjoint union. To see this observe that

$$\begin{aligned}Q_q^{(\kappa)}(A) &= h^{-1}\left(\left\{\underline{\omega} \in \{0,1\}^{\mathbb{N}_0} : \underline{\omega} = \varpi \ldots, \text{ for some } \varpi \in W_q(\kappa+1)\right\}\right) \\ &= \dot\bigcup_{i=0}^{q-1} h^{-1}\left(\left\{\underline{\omega} \in \{0,1\}^{\mathbb{N}_0} : \underline{\omega} = 1\,\underbrace{0\ldots0}_{\text{i symbols}}\,\varpi \ldots, \text{ for some } \varpi \in W_q(\kappa)\right\}\right) \\ &= \dot\bigcup_{j=1}^{q} Q_q^{(\kappa)}(A) \cap \mathcal{T}^{-j}(Q_q^{(\kappa-1)}(A)).\end{aligned}$$

It follows that

$$\mathbb{P}(\mathcal{H}_q(\kappa)) = \mathbb{P}(Q_q^{(\kappa-1)}(A)) - \mathbb{P}(Q_q^{(\kappa)}(A)).$$

Note that $Q_q^{(0)}(A) = \mathscr{W}_{[0,q+1)}(A) \setminus \mathcal{T}^{-1}(\mathscr{W}_{[0,q)}(A))$ and $\mathcal{H}_q(0) = \mathscr{W}_{[0,q+1)}(A) \setminus \mathscr{W}_{[0,q)}(A)$. Therefore, by stationarity $\mathbb{P}(Q_q^{(0)}(A)) = \mathbb{P}(\mathcal{H}_q(0)) = \mathbb{P}(\mathscr{W}_{[0,q+1)}(A)) - \mathbb{P}(\mathscr{W}_{[0,q)}(A))$. Recalling that $\mathcal{H}_q(\kappa) \subset \mathcal{H}_q(0)$ we obtain:

$$\mathbb{P}(\mathcal{H}_q(\kappa)|\mathcal{H}_q(0)) = \frac{\mathbb{P}(\mathcal{H}_q(\kappa))}{\mathbb{P}(\mathcal{H}_q(0))} = \frac{\mathbb{P}(Q_q^{(\kappa-1)}(A)) - \mathbb{P}(Q_q^{(\kappa)}(A))}{\mathbb{P}(Q_q^{(0)}(A))} = \pi_A(\kappa).$$

□

From the formula (2.6) we can easily derive a formula for the mean finite time cluster size distribution, which will see below to coincide with the reciprocal of the definition of the Extremal Index.

Theorem 2.3. *If* $\mathbb{P}(U^{(\infty)}(A)) = 0$*, then*

$$\sum_{j=1}^{\infty} j\pi_A(j) = \frac{\mathbb{P}(U^{(0)}(A))}{\mathbb{P}(Q_q^{(0)}(A))}.$$

Proof. Observe that by construction, for all $n \in \mathbb{N}$, we have $Q_q^{(\kappa)}(A) \cap Q_q^{(j)}(A) = \emptyset$, for all $\kappa \neq j$. Moreover, $U^{(0)}(A) = \bigcup_{\kappa=0}^{\infty} Q_q^{(\kappa)}(A) \cup U^{(\infty)}(A)$ and then, by assumption, we have

$$\mathbb{P}(U^{(0)}(A)) = \sum_{\kappa=0}^{\infty} \mu_\alpha(Q_q^{(\kappa)}(A)).$$

It follows that

$$\begin{aligned}\sum_{j=1}^{\infty} j\pi_A(j) &= \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} \pi_A(j) = \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} \frac{\mathbb{P}(Q_q^{(j-1)}(A)) - \mathbb{P}(Q_q^{(j)}(A))}{\mathbb{P}(Q_q^{(0)}(A))} = \frac{\sum_{i=1}^{\infty} \mathbb{P}(Q_q^{(i-1)}(A))}{\mathbb{P}(Q_q^{(0)}(A))} \\ &= \frac{\mathbb{P}(U^{(0)}(A))}{\mathbb{P}(Q_q^{(0)}(A))}.\end{aligned}$$

□

Corollary 2.4. *If* $\mathcal{T}$ *is ergodic w.r.t.* $\mathbb{P}$ *and* $\mathbb{P}(\mathscr{W}_{[0,q+1)}(A)) > 0$ *then* $\mathbb{P}(U^{(\infty)}(A)) = 0$ *and therefore the statement of the previous theorem holds.*

Proof. Let $B^\infty = \bigcup_{i=0}^{q} \mathcal{T}^{-i}(U^{(\infty)}(A))$. Observe that $\mathcal{T}^{-1}(B^\infty) \subset B^\infty$ and since by invariance of $\mathbb{P}$ we also have $\mathbb{P}(\mathcal{T}^{-1}(B^\infty)) = \mathbb{P}(B^\infty)$ then $\mathbb{P}(\mathcal{T}^{-1}(B^\infty) \triangle B^\infty) = 0$, which means that by ergodicity $\mathbb{P}(B^\infty) = 0$ or $\mathbb{P}(B^\infty) = 1$. Since $\mathscr{W}_{[0,q+1)}(A) \subset (B^\infty)^c$, then the hypothesis guarantees that $\mathbb{P}(B^\infty) \neq 1$ and the conclusion follows. □

2.2. **Point processes and the extremal index.** Our goal is to keep record of the number of occurrences of A on a certain time frame and then be able to provide statements regarding its asymptotic behaviour. We will do so by considering point process theory. The asymptotics comes to play by considering events that are rarer and rarer, *i.e.*, we will consider a nested sequence of sets A_n such that $\lim_{n\to\infty} \mathbb{P}(A_n) = 0$. In fact, we will use the framework of Extreme Value Theory where $\{X_0 \in A_n\}$ corresponds to an exceedance $\{X_0 > u_n\}$ of a threshold u_n, where u_n is converging to the right hand point of the support of the distribution function of X_0 (which may be $+\infty$). We remark that there is no loss of generality in doing so because one could always find an auxiliary stochastic process $Y_0, Y_1, \ldots$ such that $\{X_0 \in A_n\} = \{Y_0 > u_n\}$ (see [11, 12]).

We assume that the sequence of levels $(u_n)_{n\in\mathbb{N}}$ satisfies the condition:

$$\lim_{n\to\infty} n\mathbb{P}(X_0 > u_n) = \tau, \tag{2.8}$$

for some $\tau > 0$. This condition is requiring that the average frequency of exceedances of the level u_n among the n first observations is asymptotically constant. Note that this, in particular, implies that $\lim_{n\to\infty} u_n = \sup\{x \in \mathbb{R} : \mathbb{P}(X_0 \le x) < 1\}$.

Let $E = [0, \infty)$. We say that m is a *point measure* on E if $m = \sum_{i=1}^{\infty} \delta_{x_i}$, where δ_{x_i} denotes the Dirac measure supported on $x_i \in E$. We say that m is *simple* if all the x_i are distinct and that m is *Radon* if $m(K) < \infty$ for all compact $K \subset E$. Consider the space $M_p(E)$ of all the Radon point measures defined on E endowed with the vague topology. A *point process* on E is just a random element on $M_p(E)$ and we will be particularly interested on the following:

$$N_n = \sum_{i=0}^{\infty} \delta_{\frac{i}{n}} \mathbf{1}_{\{X_i > u_n\}}. \tag{2.9}$$

Note that $N_n([0, 1))$ counts the number of exceedances among the first n observations of the process. Moreover, on account of (2.8), for any interval $J \subset E$, we have that $\mathbb{E}(N_n(J)) \to \tau|J|$, where $|J|$ denotes the Lebesgue measure of J.

Our main goal is to study the weak convergence of N_n. A point process N on E is the weak limit of N_n if for any finite number of intervals of the form $J_\ell = [a_\ell, b_\ell)$, with $\ell = 1, \ldots, \varsigma$, we have that the random vector $(N_n(J_1), \ldots, N_n(J_\varsigma))$ converges in distribution to $(N(J_1), \ldots, N(J_\varsigma))$ (see [19]).

We will see that under certain conditions the weak limit N is a compound Poisson process, which can be described in the following way. Let $W_1, W_2, \ldots$ be an iid sequence of exponentially distributed random variables with mean $1/\eta > 0$, *i.e.*, $W_i \sim \text{Exp}(\eta)$. Let $T_i = \sum_{j=1}^{i} W_i$ and $D_1, D_2, \ldots$ be an iid sequence of positive integer valued random variables independent of $T_1, T_2, \ldots$. Then $N = \sum_{i=1}^{\infty} D_i \delta_{T_i}$. Typically, T_i corresponds to the time of appearance of the $i - th$ cluster and D_i the respective size. We say that n is a compound Poisson process with intensity η and multiplicity distribution given by $\pi(\kappa) = \mathbb{P}(D_1 = \kappa)$.

