

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 (EI) 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 EI for which that equality does not hold. The mechanism used to build such counterexamples is based on considering observable functions maximised 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 EI equal to 0 with that of an EI 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 cluster size distribution introduced earlier and prove that ergodicity is enough to establish that the finite versions of the reciprocal of the EI and of the mean of the cluster size distribution do coincide.

Contents

1. Introduction	1
2. Extremal analysis of stationary stochastic processes	3
3. Dynamical counterexamples	13
4. Escape of mass	20
References	23

1. Introduction

In Extreme Value Theory, the study of rare events is tied with the observation of abnormally high values among a series of realisations of a certain variable of interest, X_0, X_1, \dots , that is, one is interested in events of the form $\{X_i > u\}$, corresponding to the exceedance of a high threshold u . The choice of the level u is usually made according to the size n of the available sample, so that the average number of exceedances is asymptotically a constant, $\tau \geq 0$, as $n \rightarrow \infty$ (see (2.9)). A classical problem in this setting is the study of the asymptotic distribution of the partial maximum $M_n := \max\{X_0, \dots, X_{n-1}\}$.

For dependent data, the exceedances may present a tendency to cluster, that is, to appear in groups rather than scattered along the time line. There is then an important parameter that quantifies the intensity of clustering of extreme events. This parameter, which we will

Received 9 October 2018; revised 23 September 2019.

2010 *Mathematics Subject Classification* 37A50, 37B20 (primary), 37A25, 60G55, 60G70 (secondary).

All authors were partially supported by the joint project FAPESP (SP-Brazil) and FCT (Portugal) with reference FAPESP/19805/2014. A.C.M.F. and J.M.F. were partially supported by FCT projects PTDC/MAT-CAL/3884/2014 and PTDC/MAT-PUR/28177/2017, with national funds, and by CMUP (UIDB/00144/2020), which is funded by FCT with national (MCTES) and European structural funds through the programs FEDER, under the partnership agreement PT2020. MA is partially supposed by CNPq grant PO 305096/2019-2.

© 2020 The Authors. The publishing rights in this article are licensed to the London Mathematical Society under an exclusive licence.

denote by θ , was called extremal index (EI) by Leadbetter in [23], and is usually defined from the asymptotic distribution of M_n , namely, from the limit: $\lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) = e^{-\theta\tau}$ (see Definition 2.9). The EI 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 exceedances on a normalised time frame. Although, we defer to Subsection 2.2 its formal definition, we advance that, for a given size sample, n , and up to a given instant, t , the point processes give us the quantity

$$N_n(t) = \sum_{i=0}^{\lceil nt \rceil - 1} \mathbf{1}_{\{X_i > u_n\}}.$$

In [19], 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, that is, in particular, $N_n(t)$ was proved to converge in distribution to $N(t) = \sum_{i=1}^{N^*(t)} D_i$, where $N^*(t)$ is a Poisson random variable of mean $\theta\tau t$ independent of D_1, D_2, \dots , which is an independent and identically distributed (iid) sequence of positive integer valued random variables (see formal definition in Subsection 2.2). Under some regularity conditions, the EI can be identified as the inverse of the mean cluster size, that is, $\theta^{-1} = \mathbb{E}(D_1)$ is the average of the multiplicity distribution of the limiting compound Poisson process. When $\theta = 1$, the cluster size is 1 almost surely (a.s.) ($D_1 = 1$ a.s.) and the limiting process is a simple Poisson process.

However, in [30] a counterexample was given where the EI does not coincide with the inverse of the mean cluster size of the limiting compound Poisson process, that is, $\theta^{-1} \neq \mathbb{E}(D_1)$. 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 [29, Theorem 3.1].

More recently, a theory of extreme values for dynamical systems has been developed (see [25] and references therein). The idea is to consider stochastic processes arising from dynamical systems by evaluating a given observable along the orbits of that system. Namely, letting $T : \mathcal{X} \rightarrow \mathcal{X}$ be a discrete time dynamical system and $\varphi : \mathcal{X} \rightarrow \mathbb{R}$ a measurable function defined on the probability space $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu)$, (where μ is T -invariant), then we define X_0, X_1, \dots by $X_n = \varphi \circ T^n$, where T^n denotes the n -fold composition of T with itself. This observable φ is typically maximised at a single point ζ chosen in the phase space \mathcal{X} and then, as observed in [13], the study of the occurrence of extreme events is related to problems of entrance and recurrence times. In [15], 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 [16], the authors showed that at periodic points the Rare Events Point Process (REPP) converge to a compound Poisson process with a geometric multiplicity distribution of average θ^{-1} , that is, $\mathbb{P}(D_1 = j) = \theta(1 - \theta)^{j-1}$. Moreover, for sufficiently regular systems, a full dichotomy exists (see [4, 22]), that is, either ζ is periodic and we have clustering or ζ is non-periodic and we have the absence of clustering with $\theta = 1$ and a 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. In this case, the maximal points need not to be periodic but the maximal set that they form, that is, the set of points where the observable attains the global maximum of the observable φ , is periodic in the sense that it recurs to itself after a finite number of iterations. 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 [18]). 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 formulae for EI given in [15], the formulae for the multiplicity distributions given in [4, 16], the dynamics of the Manneville–Pomeau (MP) map and also some tools from [18].

We remark that, in the counterexample built by Smith in [30], one can show that the regenerative process is also combining the behaviour of an EI equal to 0 and an EI equal to 1, which we defer to [3]. Moreover, in all counterexamples, the EI still coincides with the reciprocal of the limit of the means of the finite time cluster size distributions. 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 1-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 [4, 16] and relate it to the one used by Robert in [28]. 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, \dots 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 $V_n : \mathcal{R}^{\mathbb{N}_0} \rightarrow \mathcal{R}$, with $V_n(x_0, x_1, \dots) = x_n$, for $n \in \mathbb{N}_0$, so that there is a natural measurable map, the shift operator $\mathcal{T} : \mathcal{R}^{\mathbb{N}_0} \rightarrow \mathcal{R}^{\mathbb{N}_0}$, given by $\mathcal{T}(x_0, x_1, \dots) = (x_1, x_2, \dots)$, which when applied later in the dynamical systems context can be identified with T .

Observe

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

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 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.

There are two main methods to identify clusters. One is called the *block declustering scheme* and the other one is the *runs declustering scheme* (see [9, 31]). The block method splits the n observations into k_n (or $k_n + 1$) blocks of length $\lfloor n/k_n \rfloor$ and then establishes that any extreme events within the same block belong to the same cluster. The runs declustering scheme consists in setting a run length, q , and establishing that any rare events separated by at most $q - 1$ non-extreme observations must belong to the same cluster, so that between two distinct clusters there must be a run of at least q non-extreme observations.

