HAL
open science

On the Computation of the Extremal Index for Time Series

Th Caby, D. Faranda, Sandro Vaienti, Pascal Yiou

To cite this version:

Th Caby, D. Faranda, Sandro Vaienti, Pascal Yiou. On the Computation of the Extremal Index for Time Series. Journal of Statistical Physics, 2020, 179 (5-6), pp.1666-1697. 10.1007/s10955-019-02423-z . hal-02450472

HAL Id: hal-02450472

https://hal.science/hal-02450472v1

Submitted on 9 Feb 2020

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

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

ON THE COMPUTATION OF THE EXTREMAL INDEX FOR TIME SERIES

Abstract. The extremal index is a quantity introduced in extreme value theory to measure the presence of clusters of exceedances. In the dynamical systems framework, it provides important information about the dynamics of the underlying systems. In this paper we provide a review of the meaning of the extremal index in dynamical systems. Depending on the observables used, this quantity can inform on local properties of attractors such as periodicity, stability and persistence in phase space, or on global properties such as the Lyapunov exponents. We also introduce a new estimator of the extremal index and shows its relation with those previously introduced in the statistical literature. We reserve a particular focus to the systems perturbed with noise as they are a good paradigm of many natural phenomena. Different kind of noises are investigated in the annealed and quenched situations. Applications to climate data are also presented.

Th. Caby [1], D. Faranda[2], S. Vaienti[3], P. Yiou[4]

arXiv:1904.04936v1 [math.DS] 9 Apr 2019

Contents

1. Introduction 2
2. The deterministic case 3
2.1. Statistical estimators of the extremal index 6
2.2. Numerical estimates of the extremal index in deterministic systems 7
3. The random case: non-stationary situations (quenched noise) 8
3.1. Sequential systems 8
3.2. Random fibred systems 10
4. The random case: stationary situations (annealed noise) 11
4.1. I.i.d. random transformations 11
4.2. Moving target 14
5. The dynamical extremal index 18
5.1. Additive noise 19
5.2. Discrete noise 19
6. Point processes: statistics of persistence 20
6.1. The deterministic case 20
6.2. The random case 22
6.3. The rotational case 23
6.4. Climate data 23
7. Conclusions 24
8. appendix 25
9. Acknowledgement 26
References 26

Date: April 11, 2019.

[1]Aix Marseille Université, Université de Toulon, CNRS, CPT, 13009 Marseille, France and Center for Nonlinear and Complex Systems, Dipartimento di Scienza ed Alta Tecnologia, Università degli Studi dell'Insubria, Como, Italy. E-mail: `caby.theo@gmail.com` ;

[2]Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL and Université Paris-Saclay, 91191 Gif-sur-Yvette, France and London Mathematical Laboratory, 8 Margravine Gardens, London, W6 8RH, UK. Email: `davide.faranda@lsce.ipsl.fr`.

[3]Aix Marseille Université, Université de Toulon, CNRS, CPT, 13009 Marseille, France. E-mail: `vaienti@cpt.univ-mrs.fr`.

[4]Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL and Université Paris-Saclay, 91191 Gif-sur-Yvette, France. E-mail: `pascal.yiou@lsce.ipsl.fr`.

1. Introduction

In the last ten years extreme value theory (EVT) has been successfully applied to the study of time series generated by dynamical systems. Illustrations can be found in the book [13] and the articles [16, 36, 35] for an exhaustive account on the formalism, the methodologies and several applications. In particular the extremal index (EI) — a quantity defined in the unit interval — has been used as a powerful statistical indicator to discriminate among different qualitative types of dynamical behaviors in the phase space of a few climate models and in real situations (an extensive overview can be found in [37]). In a series of papers [17, 7, 19, 37, 18], the EI was renamed as *local persistence* indicator, suitable to estimate the average cluster size of the trajectories within the neighborhood of a given point. The aim of this note is to give an overview on a few recent rigorous mathematical results which show that the EI is very sensitive to stochastic perturbations affecting the deterministic evolution of a system states. We will also discuss the influence of noise on the *dynamical extremal index*, which was recently introduced to characterize the local divergence of chaotic orbits[20].

The first rigorous computations of the EI for expanding and uniformly hyperbolic systems (from now on named *chaotic systems*), showed a dichotomy for the value of the EI, which is equal to 1 everywhere, except on periodic points. [24] give a detailed account of this matter. On the other hand, the EI exhibits local variations which have been also connected to the local fractal dimensions. By assuming that the data we are interested in are chaotic, one could wonder about the wider variation of the extremal index, offering a spectrum of values beyond the previous dichotomy. In a recent paper [8] we commented about a similar effect concerning the computation of the local dimensions. Indeed, while it is well known for a large class of systems that the local dimensions are all equal to the so-called information dimension D_1 with probability one, large deviation theory estimates the likelihood of deviations from this value at finite resolution. In this perspective, the spread of the experimentally observed values of the local dimensions can be thought of as originating from the multifractal structure of the invariant measure, which in turn is revealed by the non-constant value of generalized dimensions.