The weak convergence of N_n gives a lot of information about the limiting behaviour of the order statistics of a finite sample of $X_0, X_1, \ldots$. In particular, if $M_n = \max\{X_0, \ldots, X_{n-1}\}$ we have $\{M_n \le u_n\} = \{N_n([0, 1)) = 0\}$. Therefore, $\lim_{n\to\infty} \mathbb{P}(M_n \le u_n) = \mathbb{P}(N([0, 1) = 0)$. When we have a compound Poisson process in the limit then $\mathbb{P}(N([0, 1) = 0) = \mathbb{P}(W_1 > 1) = \mathrm{e}^{-\eta}$. Since in most situations $\mathbb{E}(N_n([0, 1))) = \tau$ then $\eta = \tau/\mathbb{E}(D_1)$. This motivates the following definition.

Definition 2.1. Consider a sequence $(u_n)_{n\in\mathbb{N}}$ such that (2.8) holds. We say we have an Extremal Index (EI) $0 \le \theta \le 1$ if $\lim_{n\to\infty} \mathbb{P}(M_n \le u_n) = \mathrm{e}^{-\theta\tau}$.

Usually, we have that $\theta^{-1} = \mathbb{E}(D_1)$ and the EI can be interpreted as a measure of the intensity of clustering, so that $\theta = 1$ means the absence of clustering. We will build examples where this relation between the EI and $\mathbb{E}(D_1)$ does not hold anymore.

2.3. **Convergence of point processs.** In order to obtain the convergence of the point processes introduced above we will use two conditions on the dependence structure of original stochastic process $X_0, X_1, \ldots$. We introduce the notation for all $\kappa \in \mathbb{N}_0 \cup \{\infty\}$:

$$U^{(\kappa)}(u_n) := U^{(\kappa)}([u_n, \infty)), \quad Q_q^{(\kappa)}(u_n) := Q_q^{(\kappa)}([u_n, \infty)) \quad \text{and} \quad \pi_n(\kappa) := \pi_{[u_n,\infty)}(\kappa)$$

The first condition is a sort of mixing condition specially designed for this extreme analysis with applications to dynamically generated stochastic processes. It was introduced in [14].

Condition ($Д_q(u_n)^*$)**.** We say that $Д_q(u_n)^*$ holds for the sequence $X_0, X_1, \ldots$ if for any integers $t, \kappa_1, \ldots, \kappa_\zeta$, n and any intervals of the form $I_j = [a_j, b_j)$ with $a_{j+1} \geq b_j$ for all $j = 1, \ldots \zeta - 1$ and such that $a_1 \geq t$,

$$\left| \mathbb{P}\left(Q_q^{\kappa_1}(u_n) \cap \left(\cap_{j=2}^{\zeta} N_n(I_j) = \kappa_j \right) \right) - \mathbb{P}\left(Q_q^{\kappa_1}(u_n) \right) \mathbb{P}\left(\cap_{j=2}^{\zeta} N_n(I_j) = \kappa_j \right) \right| \leq \gamma(q, n, t),$$

where for each n we have that $\gamma(q, n, t)$ is nonincreasing in t and $n\gamma(q, n, t_n) \to 0$ as $n \to \infty$, for some sequence $t_n = o(n)$.

For some fixed $q \in \mathbb{N}_0$, consider the sequence $(t_n)_{n \in \mathbb{N}}$, given by condition $Д_q(u_n)$ and let $(k_n)_{n \in \mathbb{N}}$ be another sequence of integers such that

$$k_n \to \infty \quad \text{and} \quad k_n t_n = o(n). \tag{2.10}$$

Condition ($Д'_q(u_n)^*$)**.** We say that $Д'_q(u_n)^*$ holds for the sequence $X_0, X_1, X_2, \ldots$ if there exists a sequence $(k_n)_{n \in \mathbb{N}}$ satisfying (2.10) and such that

$$\lim_{n \to \infty} n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \mathbb{P}\left(Q_q^{(0)} \cap \mathcal{T}^{-j}(U^{(0)}(u_n)) \right) = 0. \tag{2.11}$$

Note that condition $Д'_q(u_n)^*$ is just condition $D^{(q+1)}(u_n)$ in the formulation of [7, Equation (1.2)]. Essentially, it is forbidding (or making very unlikely) the appearance of two clusters in a very short period of time and, in particular, motivating our assumption that the maximum waiting time between two consecutive exceedances on the same cluster is q.

Let us define for each $n \in \mathbb{N}$

$$\theta_n = \frac{\mathbb{P}(Q_q^{(0)}(u_n))}{\mathbb{P}(U^{(0)}(u_n))}, \tag{2.12}$$

which measures the proportion of realisations of $U^{(0)}(u_n))$, *i.e.*, exceedances of u_n that do not produce another exceedance in the same cluster.

If there exists $0 \leq \theta \leq 1$ such that $\theta = \lim_{n \to \infty} \theta_n$, then under conditions $Д_q(u_n)^*$ and $Д'_q(u_n)^*$ we have that $\lim_{n \to \infty} \mathbb{P}(M_n \leq u_n) = \mathrm{e}^{-\theta\tau}$ (see [13, 15]), which means that θ is the EI. This formula for the EI has already appeared in the work of O'Brien [24].

In the case where the exceedance corresponds to hitting time to a cylinder set of at least length u_n, this formula was also used in [1, 2], with q equal to the periodicity of the cylinder.

From the study developed in [14] and as noticed in [4, Appendix B], we can state the following result which applies to general stationary stochastic processes. A full proof of this result can be seen in [10].

Theorem 2.5 ([14, 10])**.** *Let $X_0, X_1, \ldots$ satisfy conditions $Д_q(u_n)^*$ and $Д'_q(u_n)^*$, where $(u_n)_{n \in \mathbb{N}}$ is such that* (2.8) *holds. Assume that the limit $\theta = \lim_{n \to \infty} \theta_n$ exists, where θ_n is as in* (2.12) *and moreover that for each $\kappa \in \mathbb{N}$, the following limit also exists*

$$\pi(\kappa) := \lim_{n \to \infty} \pi_n(\kappa) = \lim_{n \to \infty} \frac{\left(\mathbb{P}(Q_{q,0}^{\kappa-1}(u_n)) - \mathbb{P}(Q_{q,0}^{\kappa}(u_n)) \right)}{\mathbb{P}(Q_{q,0}^{0}(u_n))}. \tag{2.13}$$

Then the REPP N_n converges in distribution to a compound Poisson process with intensity $\theta\tau$ and multiplicity distribution π given by (2.13)*.*

Observe that by Theorem 2.3, for every $n \in \mathbb{N}$, if $\mathbb{P}(U^{(\infty)}(u_n)) = 0$, then the mean of the distribution $\pi_{[u_n,\infty)}$ is the reciprocal of θ_n, *i.e.*, $\sum_{\kappa=1}^{\infty} \kappa\pi_n(\kappa) = \theta_n^{-1}$. It follows that if there exists $0 \leq \theta \leq 1$ such that $\theta = \lim_{n\to\infty} \theta_n$, then $\lim_{n\to\infty} \sum_{\kappa=1}^{\infty} \kappa\pi_n(\kappa) = \theta^{-1}$.

We are going to build an example such that, although the latter equality holds, the same does not hold for the asymptotic distribution of the cluster size, *i.e.*, $\lim_{n\to\infty} \sum_{\kappa=1}^{\infty} \kappa\pi(\kappa) \neq \theta^{-1}$.