We are going to assume a runs declustering scheme and, therefore, we consider 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. In the applications, in order to prove the convergence of the point processes we will consider a condition on the dependence structure of the stochastic process called $\mathcal{D}'_q(u_n)^*$ (see Subsection 2.3), which when satisfied implies that the two declustering approaches behave essentially in the same way (see Remark 2.10).

The choice of the run length q is quite sensitive and we will return to the subject in Subsection 2.3 when we introduce and discuss condition $\mathcal{D}'_q(u_n)^*$. However, we advance here the following interpretation for the value q . In [15], we introduced the EI in the dynamical setting and established a relation between the appearance of clustering and the existence of underlying periodic phenomena in the structure of the stochastic process. In fact, in the dynamical context, as observed, in [15] and in the subsequent papers [5, 6, 12], clustering is directly related with the periodicity of the maximal set \mathcal{M} , that is, the set of points where the observable φ achieves the global maximum. Hence, q can be interpreted as the largest of the periods of the underlying periodic phenomena present in the stochastic process.

In order to illustrate the appearance of clustering and suitable choices for q , we give some simple examples of stationary stochastic processes, some of them arising from dynamical systems.

EXAMPLE 2.1. A Maximum Moving Average process with period 2.

Let $Y_{-2}, Y_{-1}, Y_0, Y_1, \dots$ be a sequence of iid random variables with common continuous distribution function G . We define a Maximum Moving Average process X_0, X_1, \dots based on the previous sequence in the following way: for each $n \in \mathbb{N}_0$ set

$$X_n = \max\{Y_{n-2}, Y_n\}.$$

Without going into too much detail (for which we refer to [15, Appendix A]), in this case it is clear that we have an underlying periodic phenomenon of period 2. To see this, observe that if $X_0 > u$, for some large u , then there is a very good chance that $X_2 > u$, as well. In fact, in this case, clusters will be separated by at least two observations below the reference threshold and the cluster size is 2 a.s. (see end of Subsection 2.3).

EXAMPLE 2.2. A dynamically generated process based on a periodic repelling point.

Consider the discrete time dynamical system $T : [0, 1] \rightarrow [0, 1]$ given by $T(x) = 2x \bmod 1$, equipped with Lebesgue measure on the borelean subsets of $[0, 1]$ and let $\varphi : [0, 1] \rightarrow \mathbb{R} \cup \{+\infty\}$ be such that $\varphi(x) = -\log|x - 1/3|$. Define X_0, X_1, \dots by $X_n = \varphi \circ T^n$. This particular example was studied in [25, Example 4.2.1]. Observe that $\{X_0 > u\} = (1/3 - e^{-u}, 1/3 + e^{-u})$ and $\zeta = 1/3$ is a repelling periodic point of period 2, since $T(\zeta) = 2/3$ and $T^2(\zeta) = 1/3 = \zeta$. This means

that if our orbit starts very close to $1/3$, we will observe a sequence of exceedances at even time steps until eventually the orbit leaves the set $(1/3 - e^{-u}, 1/3 + e^{-u})$ and then, typically, we have to wait a long time (because we are assuming that u is large) until the system brings it back again to that set in order to observe a new cluster of exceedances. In this case, it is also clear that one should take $q = 2$.

EXAMPLE 2.3. A dynamically generated process based on two periodic repelling points.

Consider the same discrete time dynamical system of the previous example and let $\varphi : [0, 1] \rightarrow \mathbb{R} \cup \{+\infty\}$ be such that $\varphi(x) = -\log|x - 1/3|\mathbf{1}_{[0, 1/2]} - \log|x - 5/7|\mathbf{1}_{(1/2, 1]}$. Again, define X_0, X_1, \dots by $X_n = \varphi \circ T^n$. Observe that $\{X_0 > u\} = (1/3 - e^{-u}, 1/3 + e^{-u}) \cup (5/7 - e^{-u}, 5/7 + e^{-u})$ and $\zeta_1 = 1/3$ is a repelling periodic point of period 2, while $\zeta_2 = 5/7$ is a repelling periodic point of period 3. Hence, if we start with an exceedance, then either we observe a cluster of exceedances observed at even time steps (if one starts very close to ζ_1) or a cluster of exceedances observed at time steps which are multiples of 3 (if one starts very close to ζ_2). In this case, in order to be sure that a cluster has ended, one has to observe a run of at least 3 non-exceedances, which means that $q = 3$.

Having fixed a run length $q \in \mathbb{N}$, we define the sequence of nested sets $(U^{(\kappa)}(A))_{\kappa \geq 0}$ of $\mathcal{B}^{\mathbb{N}_0}$ given by

$$\begin{aligned} U^{(0)}(A) &= V_0^{-1}(A) \\ Q_q^{(0)}(A) &= U^{(0)}(A) \cap \bigcap_{i=1}^q \mathcal{T}^{-i}((U^{(0)}(A))^c), \end{aligned}$$

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

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

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

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

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 . The event $Q_q^{(\kappa-1)}(A) = U^{(\kappa-1)}(A) \setminus U^{(\kappa)}(A)$ corresponds to observing A exactly κ times within no more than q units of time between one and the next observation of A . This means, in particular, that the $\kappa + 1$ th observation of A occurs at least $q + 1$ iterations after the κ th observation of A . 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

$$\begin{aligned} h : \quad \mathcal{R}^{\mathbb{N}_0} &\rightarrow \{0, 1\}^{\mathbb{N}_0} \\ \underline{x} = (x_0, x_1, \dots) &\mapsto \underline{\omega} = \omega_0 \omega_1 \dots \end{aligned} \quad (2.4)$$

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 $\underline{x} = (x_0, x_1, \dots) \in U^{(\infty)}(A)$, then $h(\underline{x})$ is a binary sequence, which starts with a 1 and has no block of more than $q - 1$ consecutive 0. Let J be an interval contained in $[0, \infty)$. We define

$$\mathcal{W}_J(A) := \bigcap_{i \in J \cap \mathbb{N}_0} \mathcal{T}^{-i}(V_0^{-1}(A^c)). \quad (2.5)$$

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

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

This formula for the finite time cluster size distribution was used first in [16] 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 $\mathcal{D}'_q(u_n)^*$ inspired in condition $D'_p(u_n)^*$ from [16], 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 $(x_0, x_1, \dots) \in \mathcal{R}^{\mathbb{N}_0}$ is a realisation of X_0, X_1, \dots , then the beginning of cluster and the ending of cluster can be easily identified in $h(x_0, x_1, \dots)$ by the appearance of a block of at least q consecutive 0. Let $q, \kappa \in \mathbb{N}$ be fixed and consider the set of finite strings of 0 and 1 such that each string starts and ends with a 1, has exactly κ 1, which are separated by at most $q-1$ 0, that is, there is no block of q or more consecutive 0 in the string. Namely, let

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

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

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

and also set

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

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 $0011010110100 \dots \in h(\mathcal{H}_2(6))$ and $0001001011000 \dots \in h(\mathcal{H}_3(4))$.