Although finite resolution affects the estimation of the extremal index, it is unlikely that it enjoys some large deviation property. This is due to the previous dichotomy, which prevents the existence of a smooth spectrum of values for the EI. Yet, the computation of the EI is very sensitive to randomness and disturbances in the measuring process. The main object of our note is to show how the EI is affected by the presence of noise. Our starting point is to assume for the EI a general formula obtained by Keller and Liverani [33] in the context of stationary dynamical systems enjoying the so-called *spectral gap property*. For this reason we call it the *spectral formula*. It encompasses all the other rigorous formulae previously obtained for the computation of the extremal index. We will show that when the spectral formula holds, one can obtain a generalized version of the O'Brien formula, which inspired several numerical algorithms and statistical estimators (see Eq. (3.2.4) in [13]). This is discussed in Proposition 2 and Remark 2, which constitute two main contributions of this paper.

The first sections are devoted to the computation of the EI for the standard observable given by the distribution of the first visit of the trajectory in a decreasing net of balls around a given point, which we qualify as *the target set*. The system will be perturbed according to three different classes of noises: sequential, quenched and annealed. In particular, we will focus on the annealed class since it gives stationary random processes particularly suitable for the application of the spectral formula. We will show on simple but non trivial examples that noises with discrete distributions tend to make the EI less than one, while noises with continuous distributions usually make the extremal index equal to one. The latter case applies also to deterministic dynamics with the target set affected by disturbances with smooth densities, which corresponds to common physical situations.

The final section deals with a different extremal index we introduced recently with the name of Dynamical Extremal Index and that explores the mutual distances between the coordinates of a point moving in a suitable higher dimensional direct product space. This new index captures

the rate of phase space contraction and is naturally related to positive Lyapunov exponent(s). We will show again how this index is sensitive to the different kinds of perturbations described above.

We finally analyze how the random perturbations affect the statistics of the number of visits in the target regions. The expectation of the limiting law is the reciprocal of the extremal index and the differences in the value of the EI reflect in different types of Poisson compound distributions.

2. The deterministic case

Our intent is to provide a critical discussion of the application of the EI to time series. Therefore, we start defining it for particular random processes. Let us therefore consider a discrete dynamical systems defined by a map T acting on a smooth manifold X and preserving a Borel measure μ. Suppose z is a point of X and $B(z,r)$ an open ball around z and of radius r. Given any other point $x \in X$ let us consider the random variables $H_n(x)$ given by

$$H_n(x) := \{\text{first time the iterate } T^n(x) \text{ enters the open ball } B(z, e^{-u_n})\}, \tag{1}$$

where the *boundary level* u_n is defined by asking that

$$n\ \mu(B(z, e^{-u_n})) \to \tau, \tag{2}$$

for some positive number τ. Notice that by the stationarity (or invariance) of the measure μ, the quantity $\mu(B(z, e^{-u_n}))$ gives the probability that any iterate of the map T be in the open balls $B(z, e^{-u_n})$. Moreover as soon as the measure is not atomic, the measure of a ball varies continuously with the radius and this allows us to explicit u_n as a function of τ and n.

We say that we have an extreme value law for the process H_n if

$$\mu(x; H_n(x) > u_n) \to e^{-\theta\tau}, \tag{3}$$

where $0 \le \theta \le 1$ is called the *extremal index*.

We remind that the events $H_n > u_n$ are equivalent to the following ones:

$$\{M_n(x) \le u_n\}, \text{ where } M_n = \max\{\phi(x), \phi(Tx), \dots, \phi(T^{n-1}x\}, \tag{4}$$

and the observable ϕ is defined as

$$\phi(x) = -\log d(x,z). \tag{5}$$

It is remarkable that for a large class of systems whose transfer operator admits a spectral gap, the extremal index can be explicitly computed. This is done by G. Keller [32], who applied the perturbative theory developed in [33]. This spectral approach has the advantage to give a formula for the EI which holds for general target sets when their measure goes to zero. We will use it in the last section of this note when balls are replaced by tubular neighborhoods.

Let us first define the event

$$\Omega_n^{(k)}(z) := \{\phi(x) > u_n, \max_{i=1,\dots,k} \phi(T^i x) \le u_n,\ \phi(T^{k+1}x) > u_n\}, \tag{6}$$

namely the set of points in $B(z, e^{-u_n})$ whose first k iterates are outside $B(z, e^{-u_n})$ and whose $(k+1)$-th iterate falls again in $B(z, e^{-u_n})$.

Suppose that the following limit exists:

$$q_k := \lim_{n\to\infty} \frac{\mu(\Omega_n^{(k)}(z))}{\mu(B(z, e^{-u_n}))}. \tag{7}$$

Then the perturbative theory gives an estimate of the EI:

$$\theta = 1 - \sum_{k=0}^{\infty} q_k. \tag{8}$$

This formula allows us to reproduce a few existing results whenever the target point z is of a particular type. For instance, for one-dimensional expanding systems with strong mixing properties and preserving a measure absolutely continuous with respect to Lebesgue, the extremal index is 1 everywhere but in periodic points. When we have a periodic point z of (minimal) period p, only the q_{p-1} term is nonzero and it is of the form [32, 24]

$$q_{p-1} = \frac{1}{|DT^p(z)|}. \tag{9}$$

The dependence of the EI on the periodicity of the target point z means that the sojourns and returns of the point into the ball $B(z, e^{-u_n})$, which constitute the *clusters of exceedances*, keep memory either of the past orbit and of the topological structure of the target point.