3. Dynamical counterexamples

Let us consider a one-dimensional family of maps with an indifferent fixed point of the *Manneville-Pomeau* (MP) type. We will be using the particular form given in [22]. Namely, for $\alpha > 0$,

$$T = T_\alpha(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{for } x \in [0, 1/2) \\ 2x - 1 & \text{for } x \in [1/2, 1] \end{cases} \tag{3.1}$$

If $\alpha \in (0, 1)$ then there is an absolutely continuous (w.r.t. Lebesgue) invariant probability μ_α, which is the case we will restrict to. These maps have been studied in [22, 29, 18] and, for each $\alpha \in (0, 1)$, the system $([0, 1], T_\alpha, \mu_\alpha)$ has polynomial decay of correlations. That is, letting $\mathcal{H}_\beta$ denote the space of Hölder continuous functions ϕ with exponent β equipped with the norm $\|\phi\|_{\mathcal{H}_\beta} = \|\phi\|_\infty + |\phi|_{\mathcal{H}_\beta}$, where

$$|\phi|_{\mathcal{H}_\beta} = \sup_{x \neq y} \frac{|\phi(x) - \phi(y)|}{|x - y|^\beta},$$

there exists $C > 0$ such that for each $\phi \in \mathcal{H}_\beta$, $\psi \in L^\infty$ and all $t \in \mathbb{N}$,

$$\left| \int \phi \cdot (\psi \circ T^t) d\mu_\alpha - \int \phi d\mu_\alpha \int \psi d\mu_\alpha \right| \leq C \|\phi\|_{\mathcal{H}_\beta} \|\psi\|_\infty \frac{1}{t^{\frac{1}{\alpha}-1}}. \tag{3.2}$$

Let $h_\alpha = \frac{d\mu_\alpha}{dx}$. In [18], Hu showed that $h_\alpha \in L^{1+\epsilon}$, with $\epsilon < 1/\alpha - 1$, h_α is Lipschitz on $[a, 1]$ for all $0 < a < 1$ and moreover $\lim_{x\to 0} \frac{h(x)}{x^{-\alpha}} = C_0 > 0$. Hence, for small $s > 0$ we have that

$$\mu_\alpha([0, s)) \sim C_1 s^{1-\alpha}, \tag{3.3}$$

for some $C_1 > 0$, where the notation $A(s) \sim B(s)$ is used in the sense that $\lim_{s\to 0} \frac{A(s)}{B(s)} = 1$. When the constant is unimportant, we will also use the notation $A(s) \sim_c B(s)$ in the sense that there is $c > 0$ such that $\lim_{s\to 0} \frac{A(s)}{B(s)} = c$.

Let x be such that $T_\alpha(x) = y$, *i.e.*, $y = x + 2^\alpha x^{1+\alpha}$. From the properties of the invariant density and (3.3) we get that there exists $C_1 > 0$ such that

$$\mu_\alpha([0, y)) \sim C_1(x^{1-\alpha} + (1 - \alpha)2^\alpha x + o(x)) \tag{3.4}$$

$$\mu_\alpha([0, x)) \sim C_1 x^{1-\alpha}. \tag{3.5}$$

$$\mu_\alpha([x, y)) \sim_c (1 - \alpha)2^\alpha x + o(x). \tag{3.6}$$

Our goal is to study the extremal behaviour of stochastic processes arising from such dynamical systems by considering an observable function that we will denote by $\varphi : [0,1] \to \mathbb{R} \cup \{+\infty\}$ and defining the process $X_0, X_1, \dots$ by

$$X_n = \varphi \circ T_\alpha^n, \quad \text{for all } n \in \mathbb{N}_0 \tag{3.7}$$

where T_α^n denotes the n-fold composition of T_α and T_α^0 is just the identity map. The T_α invariance of μ_α guarantees that $X_0, X_1, \dots$ is stationary.

3.1. **A dynamical emulation of Smith's example.** In this case, we are going to use the idea introduced in [5] to make a balanced mixture of a behaviour associated with an EI equal to 0 with the behaviour of an EI equal to 1. For that purpose we are going to consider that the observable function φ will be maximised at two points, namely, the point $\zeta_1 = 0$, which is an indifferent fixed point, and a point $\zeta_2 \in [1/2, 1]$, whose orbit never hits the maximal set $\mathcal{C} = \{\zeta_1, \zeta_2\}$, *i.e.*, $f_\alpha^j(\zeta_2) \notin \mathcal{C}$, $\forall j \in \mathbb{N}$. One could take for example the preperiodic point $\zeta_2 \in [1/2, 1]$ such that $f(\zeta_2) = \xi$, where ξ is the periodic point of period 2 on $[0, 1/2]$. The observable function will be designed so that the chances of starting near ζ_1 or ζ_2 are equally weighed. Note that if $\mathcal{C} = \{\zeta_1\}$, by [16, Theorem 2], we would have an EI equal to 0, while, by [16, Theorem 1], if $\mathcal{C} = \{\zeta_2\}$, the EI would be equal to 1. In this case, we will obtain an EI equal to 1/2 which is the mean of the two possible values.

We take the following observable:

$$\varphi(x) = g(C_1 \mathrm{dist}(x, \zeta_1)^{1-\alpha})\mathbf{1}_{[0,\delta)} + g(2h_\alpha(\zeta_2)\mathrm{dist}(x, \zeta_2))\mathbf{1}_{(\zeta_2-\delta, \zeta_2+\delta)}, \tag{3.8}$$

for some $\delta > 0$, where dist denotes any given metric on $[0,1]$ and the function $g : [0, +\infty) \to \mathbb{R} \cup \{+\infty\}$ is such that 0 is a global maximum ($g(0)$ may be $+\infty$); g is a strictly decreasing bijection $g : V \to W$ in a neighbourhood V of 0; and has one of the three types of behaviour described for example in [23, Section 4.2.1], which are quite general and essential guarantee that we do not fall into a case of degeneracy of the limiting law for the partial maxima of the stochastic process $X_0, X_1, \dots$.

We claim that with this particular choice of type of observable φ then the process $X_0, X_1, \dots$ has an EI that does not coincide with the reciprocal of the mean cluster size distribution of the limiting process of N_n given in (2.9).

Theorem 3.1. *Consider a a map T_α defined in* (3.1) *for some $0 < \alpha < \sqrt{5} - 2$. Let φ be as in* (3.8) *and consider the stochastic process $X_0, X_1, \dots$ defined by* (3.7)*. This process admits an EI $\theta = \frac{1}{2}$. Moreover, the point process N_n defined by* (2.9) *for such stochastic process and for a sequence of levels $(u_n)_{n\in\mathbb{N}}$ satisfying* (2.8) *converges in distribution to a Poisson process N defined on the positive real line with intensity $\theta\tau$.*

Remark 3.2. Observe that the EI obtained $\theta = 1/2$ does not coincide with the reciprocal of the mean of the cluster size of the limiting process N, which in this case is 1 because it turns out that N is actually a Poisson process.

Remark 3.3. Nevertheless, recall that by Theorem 2.3 we still have that $\theta = 1/2$ is the reciprocal of the limit of the mean cluster size of the finite time point process N_n, *i.e.*,

$$\theta^{-1} = \lim_{n\to\infty} \sum_{\kappa=1} \kappa \pi_n(\kappa).$$

In order to prove this theorem, we apply Theorem 2.5. To that end we need to check conditions $Д_q(u_n)^*$ and $Д'_q(u_n)^*$ which we leave for Sections 3.3 and 3.4, respectively. We are left to prove that θ_n given in (2.12) converges to $\theta = 1/2$ and the finite time cluster size distribution π_n given by $\pi_n(\kappa) = \pi_{[u_n,\infty)}(\kappa)$ as in (2.6) converges to a degenerate distribution π such that $\pi(1) = 0$ and $\pi(\kappa) = 0$ for all $\kappa > 1$.

Letting $B_\delta(\zeta_2) = (\zeta_2 - \delta, \zeta_2 + \delta)$, we note that

$$\{\varphi(x) > u\} = \big(\{x : C_1|x|^{1-\alpha} < g^{-1}(u)\} \cap [0,\delta)\big) \cup \big(\{x : 2h_\alpha(\zeta_2)|x - \zeta_2| < g^{-1}(u)\} \cap B_\delta(\zeta_2)\big)$$

$$= \left(\left\{|x-\zeta_1| < \left(\frac{1}{C_1}g^{-1}(u)\right)^{\frac{1}{1-\alpha}}\right\} \cap [0,\delta)\right) \cup \left(\left\{|x-\zeta_2| < \frac{1}{2h_\alpha(\zeta_2)}g^{-1}(u)\right\} \cap B_\delta(\zeta_2)\right).$$

Defining now $y_n = \left(\frac{1}{C_1}g^{-1}(u_n)\right)^{1/(1-\alpha)}$ and $\delta_n = \frac{1}{2h_\alpha(\zeta_2)}g^{-1}(u_n)$, we obtain

$$U_n := U^{(0)}(u_n) = \{\varphi(x) > u_n\} = [0, y_n) \cup B_{\delta_n}(\zeta_2)$$

and by (3.4) we have

$$\begin{aligned}\mu_\alpha(U_n) &= \mu_\alpha([0,y_n)) + \mu_\alpha([\zeta_2 - \delta_n, \zeta_2 + \delta_n])\\ &\sim C_1\left(\left(\frac{1}{C_1}g^{-1}(u_n)\right)^{1/(1-\alpha)}\right)^{1-\alpha} + g^{-1}(u_n)\\ &\sim 2g^{-1}(u_n)\end{aligned}$$