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

THEOREM 2.4. *Given $q \in \mathbb{N}$, consider the distribution π_A given by (2.6). We can write*

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

REMARK 2.5. We note that the formula on the right-hand side of (2.7) can be identified precisely as the distribution of C_r^{lk} considered in [28] 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 $(\mathcal{T}^{-i}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))) \cap (\mathcal{T}^{-j}(Q_q^{(\kappa)}(A)) \cap \mathcal{T}^{-q}(Q_q^{(\kappa-1)}(A))) = \emptyset$, for all $0 \leq i \neq j \leq q-1$. To see this, assume without loss of generality that $0 \leq i < j \leq q-1$ and take $\underline{x} = (x_0, x_1, \dots) \in \mathcal{R}^{\mathbb{N}_0}$ 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 $\ell = i+1, \dots, q-1$, 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

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

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

$$\begin{aligned} Q_q^{(\kappa)}(A) &= h^{-1}(\{\underline{\omega} \in \{0, 1\}^{\mathbb{N}_0} : \underline{\omega} = \varpi \dots, \text{ for some } \varpi \in W_q(\kappa+1)\}) \\ &= \bigcup_{i=0}^{q-1} h^{-1} \left(\left\{ \underline{\omega} \in \{0, 1\}^{\mathbb{N}_0} : \underline{\omega} = 1 \underbrace{0 \dots 0}_{i \text{ symbols}} \varpi \dots, \text{ for some } \varpi \in W_q(\kappa) \right\} \right) \\ &= \bigcup_{j=1}^q Q_q^{(\kappa)}(A) \cap \mathcal{T}^{-j}(Q_q^{(\kappa-1)}(A)). \end{aligned} \tag{2.8}$$

It follows that

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

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

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

REMARK 2.6. Note that from (2.8), we have $Q_q^{(\kappa)}(A) \subset \mathcal{T}^{-j}(Q_q^{(\kappa-1)}(A))$, which by stationarity implies $\mathbb{P}(Q_q^{(\kappa)}(A)) \leq \mathbb{P}(Q_q^{(\kappa-1)}(A))$.

From 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 EI.

THEOREM 2.7. *Let $q \in \mathbb{N}$ and consider $U^{(\infty)}(A)$ defined as in (2.3) and the distribution π_A given by (2.6). 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

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

It follows that

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

COROLLARY 2.8. *If \mathcal{T} is ergodic with respect to \mathbb{P} and $\mathbb{P}(\mathcal{W}_{[0,q+1)}(A)) > 0$, 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 $\mathcal{W}_{[0,q+1)}(A) \subset (B^{\infty})^c$, then the hypothesis guarantees that $\mathbb{P}(B^{\infty}) \neq 1$ and the conclusion follows. \square

2.2. Point processes and the EI

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, that is, we will consider a nested sequence of sets A_n such that $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = 0$. 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, \dots such that $\{X_0 \in A_n\} = \{Y_0 > u_n\}$ (see [13, 14]).

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

$$\lim_{n \rightarrow \infty} n \mathbb{P}(X_0 > u_n) = \tau, \quad (2.9)$$

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 $\lim_{n \rightarrow \infty} u_n = \sup\{x \in \mathbb{R} : \mathbb{P}(X_0 \leq x) < 1\}$.

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:

$$N_n = \sum_{i=0}^{\infty} \delta_{\frac{i}{n}} \mathbf{1}_{\{X_i > u_n\}}. \quad (2.10)$$

Note that $N_n([0, 1])$ counts the number of exceedances among the first n observations of the process. Moreover, on account of (2.9), for any interval $J \subset E$, we have $\mathbb{E}(N_n(J)) \rightarrow \tau|J|$, 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 $\ell = 1, \dots, \varsigma$, we have that the random vector $(N_n(J_1), \dots, N_n(J_\varsigma))$ converges in distribution to $(N(J_1), \dots, N(J_\varsigma))$ (see [21]).

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, \dots be an iid sequence of exponentially distributed random variables with mean $1/\eta > 0$, that is, $W_i \sim \text{Exp}(\eta)$. Let $T_i = \sum_{j=1}^i W_j$ and D_1, D_2, \dots be an iid sequence of positive integer valued random variables independent of T_1, T_2, \dots . 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, \dots . In particular, if $M_n = \max\{X_0, \dots, X_{n-1}\}$ we have $\{M_n \leq u_n\} = \{N_n([0, 1]) = 0\}$. Therefore, $\lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) = \mathbb{P}(N([0, 1]) = 0)$. When we have a compound Poisson process in the limit then $\mathbb{P}(N([0, 1]) = 0) = \mathbb{P}(W_1 > 1) = e^{-\eta}$. Since in most situations $\mathbb{E}(N_n([0, 1])) = \tau$ then $\eta = \tau/\mathbb{E}(D_1)$. This motivates the following definition.

DEFINITION 2.9. Consider a sequence $(u_n)_{n \in \mathbb{N}}$ such that (2.9) holds. We say we have an EI $0 \leq \theta \leq 1$ if $\lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) = e^{-\theta\tau}$.

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 processes

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, \dots . 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 [16].

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

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

where for each n we have that $\gamma(q, n, t)$ is nonincreasing in t and $n\gamma(q, n, t_n) \rightarrow 0$ as $n \rightarrow \infty$, 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 $\mathcal{D}_q(u_n)$ and let $(k_n)_{n \in \mathbb{N}}$ be another sequence of integers such that

$$k_n \rightarrow \infty \quad \text{and} \quad k_n t_n = o(n). \quad (2.11)$$

CONDITION $(\mathcal{D}'_q(u_n)^*)$. We say that $\mathcal{D}'_q(u_n)^*$ holds for the sequence X_0, X_1, X_2, \dots if there exists a sequence $(k_n)_{n \in \mathbb{N}}$ satisfying (2.11) and such that

$$\lim_{n \rightarrow \infty} n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \mathbb{P}\left(Q_q^{(0)}(u_n) \cap \mathcal{T}^{-j}(U^{(0)}(u_n))\right) = 0. \quad (2.12)$$

Note that condition $\mathcal{D}'_q(u_n)^*$ is just condition $D^{(q+1)}(u_n)$ in the formulation of [7, Equation (1.2)].

REMARK 2.10. Observe that condition $\mathcal{D}'_q(u_n)^*$ is forbidding (or making very unlikely) the appearance of two clusters in a very short period of time (namely, within a block of size $\lfloor n/k_n \rfloor$), which means that the two declustering schemes identify clusters essentially in the same way. Moreover, in some sense, one could say that a time gap of length q inhibits the possibility of the underlying periodic phenomena creating a new exceedance that should be classified as belonging to the same cluster as the preceding exceedance because the period time has been exhausted.