This can be made more precise by defining a cluster size distribution $\pi_n(j)$ as the probability of having j returns into the ball up to a rescaled time n/k_n, for a suitable sequence $k_n = o(n)$. Section 3.2.1 in [13] provides more details on this derivation; a more formal definition is also given in section 7. It can be therefore proved that (see Sec. 3.3.3 in [13]):

$$\theta^{-1} = \lim_{n\to\infty} \sum_{j=1}^{\infty} j\pi_n(j), \tag{10}$$

which is interpreted by saying that the *EI is equal to the inverse of the average cluster size*.[5] In the applications to dynamical systems, the periodicity translates into the fact that the clustering $\pi = \lim_{n\to\infty} \pi_n$ is actually a geometric distribution of parameter $\theta \in (0, 1]$, i.e., $\pi_k = \theta(1-\theta)^{k-1}$, for every $k \in \mathbb{N}_0$ [30]. We will return to this matter in section 7.

The interpretation of the EI given in Eq. (8) has been used in applications to time series climate data and it was also emphasized the *local* character of such an indicator and its strong correlation with the fractal local dimension of the invariant measure [17, 7, 19].

Let us now consider for $k \in \mathbb{N}_0$ the event (compare with Eq. (6)):

$$A_n^{(k)} := \{\phi(x) > u_n \cap \max_{i=1,\dots,k} \phi(T^i x) \leq u_n\} \tag{11}$$

and the quantity:

$$\theta_k^{(n)} := \frac{\mu(A_n^{(k)})}{\mu(\phi(x) > u_n)} \tag{12}$$

where we set $\theta_0^{(n)} = 1, \forall n$. By introducing the event $\phi(T^{k+1} \leq u_n)$ and its complement, and passing to the limit for $n \to \infty$, we get

$$q_k = \theta_k - \theta_{k+1}, \tag{13}$$

where

$$\theta_k = \lim_{n\to\infty} \theta_k^{(n)}, \tag{14}$$

when the limit exists.

Notice that as soon as one of the two sequences $\{q_k\}_{k\in\mathbb{N}}, \{\theta_k\}_{k\in\mathbb{N}}$ is known, the other one is determined as well.

In some circumstances, the limit in Eq. (14) gives exactly the extremal index. To explain this point, we need to be more precise about the assumptions we make. The spectral approach briefly sketched above applies to a large class of dynamical systems admitting a spectral gap for the transfer (Perron-Fröbenius) operator. This means that the systems have exponential decay of correlations for smooth enough observables, usually of bounded variation type. This is not sufficient to establish the existence of the limit in Eq. (7). Computations of that limit under various circumstances are given in [32, 33, 4, 8]. The more standard approach to

[5] Abadi et al. [1] built a dynamically generated stochastic processes with an extremal index for which that equality does not hold. They considered 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. We will not consider these kind of observables in this paper.

EVT, which consists in adapting the classical Leadbetter theory for i.i.d. random variables to stationary dependent processes [34], requires milder mixing conditions that allow to get asymptotic independence for the process. The theory of O'Brien [38] was developed for strictly stationary processes verifying asymptotic independence, which is a strong probabilistic mixing condition like the ϕ-mixing (see [6] for the definition). We stress that asymptotic independence is very difficult to check in practice. Recently it has been shown that weaker conditions are enough. For our purposes, it is sufficient to remind condition $D(u_n)$ in section 2.3 of [13] successively improved in [24][6] Asymptotic independence is not enough to get the convergence of the limit distribution of the maxima. One needs to control short returns, which could be in particular affected by clustering. A powerful condition which takes care of that is:

$$D_{u_n}^{(k)} = \lim_{n\to\infty} n \sum_{j=k+1}^{[n/l_n]-1} \mu(A_n^{(k)} \cap T^{-j}(A_n^{(k)})) = 0, \tag{15}$$

where $l_n = o(n)$ is a sequence slowly diverging to infinity and verifying the $D(u_n)$ conditions (see Remark 4.1.2 in [13] for the details). We call the integer k the *clustering order*. We now summarize in the following proposition a few facts that are useful for our paper and which refer to the stationary process $\{\phi \circ T^k\}_{k\in\mathbb{N}}$, on the probability space (X, μ):

Proposition 1.
- *Suppose the sequence u_n verifies Eq. (2) and condition $D(u_n)$ holds. Moreover suppose that* $\liminf_{n\to\infty} \mu(M_n \le u_n) > 0$ *and condition in Eq. (15) holds with clustering order k. Then*

$$\mu(M_n \le u_n) - e^{-\tau\theta_k^{(n)}} \to 0,$$

  as shown by [10], *Proposition 1.1.*
- *If the limit in Eq. (14) exists, then* the extremal index is given by θ_k, *as shown by Corollary 1.3 in* [10] *or Corollary 4.1.7 in* [13]. *This is sometimes called the O'Brien formula.*
- *If condition Eq. (15) holds for some particular $k' \in \mathbb{N}$, it also holds for all $k \ge k'$ and therefore the limit in Eq. (14) for all $k \ge k'$ gives again the extremal index θ, as shown by* [24] *and* [13], *Remark 4.1.10.*

When the second item in the Proposition holds for a process with clustering order k we have

$$q_j = \theta_j - \theta_{j+1} = 0, \ \forall j \ge k. \tag{16}$$

The natural question is therefore to ask when condition (15) is verified. For a large class of dynamical systems, in particular verifying the assumptions of theorem 4.2.7 in [13], it can be shown that condition (15) holds with $k = 0$ if the target point z is not periodic, and with $k = p$ if z is periodic of prime period p (see Proposition 4.2.13 in [13]). In particular, for one-dimensional expanding systems with strongly mixing properties and preserving a measure absolutely continuous with respect to Lebesgue, the extremal index is $1 - q_{p-1} = 1 - \frac{1}{|DT^p(z)|}$ in periodic points z of (minimal) period p. For higher dimensional systems, condition (15) gives a precise value for θ (see for instance [9]).