Let τ be such that

$$2g^{-1}(u_n(\tau)) = \frac{\tau}{n} \quad \text{or equivalently} \quad u_n(\tau) = g\left(\frac{\tau}{2n}\right). \tag{3.9}$$

In this case,

$$Q^{(0)}_{p,0}(u_n) = [x_n, y_n) \cup [\zeta_2 - \delta_n, \zeta_2 + \delta_n].$$

So, by (3.6) end (3.9), we obtain

$$\mu_\alpha(Q^{(0)}_{p,0}(u_n)) \sim c(1-\alpha)2^\alpha x_n + o(x_n) + \frac{\tau}{2n}. \tag{3.10}$$

Since $\mu_\alpha([0,y_n)) \sim g^{-1}(u_n) = \frac{\tau}{2n}$, then, by (3.4), we have that

$$C_1(x_n^{1-\alpha} + (1-\alpha)2^\alpha x_n + o(x_n)) \sim \frac{\tau}{2n},$$

which implies that

$$x_n = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right). \tag{3.11}$$

Then, by (3.10),

$$\mu_\alpha(Q^{(0)}_{p,0}(u_n)) = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}.$$

In this way we easily obtain

$$\theta = \lim_{n\to+\infty} \frac{\frac{\tau}{2n} + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}{\frac{\tau}{n}} = \frac{1}{2}.$$

Recall that

$$\pi_n(k) = \frac{\mu_\alpha\left(Q_{p,0}^{(k-1)}(u_n)\right) - \mu_\alpha\left(Q_{p,0}^{(k)}(u_n)\right)}{\mu_\alpha\left(Q_{p,0}^{(0)}(u_n)\right)}.$$

In this case, $Q_{p,0}^{(1)}(u_n) = [x_n^{(1)}, x_n)$, where $T_\alpha\left(x_n^{(1)}\right) = x_n$, *i.e.*, $x_n^{(1)} + 2^\alpha\left(x_n^{(1)}\right)^{1+\alpha} = x_n$, which implies that $x_n^{(1)} = O(x_n)$.

Consequently, by (3.11)

$$\pi_n(1) = \frac{\mu_\alpha\left(Q_{p,0}^{(0)}(u_n)\right) - \mu_\alpha\left(Q_{p,0}^{(1)}(u_n)\right)}{\mu_\alpha\left(Q_{p,0}^{(0)}(u_n)\right)} = \frac{O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n} - O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}{O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}},$$

which goes to 1 as n goes to ∞ and, therefore, we must have $\pi(1) = \lim_{n\to\infty} \pi_n(1) = 1$ and $\pi(\kappa) = 0$ for all $\kappa > 1$. In any case, we can also easily check that

$$\pi_n(k) = \frac{\mu_\alpha\left(Q_{p,0}^{(k-1)}(u_n)\right) - \mu_\alpha\left(Q_{p,0}^{(k)}(u_n)\right)}{\mu_\alpha\left(Q_{p,0}^{(0)}(u_n)\right)} = \frac{O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}{O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}},$$

which goes to 0 as n goes to ∞.

3.2. **Dynamical counterexample with periodic behaviour.** As in the previous example we use a maximal set $\mathcal{C} = \{\zeta_1, \zeta_2\}$ consisting of two points, where $\zeta_1 = 0$ is again the indifferent fixed point while $\zeta_2 \in [1/2, 1]$ is a periodic point, namely, for some $p \in \mathbb{N}$, we have $T_\alpha^p(\zeta_2) = \zeta_2$ and $T_\alpha^j(\zeta_2) \notin \{\zeta_1, \zeta_2\}$, $\forall j \in \{1, \ldots, p-1\}$. As proved in [13], if $\mathcal{C} = \{\zeta_2\}$, then we would have an EI $\theta = 1 - \gamma^{-1}$, where $\gamma = DT_\alpha^p(\zeta_2)$ is the derivative of T_α^p at ζ_2. Hence, in this case, we are mixing an evenly weighed EI equal to 0 with an EI equal to $1-\gamma^{-1}$. As we will prove, the EI in this counterexample will be again the average of the two, *i.e.*, $\theta = \frac{1}{2}(1-\gamma^{-1})$, which will not coincide with the reciprocal of the mean cluster size of the limiting process.

We take, as in the previous example, the following observable:

$$\varphi(x) = g(C_1 \mathrm{dist}(x, \zeta_1)^{1-\alpha})\mathbf{1}_{[0,\delta)} + g(2h_\alpha(\zeta_2)\mathrm{dist}(x, \zeta_2))\mathbf{1}_{(\zeta_2-\delta, \zeta_2+\delta)}, \tag{3.12}$$

for some $\delta > 0$ and g as described above. In this case we also have a counterexample where the EI cannot be identified as the reciprocal of the mean limiting cluster size distribution.

Theorem 3.4. *Consider a a map T_α defined in* (3.1) *for some $0 < \alpha < \sqrt{5} - 2$. Let φ be as in* (3.12) *and consider the stochastic process $X_0, X_1, \ldots$ defined by* (3.7)*. This process admits an EI $\theta = \frac{1}{2}(1-\gamma^{-1})$, where $\gamma = DT_\alpha^p(\zeta_2)$. Moreover, the point process N_n defined by* (2.9) *for such stochastic process and for a sequence of levels $(u_n)_{n\in\mathbb{N}}$ satisfying* (2.8) *converges in distribution to a compound Poisson process N defined on the positive real line with intensity $\theta\tau$ and multiplicity distribution given by*

$$\pi(\kappa) = \gamma^{-(\kappa-1)}(1-\gamma^{-1}), \qquad \textit{for all } \kappa \in \mathbb{N}. \tag{3.13}$$

Remark 3.5. Observe that the EI obtained $\theta = \frac{1}{2}(1-\gamma^{-1})$ does not coincide with the reciprocal of the mean of the cluster size of the limiting process N, which in this case is

$$\sum_{\kappa=1} \kappa\pi(\kappa) = \sum_{\kappa=1}^{\infty} \kappa\gamma^{-(\kappa-1)}(1-\gamma^{-1}) = \frac{1}{1-\gamma^{-1}}.$$

Remark 3.6. As in the previous example, recall that by Theorem 2.3 we still have that $\theta = \frac{1}{2}(1-\gamma^{-1})$ is the reciprocal of the limit of the mean cluster size of the finite time point process N_n, *i.e.*,

$$\theta^{-1} = \lim_{n\to\infty} \sum_{\kappa=1} \kappa\pi_n(\kappa).$$

Again, in order to prove this theorem, we apply Theorem 2.5. To that end we need to check conditions $Д_q(u_n)^*$ and $Д'_q(u_n)^*$ which we leave for Sections 3.3 and 3.4, respectively. We are left to prove that θ_n given in (2.12) converges to $\theta = \frac{1}{2}(1-\gamma^{-1})$ and the finite time cluster size distribution π_n given by $\pi_n(\kappa) = \pi_{[u_n,\infty)}(\kappa)$ as in (2.6) converges to π given in (3.13).

As in the previous example, defining $y_n = \left(\frac{1}{C_1} g^{-1}(u_n)\right)^{1/(1-\alpha)}$ and $\delta_n = \frac{1}{2h_\alpha(\zeta_2)} g^{-1}(u_n)$, we obtain

$$U_n := U^{(0)}(u_n) = \{\varphi(x) > u_n\} = [0, y_n) \cup (\zeta_2 - \delta_n, \zeta_2 + \delta_n) \quad \text{and} \quad \mu_\alpha(U_n) \sim 2g^{-1}(u_n).$$

Again, we let τ to be as in (3.9). In this case,

$$Q_{p,0}^{(0)}(u_n) = [x_n, y_n) \cup (B_{\delta_n}(\zeta_2) \setminus T_\alpha^{-p}(B_{\delta_n}(\zeta_2))),$$

where as before $B_{\delta_n}(\zeta_2) = (\zeta_2 - \delta_n, \zeta_2 + \delta_n)$

So, by (3.6) and (3.9), we obtain

$$\mu_\alpha(Q_{p,0}^{(0)}(u_n)) \sim c(1-\alpha)2^\alpha x_n + o(x_n) + \frac{\tau}{2n}(1-\gamma^{-1}). \tag{3.14}$$