REMARK 2.11. Note that if condition $\mathcal{D}'_q(u_n)^*$ holds for some particular $q = q_0 \in \mathbb{N}_0$, then condition $\mathcal{D}'_q(u_n)^*$ holds for all $q \geq q_0$. This suggests that in trying to prove the convergence of REPP, one should try the values $q = q_0$ until we find the smallest one that makes $\mathcal{D}'_q(u_n)^*$ hold, establishing, in this way, the run length. In the dynamical context, the following procedure has proved very useful to find the value of q (see [5]). Let $\underline{x} \in \mathcal{R}^{\mathbb{N}_0}$ and recall the definition of the first hitting time to $A \in \mathcal{B}$, $r_A(\underline{x}) = \min\{j \in \mathbb{N} \cup \{+\infty\} : \mathcal{T}^j(\underline{x}) \in V_0^{-1}(A)\}$. The restriction of the function r_A to A is called the *first return time function* to A . We define the *first return time* to A , which we denote by $R(A)$, as the infimum of the return time function to A , that is, the non-negative integer $R(A) = \inf_{\underline{x} \in V_0^{-1}(A)} r_A(\underline{x})$. Assume that there exists $q \in \mathbb{N}_0$ such that

$$q := \min \left\{ j \in \mathbb{N}_0 : \lim_{n \rightarrow \infty} R(Q_q^{(0)}(u_n)) = \infty \right\}. \quad (2.13)$$

Then such q is the natural candidate to try to show the validity of $\mathcal{D}'_q(u_n)^*$.

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

$$\theta_n = \frac{\mathbb{P}(Q_q^{(0)}(u_n))}{\mathbb{P}(U^{(0)}(u_n))}, \quad (2.14)$$

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

If there exists $0 \leq \theta \leq 1$ such that $\theta = \lim_{n \rightarrow \infty} \theta_n$, then under conditions $\mathcal{D}_q(u_n)^*$ and $\mathcal{D}'_q(u_n)^*$ we have $\lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) = e^{-\theta\tau}$ (see [15, 17]), which means that θ is the EI. This formula for the EI has already appeared in the work of O'Brien [26].

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 [16] 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 2.12 [10, 16]. Let X_0, X_1, \dots satisfy conditions $\mathbb{D}_q(u_n)^*$ and $\mathbb{D}'_q(u_n)^*$, where $(u_n)_{n \in \mathbb{N}}$ is such that (2.9) holds. Assume that the limit $\theta = \lim_{n \rightarrow \infty} \theta_n$ exists, where θ_n is as in (2.14) and moreover that for each $\kappa \in \mathbb{N}$, the following limit also exists

$$\pi(\kappa) := \lim_{n \rightarrow \infty} \pi_n(\kappa) = \lim_{n \rightarrow \infty} \frac{\left(\mathbb{P}(Q_q^{(\kappa-1)}(u_n)) - \mathbb{P}(Q_q^{(\kappa)}(u_n)) \right)}{\mathbb{P}(Q_q^{(0)}(u_n))}. \quad (2.15)$$

Then the REPP N_n converges in distribution to a compound Poisson process with intensity $\theta\tau$ and multiplicity distribution π given by (2.15).

Observe that by Theorem 2.7, 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 , that is, $\sum_{\kappa=1}^{\infty} \kappa \pi_n(\kappa) = \theta_n^{-1}$. It follows that if there exists $0 \leq \theta \leq 1$ such that $\theta = \lim_{n \rightarrow \infty} \theta_n$, then $\lim_{n \rightarrow \infty} \sum_{\kappa=1}^{\infty} \kappa \pi_n(\kappa) = \theta^{-1}$.

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, that is, $\sum_{\kappa=1}^{\infty} \kappa \pi(\kappa) \neq \theta^{-1}$.

Before we do it, we revisit the examples introduced earlier in order to illustrate the usual behaviour.

Example 2.1 revisited

Note that by definition of the sequence $(u_n)_{n \in \mathbb{N}}$, we must have $\mathbb{P}(X_0 > u_n) \rightarrow \tau \geq 0$, as $n \rightarrow \infty$. For simplicity, let $\alpha_n = \mathbb{P}(Y_0 \leq u_n)$. Then, since $\mathbb{P}(X_0 > u_n) = 2(1 - \alpha_n) - (1 - \alpha_n)^2$ we must have $\alpha_n \rightarrow 1$ and $n(1 - \alpha_n) \rightarrow \tau/2 \geq 0$, as $n \rightarrow \infty$.

In order to exemplify the usefulness of the binary string notation introduced earlier to describe the sets $Q_q^{(\kappa)}$, let $\ell \in \mathbb{N}$, $\varsigma \in \{0, 1\}^\ell$ and define the cylinder

$$C(\varsigma) = \{\underline{\omega} \in \{0, 1\}^{\mathbb{N}} : \underline{\omega} = \varsigma \underline{\omega}^*, \text{ for some } \underline{\omega}^* \in \{0, 1\}^{\mathbb{N}}\},$$

corresponding to the binary words which start with the string ς . Then,

$$Q_q^{(0)}(u_n) = h^{-1}(C(100)) = \{X_0 > u_n, X_1 \leq u_n, X_2 \leq u_n\} = \{Y_{-2} > u_n, Y_{-1}, Y_0, Y_1, Y_2 \leq u_n\},$$

$$Q_q^{(1)}(u_n) = h^{-1}(C(1100) \cup C(10100)) =$$

$$= \{Y_0 > u_n, Y_{-1}, Y_1, Y_2, Y_3, Y_4 \leq u_n\} \cup \{Y_{-2}, Y_{-1} > u_n, Y_0, Y_1, Y_2, Y_3 \leq u_n\};$$

$$Q_q^{(2)}(u_n) = h^{-1}(C(11100) \cup C(110100) \cup C(101100) \cup C(1010100)) =$$

$$= \{Y_{-1}, Y_0 > u_n, Y_1, Y_2, Y_3, Y_4 \leq u_n\} \cup \{Y_{-2}, Y_{-1} > u_n, Y_0, Y_2, Y_3, Y_4, Y_5 \leq u_n\} \\ \cup \emptyset \cup (\{Y_2 > u_n, Y_{-1}, Y_1, Y_3, Y_4, Y_5, Y_6 \leq u_n\} \cap (\{Y_0 > u_n\} \cup \{Y_{-2} > u_n\}));$$

iid. Using that $Y_{-2}, Y_{-1}, Y_0, Y_1, \dots$ is an iid sequence, we have

$$\mathbb{P}(Q_q^{(0)}(u_n)) = (1 - \alpha_n) \alpha_n^4,$$

$$\mathbb{P}(Q_q^{(1)}(u_n)) = (1 - \alpha_n) \alpha_n^5 + (1 - \alpha_n)^2 \alpha_n^4;$$

$$\mathbb{P}(Q_q^{(2)}(u_n)) = (1 - \alpha_n)^2 \alpha_n^4 + (1 - \alpha_n)^2 \alpha_n^5 + (1 - \alpha_n)^2 \alpha_n^6 (1 + \alpha_n).$$