Remark 1. *For the preceding example around periodic points, the spectral and the standard approaches both give q_{p-1} and θ_p different from zero. Yet, the spectral approach shows that all q_j are zero for $j \ne p-1$. This is therefore coherent with Eq. (16) for which $\theta_j = \theta_{j+1}$ for $j \ge p$. Let us now consider $q_{p-1} = \theta_{p-1} - \theta_p$. The quantity θ_{p-1} is well defined as the limit of $\theta_{p-1}^{(n)}$, when n goes to infinity and it is equal to 1. This implies that the condition (15) is violated, because otherwise θ_p should be equal to 1 as well. The fact that all the θ_j, for $j \le p-1$ are equal to 1 is not surprising: it means that the full conditional measure of points in the ball $B(z, e^{-u_n})$ are outside it when iterated up to $p-1$ times.*

[6]This condition holds for instance when the invariant measure μ is mixing with decay of correlation fast enough; sometimes a rate of decay as n^{-2} is sufficient.

In the assumptions of Proposition 1, the EI is equal to θ_k when the limit in Eq. (14) defining this quantity exists and k being the clustering order. We now show that this is a particular case of a more general result.

Proposition 2. *Let us suppose that the limits in Eq. (7) defining the quantities q_k exist for any $k \geq 0$. Then the extremal index θ is given by*

$$\theta = \lim_{k\to\infty} \theta_k. \tag{17}$$

Proof. We noticed that whenever the sequence q_k exists, the same happens for the sequence θ_k. Then by a simple telescopic trick we get

$$\theta = \lim_{k\to\infty} \left(1 - \sum_{j=0}^{k} q_j\right) = \lim_{k\to\infty} \theta_{k+1}.$$

□

Remark 2. *Under the assumptions of the preceding proposition, we got that*

$$\theta = \lim_{k\to\infty} \lim_{n\to\infty} \theta_k^{(n)}. \tag{18}$$

We stress that the limits in Eq. (14) or Eq. (7) have been computed up to now and in the framework of dynamical systems only around periodic points, or around the diagonal in coupled systems (see [15]*) where at most only one of them is different from zero. We will present later on further examples of explicit computations of the q_k. In particular section 4.1.2 will give an example of a stationary random dynamical system for which all the q_k are different from zero and from each other, which implies that condition (15) is violated at all orders. In this respect our Eq. (18) is a generalization of the O'Brien formula.*

2.1. **Statistical estimators of the extremal index.** We saw in the preceding section that the value θ_p provides the extremal index if the target point z is of (minimal) period p. It is therefore not surprising that estimators based on the O'Brien formula have been proposed for the computation of the EI. In particular the following formula has been introduced to compute the extremal index associated to time series in [5] (formula 10.26): we start with a trajectory $(x_0, Tx_0, \ldots, T^{N-1}x_0)$ and for a high enough threshold u we compute:

$$\hat{\theta}_m := \frac{\sum_{i=0}^{N-1-m} \mathbf{1}(\phi(T^i x_0) > u \cap \max_{j=1,\ldots,m} \phi(T^{j+i}x_0) \leq u)/(N-m)}{\sum_{i=0}^{N-1} \mathbf{1}(\phi(T^i x_0) \geq u)/N}. \tag{19}$$

If we are in presence of a periodic point of period m for the observable $\phi(x) = -\log d(z, x)$, then $\hat{\theta}_m$ is a good approximation of the EI.

We notice that when $m = 1$, Eq. (19) gives the likelihood estimator of Süveges $\hat{\theta}_{Su}$ [39], and the obtained result is an indicator of the persistence of the system close to a state z (see for example [17]). We remind that for a high quantile p of the time series, the Süveges estimator is given by:

$$\hat{\theta}_{Su} = \frac{\sum_{i=1}^{N_c}(1-p)S_i + N + N_c - \sqrt{(\sum_{i=1}^{N_c}(1-p)S_i + N + N_c)^2 - 8N_c \sum_{i=1}^{N_c}(1-p)S_i}}{2\sum_{i=1}^{N_c}(1-p)S_i}, \tag{20}$$

where N is the number of exceedances over τ, N_c is the number of clusters of two or more exceedances, and S_i is the size of the cluster i. This estimator is presented in [39]. A *Matlab* code and numerical details related to this computation in the context of dynamical systems are available in [13].

We warn that in presence of periodicity of order, say p, the estimator of Eq. (19) is reliable when it is strictly less than 1, because all the θ_m, for $m < p$ are equal to 1. This observation applies as well to the computation of the q_k: they are all zero up to q_{p-1}.

Ferreira [22] reviews several numerical algorithms to compute the EI (including those shown before). In this paper we choose to work with the q_k and we list a few features of this approach.

- In the presence of periodicity, we do not need to check condition Eq. (15); the clustering order, if any, is obtained automatically by progressing in the evaluation of the q_j. We will show in sections 4.1.2 and 4.2 examples of stationary processes constructed with random perturbation where several q_k are different from zero.
- Eq. (8) works as well when the ball $B(z, e^{-u_n})$ is replaced by any other measurable sets whose measure goes to zero when $n \to \infty$. This comes to modify suitably the observable ϕ. A concrete example will be given in section 5.
- Whenever one of the q_j is strictly positive, we get an EI strictly less than 1. This will be extensively used in sections 4.1.2, 4.2, 5.2. Notice that it is equivalent to find a certain $\theta_k < 1$ for $k \geq 1$.