As we have seen in the previous example, $\mu_\alpha([0, y_n)) \sim g^{-1}(u_n) = \frac{\tau}{2n}$, and then, by (3.4), we have that

$$C_1(x_n^{1-\alpha} + (1-\alpha)2^\alpha x_n + o(x_n)) \sim \frac{\tau}{2n}.$$

Hence,

$$x_n = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right). \tag{3.15}$$

Then, by (3.14),

$$\mu_\alpha(Q_{p,0}^{(0)}(u_n)) = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}(1-\gamma^{-1}). \tag{3.16}$$

Gathering this information, we obtain

$$\theta = \lim_{n\to+\infty} \frac{\frac{\tau}{2n}(1-\gamma^{-1}) + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}{\frac{\tau}{n}} = \frac{1}{2}(1-\gamma^{-1}).$$

We compute now the multiplicity distribution. Observe that

$$Q_{p,0}^{(1)}(u_n) = [x_n^{(1)}, x_n) \cup (B_{\delta_n}(\zeta_2) \cap T_\alpha^{-p}(B_{\delta_n}(\zeta_2)) \setminus T_\alpha^{-2p}(B_{\delta_n}(\zeta_2))),$$

where $T_\alpha\left(x_n^{(1)}\right) = x_n$, that is, $x_n^{(1)} + 2^\alpha\left(x_n^{(1)}\right)^{1+\alpha} = x_n$, which implies that $x_n^{(1)} = O(x_n)$. Hence,

$$\mu_\alpha\left(Q_{p,0}^{(1)}(u_n)\right) = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}\gamma^{-1}(1-\gamma^{-1}) \tag{3.17}$$

Consequently, by (3.16) and (3.17), we have

$$\begin{aligned}\pi(1) = \lim_{n\to\infty}\pi_n(1) &= \lim_{n\to\infty}\frac{\mu_\alpha\left(Q_{p,0}^{(0)}(u_n)\right) - \mu_\alpha\left(Q_{p,0}^{(1)}(u_n)\right)}{\mu_\alpha\left(Q_{p,0}^{(0)}(u_n)\right)}\\ &= \lim_{n\to\infty}\frac{\frac{\tau}{2n}(1-\gamma^{-1}) - \frac{\tau}{2n}\gamma^{-1}(1-\gamma^{-1}) + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}{\frac{\tau}{2n}(1-\gamma^{-1}) + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}\\ &= 1-\gamma^{-1}\end{aligned}$$

In order to compute $\pi(2)$ we need to estimate $\mu_\alpha\left(Q_{p,0}^{(2)}(u_n)\right)$, which we do by noting

$$Q_{p,0}^{(2)}(u_n) = [x_n^{(2)}, x_n^{(1)}) \cup (B_{\delta_n}(\zeta_2) \cap T_\alpha^{-2p}(B_{\delta_n}(\zeta_2)) \setminus T_\alpha^{-3p}(B_{\delta_n}(\zeta_2))),$$

where $T_\alpha\left(x_n^{(2)}\right) = x_n^{(1)}$, *i.e.*, $x_n^{(2)} + 2^\alpha\left(x_n^{(2)}\right)^{1+\alpha} = x_n^{(1)}$, which implies that $x_n^{(2)} = O(x_n^{(1)}) = O(x_n)$. Thus,

$$\mu_\alpha\left(Q_{p,0}^{(2)}(u_n)\right) = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}\gamma^{-2}(1-\gamma^{-1}) \tag{3.18}$$

Consequently, by (3.16), (3.17) and (3.18),

$$\begin{aligned}\pi(2) = \lim_{n\to\infty}\pi_n(2) &= \lim_{n\to\infty}\frac{\frac{\tau}{2n}\gamma^{-1}(1-\gamma^{-1}) - \frac{\tau}{2n}\gamma^{-2}(1-\gamma^{-1}) + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}{\frac{\tau}{2n}(1-\gamma^{-1}) + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}\\ &= \gamma^{-1}-\gamma^{-2} = \gamma^{-1}(1-\gamma^{-1}).\end{aligned}$$

A simple inductive argument then leads to

$$\begin{aligned}\pi(\kappa) = \lim_{n\to\infty}\pi_n(\kappa)\pi_n(\kappa) &= \frac{\mu_\alpha\left(Q_{p,0}^{(\kappa-1)}(u_n)\right) - \mu_\alpha\left(Q_{p,0}^{(\kappa)}(u_n)\right)}{\mu_\alpha\left(Q_{p,0}^{(0)}(u_n)\right)}\\ &= \lim_{n\to\infty}\frac{\frac{\tau}{2n}\gamma^{-(\kappa-1)}(1-\gamma^{-1}) - \frac{\tau}{2n}\gamma^{-\kappa}(1-\gamma^{-1}) + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}{\frac{\tau}{2n}(1-\gamma^{-1}) + O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right)}\\ &= \gamma^{-(\kappa-1)} - \gamma^{-\kappa} = \gamma^{\kappa-1}(1-\gamma^{-1}).\end{aligned}$$

for all $\kappa \in \mathbb{N}$.

3.3. **The condition $Д_q(U_n)^*$.** Condition $Д_q(U_n)^*$ has been designed to be easily verified for systems with sufficiently fast decay of correlations. The argument used in [16, Section 4.2.2] allows to show that $Д_q(U_n)^*$ follows from the decay of correlations stated in (3.2), as long as $\alpha < \sqrt{5}-2$.

3.4. **The condition $Д'_q(U_n)^*$.** This subsection is dedicated to the verification of condition $Д'_q(U_n)^*$. We need to check (2.3). We will split the argument into two parts. In the first part we consider the points from $Q^{(0)}(u_n)$ that belong to a neighbourhood of ζ_2 and in the second part the points from $Q^{(0)}(u_n)$ that belong to a neighbourhood of ζ_1. Let $A_n = Q^{(0)}(u_n) \cap [0, 1/2]$ and $B_n = Q^{(0)}(u_n) \cap [1/2, 1]$.

3.4.1. *Starting in a neighbourhood of* ζ_2. We begin with the points that start in B_n. It is well known that the map T_α admits a first return time map $F_\alpha : [1/2, 1] \to [1/2, 1]$ given by $F_\alpha(x) = T_\alpha^{r_B(x)}(x)$, where $B = [1/2, 1]$ and $r_B : B \to \mathbb{N}$ is the first return time to B, *i.e.*, $r_B(x) = \inf\{j \in \mathbb{N} : f_\alpha^j(x) \in B\}$. The map F_α has $\bar\mu_\alpha = \mu_\alpha|B$ as an invariant probability measure, is piecewise expanding and in particular qualifies as Rychlik map. Therefore, F_α has a strong form of decay of correlations, namely, there exist $C, a > 0$ such that for all bounded variation functions ϕ against all L^1 functions ψ we have

$$\left| \int \phi \cdot (\psi \circ F_\alpha^t) d\bar\mu_\alpha - \int \phi d\bar\mu_\alpha \int \psi d\bar\mu_\alpha \right| \le C \|\phi\|_{BV} \|\psi\|_1 \mathrm{e}^{-at}. \tag{3.19}$$

Let $D_n = U^{(0)}(u_n) \cap B$, $E_n = U^{(0)}(u_n) \setminus D_n$ and $\tilde E_n = T_\alpha^{-1}(E_n) \cap B$. We observe that if $x \in B_n \cap T_\alpha^{-j} D_n$ then there exists $i \le j$ such that $x \in B_n \cap F_\alpha^{-i} D_n$ and, moreover, if $x \in B_n \cap T_\alpha^{-j} E_n$ then there exists $i \le j - 1$ such that $x \in B_n \cap F_\alpha^{-i} \tilde E_n$. Therefore,

$$n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \bar\mu_\alpha \left(B_n \cap T_\alpha^{-j}(U^{(0)}(u_n)) \right) \le n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \bar\mu_\alpha \left(B_n \cap F_\alpha^{-j}(D_n \cup \tilde E_n) \right).$$