Using the formulae for the EI, (2.14), and the multiplicity distribution, (2.6), we obtain

$$\begin{aligned}\theta_n &= \frac{(1 - \alpha_n)\alpha_n^4}{2(1 - \alpha_n) - (1 - \alpha_n)^2} \xrightarrow{n \rightarrow \infty} \frac{1}{2} = \theta, \\ \pi_n(1) &= \frac{2(1 - \alpha_n)^2\alpha_n^4}{(1 - \alpha_n)\alpha_n^4} \xrightarrow{n \rightarrow \infty} 0 = \pi(1) \\ \pi_n(2) &= \frac{(1 - \alpha_n)\alpha_n^5 - (1 - \alpha_n)^2\alpha_n^5 - (1 - \alpha_n)^2\alpha_n^6(1 + \alpha_n)}{(1 - \alpha_n)\alpha_n^4} \xrightarrow{n \rightarrow \infty} 1 = \pi(2).\end{aligned}$$

Observing that for $\kappa > 2$, all the terms of $\mathbb{P}(Q_q^{(\kappa)}(u_n))$ include a factor $(1 - \alpha_n)^2$, one easily verifies that $\pi(\kappa) = 0$ for all such $\kappa > 2$.

Also note that condition $\mathbb{D}_q(u_n)^*$ follows trivially from the fact that the process is 2-dependent.

Regarding condition $\mathbb{D}'_q(u_n)$, observe that, for $j = 3, 4$, we have

$$Q_q^{(0)}(u_n) \cap \{X_j > u_n\} = \{Y_{-2}, Y_j > u_n, Y_{-1}, Y_0, Y_1, Y_2 \leq u_n\}$$

and for $j \geq 5$, we have

$$Q_q^{(0)}(u_n) \cap \{X_j > u_n\} = \{Y_{-2} > u_n, Y_{-1}, Y_0, Y_1, Y_2 \leq u_n\} \cap (\{X_{j-2} \geq u_n\} \cup \{X_j \geq u_n\}).$$

Then, clearly, $\mathbb{P}(Q_q^{(0)}(u_n) \cap \{X_j > u_n\}) \leq 2(1 - \alpha_n)^2\alpha_n^4$, for all $j \geq 3$. Hence,

$$\sum_{j=3}^{\lfloor n/k_n \rfloor} n\mathbb{P}(Q_q^{(0)}(u_n) \cap \{X_j > u_n\}) \leq \frac{n}{k_n} 2(1 - \alpha_n)^2\alpha_n^4 \xrightarrow{n \rightarrow \infty} 0,$$

by the properties of α_n .

Example 2.2 revisited

We take $u_n = -\log \tau + \log(2n)$, so that $\mu(X_0 > u_n) = \tau/n$. We observe that $U^{(0)}(u_n) = (1/3 - 1/2n, 1/3 + 1/2n)$ and for $\kappa \in \mathbb{N}_0$ we have

$$Q_q^{(\kappa)}(u_n) = \left(\frac{1}{3} - \frac{1}{4^\kappa} \frac{\tau}{2n}, \frac{1}{3} - \frac{1}{4^{\kappa+1}} \frac{\tau}{2n}\right) \cup \left(\frac{1}{3} + \frac{1}{4^{\kappa+1}} \frac{\tau}{2n}, \frac{1}{3} + \frac{1}{4^\kappa} \frac{\tau}{2n}\right) \Rightarrow \mu(Q_q^{(\kappa)}(u_n)) = \frac{1}{4^\kappa} \frac{3}{4} \frac{\tau}{n}.$$

It follows that

$$\theta_n = \frac{3\tau/4n}{\tau/n} \xrightarrow{n \rightarrow \infty} 3/4 = \theta \quad \text{and} \quad \pi_n(\kappa) = \frac{\frac{1}{4^{\kappa-1}} \frac{3}{4} \frac{\tau}{n} - \frac{1}{4^\kappa} \frac{3}{4} \frac{\tau}{n}}{\frac{3\tau}{4n}} \xrightarrow{n \rightarrow \infty} \frac{1}{4^{\kappa-1}} \frac{3}{4} = \pi(\kappa).$$

Note that $\sum_{\kappa \geq 1} \kappa \pi(\kappa) = 4/3 = \theta^{-1}$.

The convergence of the REPP to a compound Poisson process with a geometric multiplicity distribution is assured by the validity of conditions $\mathbb{D}_q(u_n)^*$ and $\mathbb{D}'_q(u_n)^*$, which follows from the fact that the system has decay of correlations against L^1 (see [25, Chapter 4] for further details).

Example 2.3 revisited

Adjusting the computations of the previous example, we take $u_n = -\log \tau + \log(4n)$, so that $\mu(X_0 > u_n) = \tau/n$. We observe that $U^{(0)}(u_n) = (1/3 - 1/4n, 1/3 + 1/4n) \cup (5/7 - 1/4n, 5/7 + 1/4n)$ and for $\kappa \in \mathbb{N}_0$ we have

$$Q_q^{(\kappa)}(u_n) = \left(\frac{1}{3} - \frac{1}{4^\kappa} \frac{\tau}{4n}, \frac{1}{3} - \frac{1}{4^{\kappa+1}} \frac{\tau}{4n} \right) \cup \left(\frac{1}{3} + \frac{1}{4^{\kappa+1}} \frac{\tau}{4n}, \frac{1}{3} + \frac{1}{4^\kappa} \frac{\tau}{4n} \right) \cup \\ \cup \left(\frac{5}{7} - \frac{1}{8^\kappa} \frac{\tau}{4n}, \frac{5}{7} - \frac{1}{8^{\kappa+1}} \frac{\tau}{4n} \right) \cup \left(\frac{5}{7} + \frac{1}{8^{\kappa+1}} \frac{\tau}{4n}, \frac{5}{7} + \frac{1}{8^\kappa} \frac{\tau}{4n} \right).$$

Observing that $\mu(Q_q^{(\kappa)}(u_n)) = \frac{1}{4^\kappa} \frac{3}{4} \frac{\tau}{2n} + \frac{1}{8^\kappa} \frac{7}{8} \frac{\tau}{2n}$, we have

$$\theta_n = \frac{\frac{3}{4} \frac{\tau}{2n} + \frac{7}{8} \frac{\tau}{2n}}{\tau/n} = \frac{13}{16} = \theta \\ \pi_n(\kappa) = \frac{\frac{1}{4^{\kappa-1}} \left(\frac{3}{4}\right)^2 \frac{\tau}{2n} - \frac{1}{8^{\kappa-1}} \left(\frac{7}{8}\right)^2 \frac{\tau}{2n}}{\frac{3}{4} \frac{\tau}{2n} - \frac{7}{8} \frac{\tau}{2n}} = \frac{\frac{1}{4^{\kappa-1}} \left(\frac{3}{4}\right)^2 - \frac{1}{8^{\kappa-1}} \left(\frac{7}{8}\right)^2}{\frac{13}{8}} = \pi(\kappa).$$

Note that $\sum_{\kappa \geq 1} \kappa \pi(\kappa) = 4/3 = \theta^{-1}$.