In practice, one cannot compute all the q_j values and must stop at an order m. We stress that the sequence $\{q_j\}$ is rapidly decreasing to 0 and the quantity $\theta_m = 1 - q_0 - \cdots - q_{m-1}$, where m should be taken not too big, should give a fair estimate of θ. The approximate value $\hat{q}_j$ for q_j is obtained by considering a trajectory of length N as

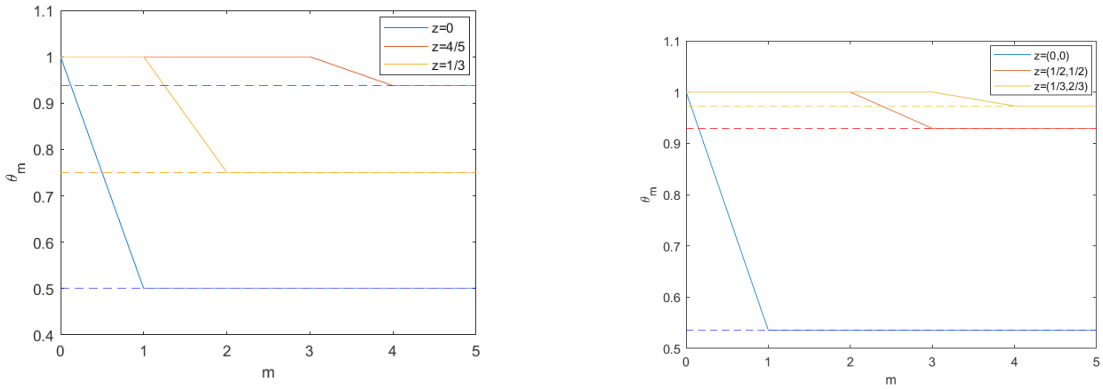
$$\hat{q}_j = \frac{\sum_{i=0}^{N-2-m} \mathbf{1}(\phi(T^i x_0) \geq u \cap \max_{l=1,\ldots,m} \phi(T^{l+i} x_0) < u \cap \phi(T^{i+j+1} x_0) \geq u)/(N-1-m)}{\sum_{i=0}^{N-1} \mathbf{1}(\phi(T^i x_0) \geq u)/N}. \tag{21}$$

2.2. **Numerical estimates of the extremal index in deterministic systems.** We present below the computations for different maps and various target points z and show the difference between the approximate estimate $\hat{\theta}_m$ of θ_m (we performed our computations at the order $m = 5$) and Süveges estimate (with order 0). We find that $\hat{\theta}_m$ obtains values close to the predictions of Eq. (9), while the other estimate gives 1 for points of period larger than 1.

Application	z	period	Theoretical value	$\hat{\theta}_{Su}$	$\hat{\theta}_5$	Uncertainty
$2x$ mod1	$4/5$	4	0.9375	1	0.9375	0.0024
$2x$ mod1	0	1	0.5	0.5007	0.5005	0.0028
$2x$ mod1	$1/3$	2	0.75	1	0.7508	0.0017
$2x$ mod1	$1/\pi$	not periodic	1	1	1	0
Cat's map	$(1/3, 2/3)$	4	0.9730 [9]	1	0.9732	$8.3.10^{-4}$
Cat's map	$(1/2, 1/2)$	3	0.9291 [9]	1	0.9292	$9.32.10^{-4}$
Cat's map	$(0, 0)$	1	0.5354 [9]	0.5352	0.5350	0.0018
Cat's map	$(1/\sqrt{2}, \pi - 3)$	not periodic	1	1	1	0

TABLE 1. Comparison of estimates of θ found with the different methods for the $2x-$mod1 map and Arnol'd cat map [3], defined in $\mathbb{T}^2$ by $T(x, y) = (x+y, x+2y)$. For all of the computations, we averaged our results over 20 trajectories of 5.10^7 points and took as a threshold the 0.999-quantile of the observable distribution. The uncertainty is the standard deviation of the results.

We observe in figure 1 that for a target point of period k, $\hat{\theta}_m$ is equal to 1 for $m < k$, is equal to θ for $m \geq k$, due to the fact that $\hat{q}_{k-1}$ is non zero. For $m \neq k$, we also have $\hat{q}_{m-1} = 0$, as expected. Although the Süveges estimator is not in principle suitable to detect periodic points with period strictly larger than 1, it is very often used to analyze time series since it is particularly simple to implement. But there is a more interesting reason. If the process is not periodic, for instance generated by random dynamical systems like those presented in the next sections, whenever the Süveges estimator is strictly less than one, the same happens for the EI, thanks to our Proposition 2.



(A) $2x \mod 1$ (B) Cat's map

FIGURE 1. Evolution of $\hat{\theta}_m$ with m, for different target points and different maps. The numerical values are found in table 4 and the parameter used in the text. The dashed line are the theoretical values of θ.