We will use decay of correlations against L^1 of the first return time induced map F_α to estimate the last quantity on the right. Let $R_n = \inf\{j \in \mathbb{N} : B_n \cap F_\alpha^{-j}(D_n \cup \tilde E_n) \neq \emptyset\}$. In both examples described in sections 3.1 and 3.2, we have that $R_n \to \infty$ as $n \to \infty$. In the first situation, this follows since $B_n = D_n$ get arbitrarily small and close to ζ_2, while E_n gets arbitrarily small and close to ζ_1 and the orbit of ζ_2 does not hit $\mathcal{C} = \{\zeta_1, \zeta_2\}$. In the second situation, it follows because B_n, D_n get arbitrarily small and close to ζ_2, while E_n gets arbitrarily small and close to ζ_1 and since ζ_2 is a repelling periodic point, by construction of B_n its points take an arbitrarily increasing amount of time before having the opportunity to return to D_n. Using this observation, (3.19), with $\phi = \mathbf{1}_{B_n}$ and $\psi = \mathbf{1}_{D_n \cap \tilde E_n}$, the facts that $\bar\mu_\alpha(D_n \cup \tilde E_n) = O\left(n^{-\frac{1}{1-\alpha}}\right) + O\left(n^{-1}\right) = O\left(n^{-1}\right)$ and $\|\phi\|_{BV} \le 6$, we have

$$\begin{aligned}
n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \bar\mu_\alpha \left(B_n \cap F_\alpha^{-j}(D_n \cup \tilde E_n) \right) &= n \sum_{j=R_n}^{\lfloor n/k_n \rfloor - 1} \bar\mu_\alpha \left(B_n \cap F_\alpha^{-j}(D_n \cup \tilde E_n) \right) \\
&\le n \sum_{j=R_n}^{\lfloor n/k_n \rfloor - 1} \bar\mu_\alpha(B_n) \bar\mu_\alpha(D_n \cup \tilde E_n) + 6Cn\bar\mu_\alpha(D_n \cup \tilde E_n) \sum_{j=R_n}^{\lfloor n/k_n \rfloor - 1} \mathrm{e}^{-aj} \\
&\le O\left(\frac{1}{k_n}\right) + O\left(\sum_{j=R_n}^{\infty} \mathrm{e}^{-aj} \right) \xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

3.4.2. *Starting in a neighbourhood of ζ_1.* Using the notation above, we start now with points in A_n. Observe that by definition of A_n and the properties of T_α, a point of $x\in A_n$ can only return to $U^{(0)}(u_n)$ after hitting the set B. Then if it hits D_n it returns to $U^{(0)}(u_n)$ or if it hits $\tilde{E}_n$, it will return in the following iterate. Otherwise, if it hits $B\setminus(D_n\cup\tilde{E}_n)$, we must wait until its orbit hits B again to have another chance of returning to $U^{(0)}(u_n)$. Hence, in order to check (2.3), we need to estimate

$$n\sum_{j=q+1}^{\lfloor n/k_n\rfloor-1}\mu_\alpha\left(A_n\cap T_\alpha^{-j}(D_n\cup\tilde{E}_n)\right)=n\sum_{j=R_n}^{\lfloor n/k_n\rfloor-1}\mu_\alpha\left(A_n\cap T_\alpha^{-j}(D_n\cup\tilde{E}_n)\right),$$

where $R_n=\inf\{j\in\mathbb{N}:A_n\cap T_\alpha^{-j}(D_n\cup\tilde{E}_n)\neq\emptyset\}$. Let $P_\alpha:L^1(\text{Leb})\to L^1(\text{Leb})$ denote the transfer or Perron-Frobenius operator given by duality from the equation

$$\int\phi\cdot\psi\circ T_\alpha\,dx=\int P_\alpha(\phi)\cdot\psi\,dx,$$

where $\phi\in L^1(\text{Leb})$ and $\psi\in L^\infty(\text{Leb})$. Now, recalling that $h_\alpha>0$, we have

$$\begin{aligned}\mu_\alpha\left(A_n\cap T_\alpha^{-j}(D_n\cup\tilde{E}_n)\right)&=\int\mathbf{1}_{A_n}\cdot\mathbf{1}_{D_n\cup\tilde{E}_n}\circ T_\alpha^j\cdot h_\alpha\,dx=\int P_\alpha^j(\mathbf{1}_{A_n}h_\alpha)\cdot\mathbf{1}_{D_n\cup\tilde{E}_n}\,dx\\&=\int\frac{P_\alpha^j(\mathbf{1}_{A_n}h_\alpha)}{h_\alpha}\cdot\mathbf{1}_{D_n\cup\tilde{E}_n}\cdot h_\alpha\,dx\leq\mu_\alpha(D_n\cup\tilde{E}_n)\sup_{x\in D_n\cup\tilde{E}_n}\frac{P_\alpha^j(\mathbf{1}_{A_n}h_\alpha)}{h_\alpha}.\end{aligned}$$

Following now the same argument used in [16, Section 4.2.1] to estimate $\frac{P_\alpha^j(\mathbf{1}_{A_n}h_\alpha)}{h_\alpha}$, with the necessary adjustments (note that here $A_n=[x_n,y_n)$ where $x_n\sim_c\frac{1}{n^{1/(1-\alpha)}}$ while in [16, Section 4.2.1] x_n was such that $x_n\sim_c\frac{1}{n}$) we obtain for some $C>0$,

$$\frac{P_\alpha^j(\mathbf{1}_{A_n}h_\alpha)}{h_\alpha}\leq C\frac{1}{n^{1/(1-\alpha)}}.$$

Therefore, recalling that $\mu_\alpha(D_n\cup\tilde{E}_n)=O\left(n^{-1}\right)$, we have

$$n\sum_{j=R_n}^{\lfloor n/k_n\rfloor-1}\mu_\alpha\left(A_n\cap T_\alpha^{-j}(D_n\cup\tilde{E}_n)\right)\leq n\frac{n}{k_n}\mu_\alpha(D_n\cup\tilde{E}_n)\frac{C}{n^{1/(1-\alpha)}}=O\left(\frac{1}{k_nn^{\alpha/(1-\alpha)}}\right)\xrightarrow[n\to\infty]{}0.$$

4. Escape of mass

We note that, in the counterexamples that we built, there exists an escape of mass, which is responsible for difference between the mean of the finite time cluster size distribution (associated to the point process N_n) and the mean of the limiting cluster size distribution (associated to the limiting process N). The loss of mass can be immediately detected by looking at the average number of rare events (exceedances of $u_n(\tau)$) recorded by both N_n and N. Indeed, observe that $\mathbb{E}(N_n([0,1))=\tau$, where τ is given by (2.8), while $\mathbb{E}(N([0,1))=\frac{1}{2}\tau$, in the case considered in Section 3.1, and $\mathbb{E}(N([0,1))=\frac{1}{2}(1-\gamma^{-1})\tau$, in the case considered in Section 3.2 (recall that $\gamma>1$). This means that the limiting processes have lost half of the mass relative to extremal events detected, in the first case, and more than half, in the second case.

The main goal of this section is to try to provide an explanation for the question: how did mass disappear?

We consider two dimensional point processes as studied in [9], namely,

$$N_n^{(2)} = \sum_{j=0}^{\infty} \delta_{\left(j/n, u_n^{-1}(X_j)\right)} \tag{4.1}$$

We are assuming that for each $n \in \mathbb{N}$, the threshold function $u_n(\tau)$ is continuous and strictly decreasing in τ. We can define the inverse function u_n^{-1}. This function can be though of as the asymptotic frequency associated to a given threshold on the. range of the r.v. X_0. This point process is defined on the bi-dimensional space $E^2 = [0, +\infty) \times [0, +\infty)$ and keeps record both of the times of occurrence of events and also of their severity, in the sense that a point with a vertical coordinate close to 0 corresponds to a severe or abnormally high observation (whose corresponding asymptotic frequency is very low, *i.e.*, very few exceedances of the corresponding threshold are expected).

The weak convergence of these point processes is a very powerful tool to obtain other results such as convergence of record point processes extremal processes, limiting laws for the maxima, which can all be settled very easily through the continuous mapping theorem and a suitable projection (see [25], for example). In particular, note that if we define $H_\tau : E^2 \to E$, by $H_\tau(t, y) = t \cdot \mathbf{1}_{[0,\tau)}(y)$ then $H_\tau(N_n^{(2)}) = N_n$.