Again, the convergence of the REPP to a compound Poisson process is assured by the validity of conditions $\mathcal{D}_q(u_n)^*$ and $\mathcal{D}'_q(u_n)^*$, which follows from the fact that the system has decay of correlations against L^1 (see [25, Chapter 4] for further details).

3. Dynamical counterexamples

Let us consider a 1-dimensional family of maps with an indifferent fixed point of the MP type. We will be using the particular form given in [24]. Namely, for $\alpha > 0$,

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

If $\alpha \in (0, 1)$, then there is an absolutely continuous (with respect to Lebesgue) invariant probability μ_α , which is the case we will restrict to. These maps have been studied in [20, 24–32] 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}_β 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 $t \in \mathbb{N}$,

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

Let $h_\alpha = \frac{d\mu_\alpha}{dx}$. In [20], 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 $\lim_{x \rightarrow 0} \frac{h(x)}{x^{-\alpha}} = C_0 > 0$. Hence, for small $s > 0$, we have

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

for some $C_1 > 0$, where the notation $A(s) \sim B(s)$ is used in the sense that $\lim_{s \rightarrow 0} \frac{A(s)}{B(s)} = 1$. 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 $\lim_{s \rightarrow 0} \frac{A(s)}{B(s)} = c$.

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

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

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

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

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 $\varphi : [0, 1] \rightarrow \mathbb{R} \cup \{+\infty\}$ and defining the process X_0, X_1, \dots by

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

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 $\mathcal{C} = \{\zeta_1, \zeta_2\}$, that is, $f_\alpha^j(\zeta_2) \notin \mathcal{C}$, $\forall j \in \mathbb{N}$. 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 [18, Theorem 2], we would have an EI equal to 0, while, by [18, 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:

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

for some $\delta > 0$, where dist denotes any given metric on $[0, 1]$ and the function $g : [0, +\infty) \rightarrow \mathbb{R} \cup \{+\infty\}$ is such that 0 is a global maximum ($g(0)$ may be $+\infty$); g is a strictly decreasing bijection $g : V \rightarrow W$ in a neighbourhood V of 0; and has one of the three types of behaviour described, for example, in [25, 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 (2.10).

THEOREM 3.1. *Consider 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 (2.10) for such stochastic process and for a sequence of levels $(u_n)_{n \in \mathbb{N}}$ satisfying (2.9) 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 2.7 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 , that is,

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

In order to prove this theorem, we apply Theorem 2.12. To that end we need to check conditions $\mathbb{D}_q(u_n)^*$ and $\mathbb{D}'_q(u_n)^*$ which we leave for Subsections 3.3 and 3.4, respectively. We are left to prove that θ_n given in (2.14) 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

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

Defining now $y_n = (\frac{1}{C_1} g^{-1}(u_n))^{1/(1-\alpha)}$ 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

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

In this case,

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

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

$$\mu_\alpha(Q_p^{(0)}(u_n)) \sim c(1-\alpha)2^\alpha x_n + o(x_n) + \frac{\tau}{2n}. \quad (3.10)$$

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

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

which implies

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

Then, by (3.10),

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

In this way, we easily obtain

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

Recall that

$$\pi_n(k) = \frac{\mu_\alpha(Q_p^{(k-1)}(u_n)) - \mu_\alpha(Q_p^{(k)}(u_n))}{\mu_\alpha(Q_p^{(0)}(u_n))}.$$

In this case, $Q_p^{(1)}(u_n) = [x_n^{(1)}, x_n]$, where $T_\alpha(x_n^{(1)}) = x_n$, that is, $x_n^{(1)} + 2^\alpha(x_n^{(1)})^{1+\alpha} = x_n$, which implies $x_n^{(1)} = O(x_n)$.

Consequently, by (3.11)

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

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

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

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 $T_\alpha^j(\zeta_2) \notin \{\zeta_1, \zeta_2\}$, $\forall j \in \{1, \dots, p-1\}$. As proved in [15], 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, that is, $\theta = \frac{1}{2}(1 - \gamma^{-1})$, 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:

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

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 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, \dots 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 (2.10) for such stochastic process and for a sequence of levels $(u_n)_{n \in \mathbb{N}}$ satisfying (2.9) converges in distribution to a compound Poisson process N defined on the positive real line with intensity $\theta\tau$ and multiplicity distribution given by*

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

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

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

REMARK 3.6. As in the previous example, recall that by Theorem 2.7 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 , that is,

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

Again, in order to prove this theorem, we apply Theorem 2.12. To that end, we need to check conditions $\mathbb{D}_q(u_n)^*$ and $\mathbb{D}'_q(u_n)^*$ which we leave for Subsections 3.3 and 3.4, respectively. We are left to prove that θ_n given in (2.14) 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 $y_n = (\frac{1}{C_1} g^{-1}(u_n))^{1/(1-\alpha)}$ 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,

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

whereas before $B_{\delta_n}(\zeta_2) = (\zeta_2 - \delta_n, \zeta_2 + \delta_n)$.

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

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

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

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

Hence,

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

Then, by (3.14),

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

Gathering this information, we obtain

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

We compute now the multiplicity distribution. Observe that

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

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

$$\mu_\alpha(Q_p^{(1)}(u_n)) = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}\gamma^{-1}(1 - \gamma^{-1}). \quad (3.17)$$

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

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

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

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

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

$$\mu_\alpha(Q_p^{(2)}(u_n)) = O\left(\left(\frac{\tau}{2n}\right)^{1/(1-\alpha)}\right) + \frac{\tau}{2n}\gamma^{-2}(1 - \gamma^{-1}). \quad (3.18)$$

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

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

A simple inductive argument then leads to

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

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

3.3. The condition $\mathbb{D}_q(U_n)^*$

Condition $\mathbb{D}_q(U_n)^*$ has been designed to be easily verified for systems with sufficiently fast decay of correlations. The argument used in [18, Section 4.2.2] allows to show that $\mathbb{D}_q(U_n)^*$ follows from the decay of correlations stated in (3.2), as long as $\alpha < \sqrt{5} - 2$.