Instead, as we saw in the table above, the difference with real situations of larger periodicity could be important. In such cases, the estimate $\hat{\theta}_m$ performs systematically better. We will use this estimator in the numerical computations of this paper.
To evaluate the differences between the two estimates in high dimensional datasets, we test them on the North Atlantic daily sea-level pressure data described in [17]. For each atmospheric state z, the EI index associated to the observable $-\log \mathrm{dist}(z, .)$ was evaluated using the Süveges estimate and called the *inverse persistence* of z. The two estimations are shown in Figure 3
Estimates using $\hat{\theta}_5$ for these data are systematically lower than those with the order 0 method (in average 0.065 less), due to the contribution of some $\hat{q}_k$ for $k > 0$ (see the empirical distribution in figure 2). In Figure 3 we present a scatter plot of the daily values of θ obtained with the two estimators. We see that there is a strong linear relation between the two estimates with an offset of about 0.1 days^{-1}. The estimation of the cross-correlation coefficient, namely the zeroth lag of the normalized covariance function, yields 0.94, meaning that the information contained in the two estimators is the same except for a restricted set of states z. Moreover, if we remove the time averaged values $\langle \hat{\theta}_{Su} \rangle$ of the Süveges estimator and $\langle \hat{\theta}_5 \rangle$ of the new estimator from the time series, we can compare the distribution of $\hat{\theta}_{Su} - \langle \hat{\theta}_{Su} \rangle$ with that of $\hat{\theta}_5 - \langle \hat{\theta}_5 \rangle$. For this comparison, we have used a two sided Kolmogorov-Smirnov test and verified the null hypothesis that the estimators are from the same continuous distribution at the 5% confidence level.
The lowest differences between the two estimators are found for patterns close to the average field z. This corresponds to anomalies close to 0 hPa in Figure (3b). Instead, large differences between estimators correspond to a peculiar z pattern, consisting of a deep low pressure anomaly over the north of the domain and an anticyclonic anomaly over the southern part of the domain. This pattern resembles those observed in [17] for the minima of the local dimensions $d(z)$. This analysis confirms that the two estimators are different even beyond the simple shift in the values. The patterns where the two estimators mostly diverge can be thought as the pre-images of higher order periodic points of the attractor underlying the mid-latitude atmospheric circulation. In this sense, the combined use of both $\hat{\theta}_{Su}$ and $\hat{\theta}_5$ estimators could be useful to detect these special points of the dynamics in climate data and other natural phenomena.

3. THE RANDOM CASE: NON-STATIONARY SITUATIONS (QUENCHED NOISE)

The dynamics and the consequent detection of the extremal index could be affected in different ways. We begin with the dynamics by distinguishing first two cases:

3.1. **Sequential systems.** These systems are defined by concatenating maps chosen in some set, usually in the close neighborhood of a given map. As a probability measure one usually

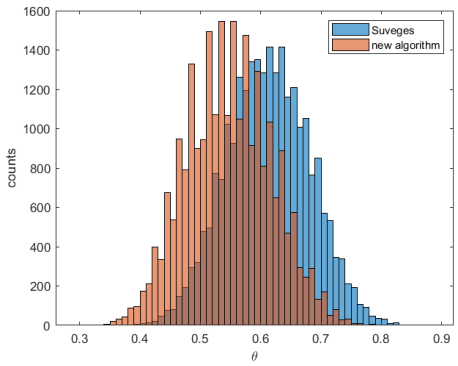


FIGURE 2. Comparison on the distribution of θ with the two estimates. The blue distribution is obtained using $\hat{\theta}_{Su}$ and the red one using $\hat{\theta}_5$ (for both taking as a threshold the 0.99-quantile of the observable distribution).

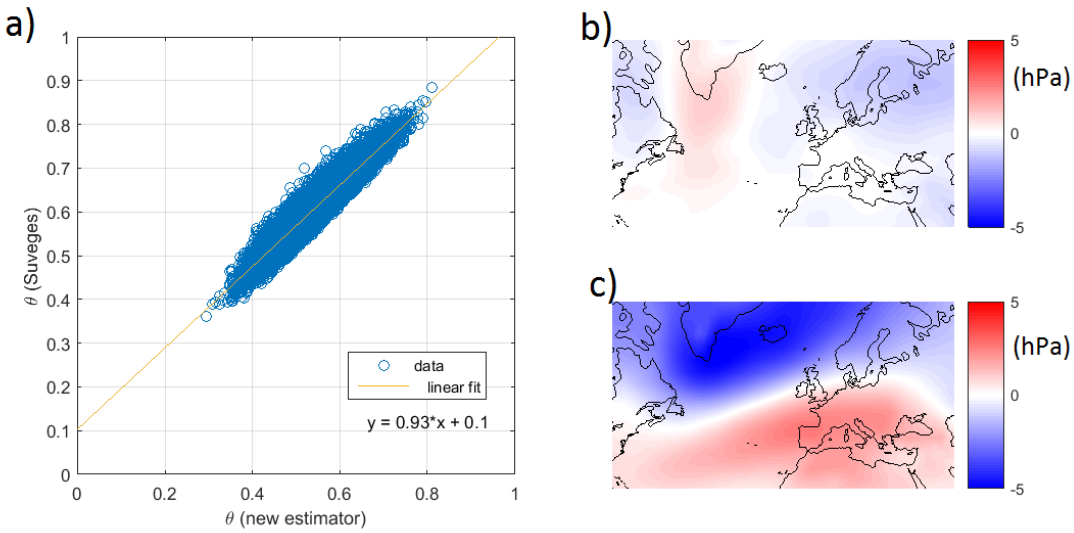


FIGURE 3. a) Scatter plot of $\hat{\theta}_{Su}$ (the Süveges estimator) vs $\hat{\theta}_5$ (the new estimator introduced in this work). Average map of the 5% sea-level pressure patterns such that the residual between $\hat{\theta}_{Su}$ and $\hat{\theta}_5$ are smaller than the 5% percentile (b) or larger (c) than the 95% percentile, for both taking as a threshold the 0.99-quantile of the observable distribution.