4.1. **The regular periodic case when the EI is the reciprocal of the mean limiting cluster size distribution.** In order to understand how the mass escapes, we are going to consider first the usual case where the EI coincides with the reciprocal of the cluster size distribution. Suppose that the observable $\varphi : [0, 1] \to \mathbb{R} \cap \{+\infty\}$ is maximised at a single periodic point $\zeta \in [1/2, 1]$, *i.e.*,

$$\varphi(x) = g(|x - \zeta|), \tag{4.2}$$

where g is as above and, for definiteness ζ is the periodic point of period 2 sitting on $[1/2, 1]$. Let $\gamma = DT_\alpha(T_\alpha(\zeta)) \cdot DT_\alpha(\zeta)$. Considering a stochastic process $X_0, X_1, \ldots$ defined as in 3.7 for such φ and given a sequence $(u_n(\tau))_{n\in\mathbb{N}}$ as in (2.8), by [14], we have that there exists an EI $\theta = (1 - \gamma^{-1})$ and N_n given in (2.9) converges to a compound Poisson process N of intensity $\theta\tau$ with a geometric cluster size distribution, *i.e.*, $\pi(\kappa) = \mathbb{P}(D_i = \kappa) = \theta(1 - \theta)^{\kappa-1}$. Moreover, by [9], we have that $N_n^{(2)}$ converges weakly to

$$N^{(2)} = \sum_{i,j=1}^{\infty} \sum_{\ell=0}^{\infty} \delta_{(T_{i,j}, \gamma^\ell \cdot U_{i,j})}, \tag{4.3}$$

where the matrices $(T_{i,j})_{i,j\in\mathbb{N}}$ and $(U_{i,j})_{i,j\in\mathbb{N}}$ are mutually independent and obtained in the following way. Let $(W_{i,j})_{i,j\in\mathbb{N}}$ be a matrix of iid r.v. with common $\mathrm{Exp}(\theta)$ distribution and consider $(T_{i,j})_{i,j\in\mathbb{N}}$ given by: $T_{i,j} = \sum_{\ell=1}^{j} W_{i,\ell}$. Note that the rows of $(T_{i,j})_{i,j\in\mathbb{N}}$ are independent. Let $(U_{i,j})_{i,j\in\mathbb{N}}$ be a matrix of independent r.v. such that, for all $j \in \mathbb{N}$, the r.v. $U_{i,j} \overset{D}{\sim} \mathcal{U}_{(i-1,i]}$, *i.e.*, $U_{i,j}$ has a uniform distribution on the interval $(i - 1, i]$.

The point process $N^{(2)}$ can be described in the following way, first one obtains the points of bi-dimensional Poisson process on E^2 with $\theta \cdot \mathrm{Leb}$ as its intensity measure, where $\theta \cdot \mathrm{Leb}([a, b) \times [c, d)) = \theta(b - a)(d - c)$, and then for every such point created we put a vertical pile of points above it, such that the distance to the original point follows a geometric law, namely, their second coordinate is the original one multiplied by a power of γ. The idea is that the observations within a cluster in $N_n^{(2)}$ appear closer and closer in time and as n goes to ∞, eventually, they get to be aligned on the same vertical line for $N^{(2)}$. On the other hand, the dynamics near ζ tell us that if an orbit enters a very close neighbourhood of ζ then it gets repelled away at a rate given by γ, which explains the vertical distribution of the points. To be more precise, we note that $u_n^{-1}(z) \sim n2h_\alpha(\zeta)g^{-1}(z)$. Now, say that X_j is so large that $u_n^{-1}(X_j) = 1$, which means that the point $(j/n, 1)$ is charged by the point process $N_n^{(2)}$. Then $|T_\alpha^j(x) - \zeta| \sim \frac{1}{2h_\alpha(\zeta)n}$. Since $DT_\alpha^2(\zeta) = \gamma$, then for large n it follows that $|T_\alpha^{j+2}(x) - \zeta| \sim \frac{\gamma}{2h_\alpha(\zeta)n}$, $|T_\alpha^{j+4}(x) - \zeta| \sim \frac{\gamma^2}{2h_\alpha(\zeta)n}$ and so forth. Recalling that the points $(\frac{j+2}{n}, u_n^{-1}(X_{j+2})), (\frac{j+4}{n}, u_n^{-1}(X_{j+4})), \ldots$ will also be charged by $N_n^{(2)}$ and since by the previous computations and the form of φ we have $u_n^{-1}(X_{j+2}) \sim \gamma, u_n^{-1}(X_{j+4}) \sim \gamma^2, \ldots$, then one realises that, in the limit process $N^{(2)}$, these cluster points get vertically aligned and distributed according to the powers of γ.

Also observe that $H_\tau(N^{(2)}) = N$. In fact, exceedances of the level $u_n(\tau)$ correspond to points with second coordinate less than τ and the cluster size can be easily interpreted as the number of points in each vertical pile still below the threshold τ that project on the same time event. See Figure 1.

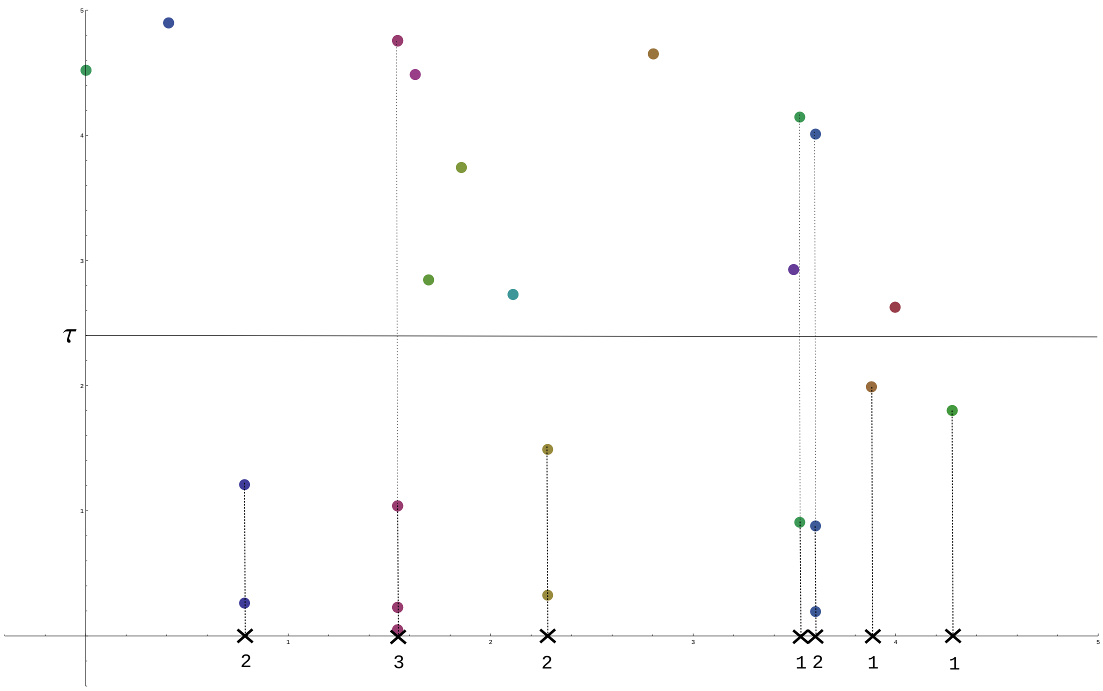

FIGURE 1. Simulation of the bi-dimensional process $N^{(2)}$ in the case of the periodic point where the observable φ is given by 4.2. The picture also shows how the projection H_τ works to obtain $N = H_\tau(N^{(2)})$, which is a compound Poisson process on the line, where the crosses represent the Poisson time events and the numbers below them the respective multiplicity (cluster size).

4.2. **The dynamical counterexample with no periodicity mixed with the indifferent fixed point.** Assume now that the observable φ is given as in (3.8). In this case the limiting process is a bi-dimensional Poisson process with intensity measure $\frac{1}{2}\cdot$Leb, which can be written as:

$$N^{(2)}=\sum_{i,j=1}^{\infty}\delta_{(T_{i,j},\,U_{i,j})},$$

where $T_{i,j}$ and $U_{i,j}$ are as above with $\theta=1/2$. Note that in this case there are no vertical piles of points as before. However, the process $N_n^{(2)}$ does have clustering points. There are two phenomena that help to explain how do they disappear.