3.4. The condition $\mathbb{D}'_q(U_n)^*$

This subsection is dedicated to the verification of condition $\mathbb{D}'_q(U_n)^*$. 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 $F_\alpha : [1/2, 1] \rightarrow [1/2, 1]$ given by $F_\alpha(x) = T_\alpha^{r_B(x)}(x)$, where $B = [1/2, 1]$ and $r_B : B \rightarrow \mathbb{N}$ is the first return time to B , that is, $r_B(x) = \inf\{j \in \mathbb{N} : f_\alpha^j(x) \in B\}$. 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

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

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 $i \leq j$ 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 $i \leq j-1$ such that $x \in B_n \cap F_\alpha^{-i}\tilde{E}_n$. Therefore,

$$n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \bar{\mu}_\alpha(B_n \cap T_\alpha^{-j}(U^{(0)}(u_n))) \leq n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \bar{\mu}_\alpha(B_n \cap F_\alpha^{-j}(D_n \cup \tilde{E}_n)).$$

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 Subsections 3.1 and 3.2, we have $R_n \rightarrow \infty$ as $n \rightarrow \infty$. 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 $\psi = \mathbf{1}_{D_n \cup \tilde{E}_n}$, the facts that $\bar{\mu}_\alpha(D_n \cup \tilde{E}_n) = O(n^{-\frac{1}{1-\alpha}}) + O(n^{-1}) = O(n^{-1})$ and $\|\phi\|_{BV} \leq 6$, we have

$$\begin{aligned} n \sum_{j=q+1}^{\lfloor n/k_n \rfloor - 1} \bar{\mu}_\alpha(B_n \cap F_\alpha^{-j}(D_n \cup \tilde{E}_n)) &= n \sum_{j=R_n}^{\lfloor n/k_n \rfloor - 1} \bar{\mu}_\alpha(B_n \cap F_\alpha^{-j}(D_n \cup \tilde{E}_n)) \\ &\leq n \sum_{j=R_n}^{\lfloor n/k_n \rfloor - 1} \bar{\mu}_\alpha(B_n) \bar{\mu}_\alpha(D_n \cup \tilde{E}_n) + 6Cn \bar{\mu}_\alpha(D_n \cup \tilde{E}_n) \sum_{j=R_n}^{\lfloor n/k_n \rfloor - 1} e^{-aj} \\ &\leq O\left(\frac{1}{k_n}\right) + O\left(\sum_{j=R_n}^{\infty} e^{-aj}\right) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

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

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

where $R_n = \inf\{j \in \mathbb{N} : A_n \cap T_\alpha^{-j}(D_n \cup \tilde{E}_n) \neq \emptyset\}$. Let $P_\alpha : L^1(\text{Leb}) \rightarrow L^1(\text{Leb})$ 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

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

Following now the same argument used in [18, 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 [18, 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 $\mu_\alpha(D_n \cup \tilde{E}_n) = O(n^{-1})$, we have

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

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 (2.9), while $\mathbb{E}(N([0, 1])) = \frac{1}{2}\tau$, in the case considered in Subsection 3.1, and $\mathbb{E}(N([0, 1])) = \frac{1}{2}(1 - \gamma^{-1})\tau$, in the case considered in Subsection 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 2-dimensional point processes as studied in [11], namely,

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

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 thought of as the asymptotic frequency associated to a given threshold on the range of the random variable (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, that is, 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, for example, [27]). In particular, note that if we define $H_\tau : E^2 \rightarrow E$, 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 $\varphi : [0, 1] \rightarrow \mathbb{R} \cap \{+\infty\}$ is maximised at a single periodic point $\zeta \in [1/2, 1]$, that is,

$$\varphi(x) = g(|x - \zeta|), \quad (4.2)$$

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, \dots defined as in 3.7 for such φ and given a sequence $(u_n(\tau))_{n \in \mathbb{N}}$ as in (2.9), by [16], we have that there exists an EI $\theta = (1 - \gamma^{-1})$ and N_n given in (2.10) converges to a compound Poisson process N of intensity $\theta\tau$ with a geometric cluster size distribution, that is, $\pi(\kappa) = \mathbb{P}(D_i = \kappa) = \theta(1 - \theta)^{\kappa-1}$. Moreover, by [11], we have that $N_n^{(2)}$ converges weakly to

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

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 $\text{Exp}(\theta)$ 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. $U_{i,j} \stackrel{D}{\sim} \mathcal{U}_{(i-1, i]}$, that is, $U_{i,j}$ 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 $\theta \cdot \text{Leb}$ as its intensity measure, where $\theta \cdot \text{Leb}([a, b] \times [c, d]) = \theta(b-a)(d-c)$, 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 $(\frac{j+2}{n}, u_n^{-1}(X_{j+2}))$, $(\frac{j+4}{n}, u_n^{-1}(X_{j+4}))$, \dots will also be charged by $N_n^{(2)}$ and since by the previous computations and the form of φ we have $u_n^{-1}(X_{j+2}) \sim \gamma$, $u_n^{-1}(X_{j+4}) \sim \gamma^2, \dots$, 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

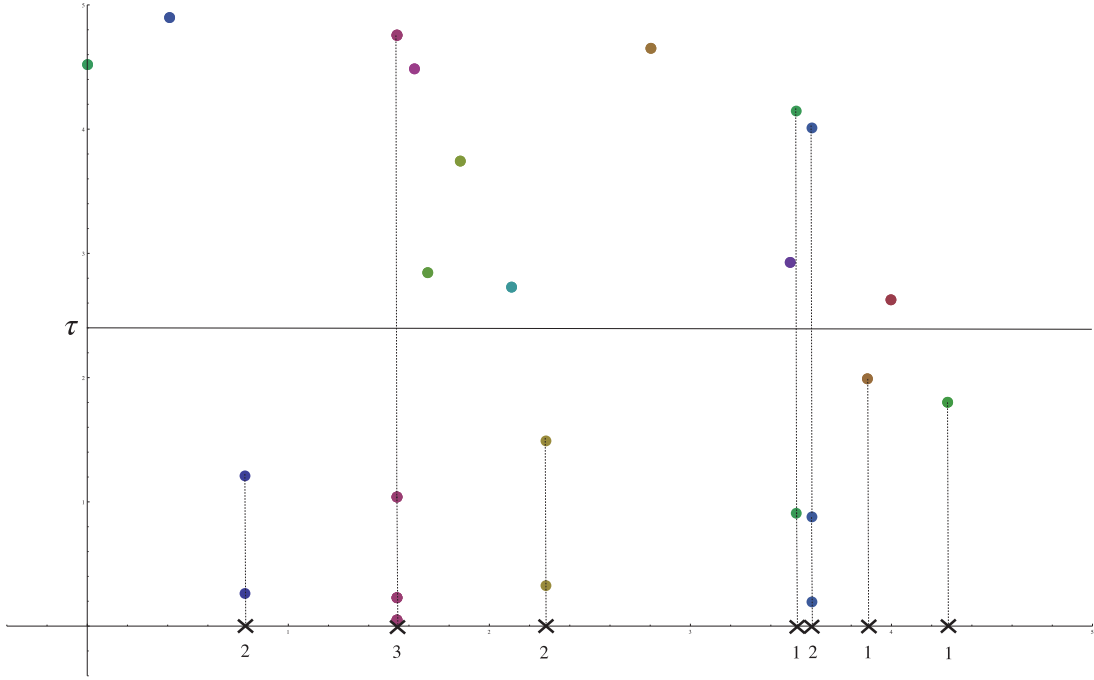


FIGURE 1 (colour online). Simulation of the bi-dimensional process $N^{(2)}$ in the case of the periodic point where the observable φ is given by 4.2. The picture also shows how the projection H_τ works to obtain $N = H_\tau(N^{(2)})$, which is a compound Poisson process on the line, where the crosses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