takes some ambient measure like Lebesgue, which makes the process given by the sequences of concatenations *non-stationary*. The extremal value theory and the extremal index must be redefined (see [25], in particular Eq. (2.8)). In section 4.5 of the aforementioned paper, we showed an example of a sequential system modeled on maps chosen in the neighborhood $\mathcal{U}$ of a given β-transformation T_β. According to suitable choices of $\mathcal{U}$ it is possible to show either that the EI of the unperturbed map and of the sequential system are the same, or the two differ, in particular when the elements of the concatenation are far enough from T_β. In this case the EI is simply 1. Sequential systems are a model for non-autonomous physical systems that do not admit a natural stationary probability measure. Statistical limit theorems can be proved as well in this context provided the observables are suitably centered (see [12, 29] for more details). The previous example with β-transformations suggests that when the maps are far enough from each other, the EI converges to 1.
To the best of our knowledge, there are not numerical investigations of extreme value behaviors for sequential systems. We propose one of them in the Appendix.

3.2. **Random fibred systems.** The sequential systems are in some sense too general, since there are no real prescriptions on the choice of the maps. A much more realistic class of non-autonomous systems is given by random fibred systems. They are constructed by taking a driving map σ that preserves a probability measure ν on the measurable space Ω, and which codes a family of transformations f_ω, for $\omega \in \Omega$ on the fiber X *via* the composition rule

$$f_\omega^n : x = f_{\sigma^{n-1}\omega} \circ \cdots \circ f_\omega. \tag{22}$$

The driving system encodes the external influences on the system of interest: it acts deterministically in the choice of the evolution transformations on the fibers. As L. Arnold wrote at the beginning of his monograph [2]:

Imagine a mechanism which at each discrete time n tosses a (possibly complicated, many-sided) coin to randomly select a mapping ϕ_n by which a given point x_n is moved to $x_{n+1} = \phi(x_n)$. The selection mechanism is permitted at time n to remember the choices made prior to n, and even to foresee the future. The only assumption made is that the same mechanism is used at each step. This scenario, called product of random mapping*, is one of the prototypes of a random dynamical systems.*

The mechanism used at each step is just the driving system. It is interesting to observe for the objectives of this paper, that random fibred systems have been used to analyze the transport phenomena in non-autonomous dynamical systems, such as geophysical flows (see the enlightening review article [27]). In order to perform statistics on these systems, we notice that a family of sample measures μ_ω lives on the fiber, which verifies the quasi-invariant equation $(f_\omega)_*\mu_\omega = \mu_{\sigma\omega}$, where $(f_\omega)_*$ is the push-forward of the measure. These sample measures will be taken as the probability measures that describe the statistical properties along the fiber and they do not give rise to stationary processes. The relation between the sample measures and ν is that the measure $\mu := \int \mu_\omega d\nu(\omega)$ is preserved by the skew (deterministic) transformation $F(x,\omega) = (f_\omega(x), \sigma\omega)$, acting on the product space $X \times \Omega$. The extremal index can be suitably defined and it could be different from 1 (see Corollary 5.1 and Theorem 5.3 in [25] dealing with *random subshift*). Transparent examples are constructed in the following subsections.

3.2.1. *Random Lasota-Yorke maps.* [23]. Take $\Omega = I^{\mathbb{Z}}$, where $I = (1, \ldots, m)$ is a finite alphabet with m letters. We associate to each letter a piecewise expanding map of the interval f_ω satisfying some smoothness and distortion standard assumptions. The map σ is therefore the bilateral shift and ν any ergodic shift-invariant non-atomic probability measure, for instance the Bernoulli measure with weights $p_1, \ldots, p_m$. The sequence of concatenations is given by $f_\omega^k = f_{\omega_k} \circ \cdots \circ f_{\omega_1}$, with $\omega_j \in I$ and $j = 1, \ldots, k$ are the first k symbols of the word ω.[7] In this setting it can proved that the sample measures are equivalent to the Lebesgue measure m and for m-almost any target point $z \in X$ and ν almost any realization $\omega \in \Omega$, the EI is equal to 1 (see [23]). This corresponds to a quenched result since it depends on the choice of the sequences ω.

3.2.2. *Rotations.* Another more physical quenched example is constructed in the following way. Let us take as Ω the unit circle $\mathbb{S}^1$ and as σ the irrational rotation: $\sigma(\omega) = \omega + \alpha - \text{mod}1$, $\omega \in \mathbb{S}^1$ with $\alpha \in \mathbb{R}$. Then we define $T(x) = 3x - \text{mod}1$ and make the correspondence:

$$\begin{aligned}
\omega \to f_\omega \text{ such that } f_\omega(x) &= T(x) + \omega - \text{mod}1. \\
\sigma\omega \to f_{\sigma\omega} \text{ such that } f_{\sigma\omega}(x) &= T(x) + \sigma\omega - \text{mod}1. \\
&\vdots \\
\sigma^k\omega \to f_{\sigma^k\omega} \text{ such that } f_{\sigma^k\omega}(x) &= T(x) + \sigma^k\omega - \text{mod}1.
\end{aligned}$$

and so on, by composing after that as: $f_{\sigma^k\omega} \circ \cdots f_\omega$.