On one hand, if we consider that X_j is a very large observation that results from the orbit entering a very small vicinity of 0 at time j. From (3.9), we have that $u_n(\tau)=g(\tau/(2n))$, which implies that $u_n^{-1}(z)=2ng^{-1}(z)$. For definiteness, let us assume that $u_n^{-1}(X_j)=1$, which means that the point $(j/n,1)$ is charged by the point process $N_n^{(2)}$. Moreover, since $\zeta_1=0$ is an indifferent fixed point then the orbit will linger around 0 for a long time which creates clustering and the points $\left(\frac{j+1}{n},u_n^{-1}(X_{j+1})\right)$, $\left(\frac{j+2}{n},u_n^{-1}(X_{j+2})\right)$, ..., which are also charged by $N_n^{(2)}$, will still be close to $(j/n,1)$. As in the previous example, the points on the same cluster will end up vertically aligned because of the horizontal contraction caused by the normalisation consisting on dividing by n. However, in this case, something interestingly different occurs in the vertical direction. Namely, since $X_j\sim g(C_1(T_\alpha(x))^{1-\alpha})$, then $u_n^{-1}(X_j)\sim 2nC_1(T_\alpha^j(x))^{1-\alpha}$, which in turn implies that $T_\alpha^j(x)\sim\left(\frac{1}{2nC_1}\right)^{\frac{1}{1-\alpha}}$. Now, observe that

$$u_n^{-1}(X_{j+1})\sim 2nC_1\left(\left(\frac{1}{2nC_1}\right)^{\frac{1}{1-\alpha}}+2^\alpha\left(\frac{1}{2nC_1}\right)^{\frac{1+\alpha}{1-\alpha}}\right)^{1-\alpha}\sim 1+(1-\alpha)2^\alpha\left(\frac{1}{2nC_1}\right)^{\frac{\alpha+\alpha^2}{1-\alpha}}\sim 1.$$

Similarly, we obtain that $u_n^{-1}(X_{j+2})\sim 1$ and so on. Therefore, not only the points of the same cluster get vertically aligned but they also get horizontally aligned, *i.e.*, they collapse to a single point in $N^{(2)}$. So these clusters collapse to one point. On the other hand the appearance of a cluster becomes less and less frequent since the mass concentrated at each cluster (which collapses to one point in the limit) is growing and must be compensated by a smaller and smaller frequency so that the mean of the mass in the clusters observed in $N_n^{(2)}$ below the threshold τ is approximately $\tau/2$. (Recall that the remaining $\tau/2$ correspond to the mass points associated with entrances near ζ_2 for which there is no clustering). In fact, the frequency of clusters of exceedances above $u_n(\tau)$ observed in $N_n^{(2)}$ is of the order of $\frac{\mu_\alpha([x_n,y_n))}{\mu_\alpha(U_n)}\frac{\tau}{2}$ which is becomes negligible when compared to the mean frequency $\tau/2$ corresponding to the exceedances with no clustering coming from entrances near ζ_2. In the limit their asymptotic time frequency is actually 0. Hence, in $N^{(2)}$ we only observe the contribution from the entrances near ζ_2. This explains the loss of half of the mass.

4.3. **The dynamical counterexample with a periodic point mixed with the indifferent fixed point.** For the observable φ given in (3.12), the limiting bi-dimensional process process can be written as:

$$N^{(2)}=\sum_{i,j=1}^{\infty}\sum_{\ell=0}^{\infty}\delta_{(T_{i,j},\,\gamma^\ell\cdot U_{i,j})},$$

where $T_{i,j}$ and $U_{i,j}$ are as above with $\theta = \frac{1}{2}(1-\gamma^{-1})$. Recall that $T_{i,j}$ are defined as sums of the waiting times $W_{i,j}$ which follow an $\mathrm{Exp}(\theta)$ distribution. In this case, we have two types of clustering of exceedances observed in $N_n^{(2)}$, namely the ones corresponding to entrances in U_n near $\zeta_1 = 0$ and entrances near the periodic point ζ_2. As in the previous case, the first type of clusters collapse to one point and since their asymptotic frequency is 0, the limiting process $N^{(2)}$ does not show any sign of their appearance. In fact, in $N^{(2)}$ one can only detect the presence of the second type of clusters, which are identical to the ones described in the periodic case example, except for the fact that their asymptotic frequency is half of what one would see in the periodic case.

References

[1] M. Abadi. Hitting, returning and the short correlation function. *Bull. Braz. Math. Soc. (N.S.)*, 37(4):593–609, 2006.

[2] M. Abadi, L. Cardeño, and S. Gallo. Potential well spectrum and hitting time in renewal processes. *J. Stat. Phys.*, 159(5):1087–1106, 2015.

[3] M. Abadi, A. C. M. Freitas, and J. M. Freitas. Clustering indeces and decay of correlations in non-Markovian models. in preparation, 2018.

[4] H. Aytaç, J. M. Freitas, and S. Vaienti. Laws of rare events for deterministic and random dynamical systems. *Trans. Amer. Math. Soc.*, 367(11):8229–8278, 2015.

[5] D. Azevedo, A. C. M. Freitas, J. M. Freitas, and F. B. Rodrigues. Clustering of extreme events created by multiple correlated maxima. *Phys. D*, 315:33–48, 2016.

[6] D. Azevedo, A. C. M. Freitas, J. M. Freitas, and F. B. Rodrigues. Extreme Value Laws for Dynamical Systems with Countable Extremal Sets. *J. Stat. Phys.*, 167(5):1244–1261, 2017.

[7] M. R. Chernick, T. Hsing, and W. P. McCormick. Calculating the extremal index for a class of stationary sequences. *Adv. in Appl. Probab.*, 23(4):835–850, 1991.

[8] T. Downarowicz and Y. Lacroix. The law of series. *Ergodic Theory Dynam. Systems*, 31(2):351–367, 2011.

[9] A. C. M. Freitas, J. M. Freitas, and M. Magalhães. Complete convergence and records for dynamically generated stochastic processes,. Preprint arXiv:1707.07071, July 2017.

[10] A. C. M. Freitas, J. M. Freitas, and M. Magalhães. Convergence of marked point processes of excesses for dynamical systems. *J. Eur. Math. Soc. (JEMS)*, 20(9):2131–2179, September 2018.

[11] A. C. M. Freitas, J. M. Freitas, and M. Todd. Hitting time statistics and extreme value theory. *Probab. Theory Related Fields*, 147(3-4):675–710, 2010.

[12] A. C. M. Freitas, J. M. Freitas, and M. Todd. Extreme value laws in dynamical systems for non-smooth observations. *J. Stat. Phys.*, 142(1):108–126, 2011.

[13] A. C. M. Freitas, J. M. Freitas, and M. Todd. The extremal index, hitting time statistics and periodicity. *Adv. Math.*, 231(5):2626–2665, 2012.

[14] A. C. M. Freitas, J. M. Freitas, and M. Todd. The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics. *Comm. Math. Phys.*, 321(2):483–527, 2013.

[15] A. C. M. Freitas, J. M. Freitas, and M. Todd. Speed of convergence for laws of rare events and escape rates. *Stochastic Process. Appl.*, 125(4):1653–1687, 2015.

[16] A. C. M. Freitas, J. M. Freitas, M. Todd, and S. Vaienti. Rare events for the Manneville-Pomeau map. *Stochastic Process. Appl.*, 126(11):3463–3479, 2016.

[17] T. Hsing, J. Hüsler, and M. R. Leadbetter. On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields*, 78(1):97–112, 1988.

[18] H. Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Ergodic Theory Dynam. Systems*, 24(2):495–524, 2004.

[19] O. Kallenberg. *Random measures*. Akademie-Verlag, Berlin, fourth edition, 1986.

[20] G. Keller. Rare events, exponential hitting times and extremal indices via spectral perturbation. *Dyn. Syst.*, 27(1):11–27, 2012.

[21] M. R. Leadbetter. Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete*, 65(2):291–306, 1983.

[22] C. Liverani, B. Saussol, and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems*, 19(3):671–685, 1999.
[23] V. Lucarini, D. Faranda, A. C. M. Freitas, J. M. Freitas, M. Holland, T. Kuna, M. Nicol, and S. Vaienti. *Extremes and Recurrence in Dynamical Systems*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, Hoboken, NJ, 2016.
[24] G. L. O'Brien. Extreme values for stationary and Markov sequences. *Ann. Probab.*, 15(1):281–291, 1987.
[25] S. I. Resnick. *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987.
[26] C. Y. Robert. Automatic declustering of rare events. *Biometrika*, 100(3):587–606, 2013.
[27] M. R. Rychlik. Another proof of Jakobson's theorem and related results. *Ergodic Theory Dynam. Systems*, 8(1):93–109, 1988.
[28] R. L. Smith. A counterexample concerning the extremal index. *Adv. in Appl. Probab.*, 20(3):681–683, 1988.
[29] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999.

Miguel Abadi, Instituto de Matemática e Estatística, Universidade de S. Paulo, Rua do Matão 1010, Cid. Universitaria, 05508090 - São Paulo, SP - Brasil

E-mail address: `leugim@ime.usp.br`

URL: `http://miguelabadi.wixsite.com/miguel-abadi`

Ana Cristina Moreira Freitas, Centro de Matemática & Faculdade de Economia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal

E-mail address: `amoreira@fep.up.pt`

URL: `http://www.fep.up.pt/docentes/amoreira/`

Jorge Milhazes Freitas, Centro de Matemática & Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

E-mail address: `jmfreita@fc.up.pt`

URL: `http://www.fc.up.pt/pessoas/jmfreita/`