of points in each vertical pile still below the threshold τ that project on the same time event (see Figure 1).

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 $\frac{1}{2} \cdot \text{Leb}$, 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 $u_n(\tau) = g(\tau/(2n))$, which implies $u_n^{-1}(z) = 2ng^{-1}(z)$. For definiteness, let us assume $u_n^{-1}(X_j) = 1$, 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 $(\frac{j+1}{n}, u_n^{-1}(X_{j+1}))$, $(\frac{j+2}{n}, u_n^{-1}(X_{j+2}))$, \dots , 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 $T_\alpha^j(x) \sim (\frac{1}{2nC_1})^{\frac{1}{1-\alpha}}$. 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 $u_n^{-1}(X_{j+2}) \sim 1$ and so on. Therefore, do not only the points of the same cluster get vertically aligned but they also get horizontally aligned, that is, 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 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 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 $\text{Exp}(\theta)$ 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 (2006) 593–609.
2. M. ABADI, L. CARDEÑO and S. GALLO, ‘Potential well spectrum and hitting time in renewal processes’, *J. Stat. Phys.* 159 (2015) 1087–1106.
3. M. ABADI, A. C. M. FREITAS and J. M. FREITAS, ‘Clustering indices and decay of correlations in non-Markovian models’, *Nonlinearity* 32 (2019) 4853–4870.
4. H. AYTAÇ, J. M. FREITAS and S. VAIENTI, ‘Laws of rare events for deterministic and random dynamical systems’, *Trans. Amer. Math. Soc.* 367 (2015) 8229–8278.
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 (2016) 33–48.
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 (2017) 1244–1261.
7. M. R. CHERNICK, T. HSING and W. P. MCCORMICK, ‘Calculating the extremal index for a class of stationary sequences’, *Adv. Appl. Probab.* 23 (1991) 835–850.

8. T. DOWNAROWICZ and Y. LACROIX, 'The law of series', *Ergodic Theory Dynam. Systems* 31 (2011) 351–367.
9. C. A. T. FERRO, 'Statistical methods for clusters of extreme values', PhD Thesis, Lancaster University, Lancaster, Lancashire, 2003.
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 (2018) 2131–2179.
11. A. C. M. FREITAS, J. M. FREITAS, M. MAGALHÃES, 'Complete convergence and records for dynamically generated stochastic processes', *Trans. Amer. Math. Soc.* 373 (2020) 435–478.
12. A. C. M. FREITAS, J. M. FREITAS, F. B. RODRIGUES and J. V. SOARES, 'Rare events for cantor target sets', *Comm. Math. Phys.*, to appear, arXiv:1903.07200.
13. A. C. M. FREITAS, J. M. FREITAS and M. TODD, 'Hitting time statistics and extreme value theory', *Probab. Theory Related Fields* 147 (2010) 675–710.
14. A. C. M. FREITAS, J. M. FREITAS and M. TODD, 'Extreme value laws in dynamical systems for non-smooth observations', *J. Stat. Phys.* 142 (2011) 108–126.
15. A. C. M. FREITAS, J. M. FREITAS and M. TODD, 'The extremal index, hitting time statistics and periodicity', *Adv. Math.* 231 (2012) 2626–2665.
16. 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 (2013) 483–527.
17. 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 (2015) 1653–1687.
18. A. C. M. FREITAS, J. M. FREITAS, M. TODD and S. VAIENTI, 'Rare events for the Manneville–Pomeau map', *Stochastic Process. Appl.* 126 (2016) 3463–3479.
19. T. HSING, J. HÜSLER and M. R. LEADBETTER, 'On the exceedance point process for a stationary sequence', *Probab. Theory Related Fields* 78 (1988) 97–112.
20. H. HU, 'Decay of correlations for piecewise smooth maps with indifferent fixed points', *Ergodic Theory Dynam. Systems* 24 (2004) 495–524.
21. O. KALLENBERG, *Random measures*, 4th edn (Akademie, Berlin, 1986).
22. G. KELLER, 'Rare events, exponential hitting times and extremal indices via spectral perturbation', *Dyn. Syst.* 27 (2012) 11–27.
23. M. R. LEADBETTER, 'Extremes and local dependence in stationary sequences', *Z. Wahrsch. Verw. Gebiete* 65 (1983) 291–306.
24. C. LIVERANI, B. SAUSSOL and S. VAIENTI, 'A probabilistic approach to intermittency', *Ergodic Theory Dynam. Systems* 19 (1999) 671–685.
25. 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).
26. G. L. O'BRIEN, 'Extreme values for stationary and Markov sequences', *Ann. Probab.* 15 (1987) 281–291.
27. S. I. RESNICK, *Extreme values, regular variation, and point processes*, Applied Probability. A Series of the Applied Probability Trust 4 (Springer, New York, 1987).
28. C. Y. ROBERT, 'Automatic declustering of rare events', *Biometrika* 100 (2013) 587–606.
29. M. R. RYCHLIK, 'Another proof of Jakobson's theorem and related results', *Ergodic Theory Dynam. Systems* 8 (1988) 93–109.
30. R. L. SMITH, 'A counterexample concerning the extremal index', *Adv. in Appl. Probab.* 20 (1988) 681–683.
31. R. L. SMITH and I. WEISSMAN, 'Estimating the extremal index', *J. Roy. Statist. Soc. Ser. B* 56 (1994) 515–528.
32. L.-S. YOUNG, 'Recurrence times and rates of mixing', *Israel J. Math.* 110 (1999) 153–188.

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

leugim@ime.usp.br
[http://miguelabadi.wixsite.com/
miguel-abadi](http://miguelabadi.wixsite.com/miguel-abadi)

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

jmfreira@fc.up.pt
<http://www.fc.up.pt/pessoas/jmfreira/>

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

amoreira@fep.up.pt
<http://www.fep.up.pt/docentes/amoreira/>