[7]Notice that this is equivalent to Eq. (22) by defining the map $\omega \to f_{(\omega)'}$, where $(\omega)'$ is the ω_1 coordinate of ω.

This last example was implemented numerically using the estimate $\hat{\theta}_5$ for different trajectories and choices of α, either irrational or rational. We found an extremal index equal or very close to 1 in all cases (see table 2).

Remark 3. *In the preceding two examples we considered the extremal index, but we did not indicate any method to compute it. We are in fact outside the stationary regime which provided us with the various formulae described in the previous sections. Actually a definition of the EI in the non-stationary case is given in formula (2.8) in* [25]. *This formula relies also on a suitable definition of the boundary level u_n given in Eq. (2.2) in the aforementioned paper. The above formula (2.8) is a slight modification of the O'Brien formula and it could be reasonably recovered numerically by using Eq. (19) or the Süveges formula when we are not in presence of clustering. Although the EI was computed rigorously for the Random Lasota-Yorke maps as indicated in section 3.2.1, we do not dispose of a rigourous argument for the rotations and the numerical results we presented have been obtained with the $\hat{\theta}_5$ estimate. We observe that we could* assume *as a definition of the extremal index in these non-stationary situations the reciprocal of the expectation given by the distribution of the number of visits (see Eq. (46) in section 6). We will show in section 6.3 that the value for the EI computed in that way is the same as that provided by the $\hat{\theta}_5$ estimate.*

Trajectory	$\alpha = 1/\pi$	$\alpha = \sqrt{2} - 1$	$\alpha = 4/5$
1	0.9995	0.9995	1
2	0.9995	0.9996	1
3	0.9997	0.9998	1
4	0.9997	0.9995	1
5	0.9996	0.9996	1
6	0.9994	0.9996	1

TABLE 2. Values of θ found for different values of α and for different trajectories of length 10^7. We used the estimate $\hat{\theta}_5$ and the 0.999−quantile of the observable distribution as a threshold.

4. The random case: stationary situations (annealed noise)

The preceding two situations dealt with non-stationary processes. We now focus on stationary processes given (i) by i.i.d. randomly chosen transformations, and (ii) by i.i.d. moving target sets.

4.1. I.i.d. random transformations.

We now suppose that the maps f_ω are no longer driven by the measure-preserving map σ, but are chosen in an i.i.d. manner.

4.1.1. *Additive noise.* A common way to build such a process is to choose a map f and to construct the family of maps $f_\xi = f + \epsilon\xi$, where ξ a random variable sampled from some distribution $\mathbb{G}$. This distribution can be the uniform distribution on some small ball of radius ϵ around 0, where ϵ is the intensity of the noise. The iteration of the single map f is now replaced by the concatenation $f_{\xi_n} \circ \cdots \circ f_{\xi_1}$, where the ξ_k are i.i.d. random variables with distribution $\mathbb{G}$. If T has good expanding or hyperbolic properties, it is possible to show the existence of the so-called *stationary measure* ρ_s, verifying for any real bounded function g: $\int g d\rho_s = \int \int g \circ T_\xi d\rho_s d\mathbb{G}$, (see [13] Chap. 7, for a general introduction to the matter). If we now take the probability product measure $\mathbb{Q} := \mathbb{G}^{\mathbb{N}} \times \rho_s$, any process of type $\{g(f_{\xi_n} \circ \cdots \circ f_{\xi_1})\}_{n\in\mathbb{N}}$ is $\mathbb{Q}$-stationary. The extremal index is obtained again by applying Eq. (8) where the sets $\Omega_n^{(k)}(z)$ and $\mu(B(z, e^{-u_n}))$ are now weighted with the measure $\mathbb{Q}$. The expectation being taken with respect to $\mathbb{G}$ too, makes this an *annealed* type random perturbation, while the fibred perturbation described above is a *quenched* one, the expectations depending on the realization

ω. In the article [4] we rigorously proved that for a large class of piecewise expanding maps of the interval perturbed with additive noise with the uniform distribution $\mathbb{G}$, the extremal index is 1 as soon as ϵ becomes positive. The same happens for other kinds of maps and smooth distributions, as we showed in [14] almost numerically. We believe that this mostly happens for the systems perturbed with i.i.d. noise, since the latter destroys all the periodic points. Nevertheless it is possible to construct annealed examples giving an EI less than one. This will be done in the next section.

4.1.2. *Discrete noise.* Let us consider:

- Two maps $f_0 = 2x - \text{mod}1$ and $f_1 = 2x + b - \text{mod}1$, $0 < b < 1$.
- A *fixed* target set around 0: $H_n(0) := B(0, e^{-u_n})$, the ball of radius e^{-u_n} around 0. Notice that 0 is a fixed point of f_0, but not of f_1.

The indices $\xi_1, \dots, \xi_n$ in the concatenation $f_{\xi_n} \circ \cdots \circ f_{\xi_1}$ are i.i.d. Bernoulli random variables with distribution, for instance, $\mathbb{G} = p\delta_0 + (1-p)\delta_1$, with $p = \frac{1}{2}$. The spectral theory developed in [4] applies as well to this case, giving an absolutely continuous stationary measure ρ_s. Proposition 5.3 in [4] gives for q_0:

$$q_0 = \lim_{n\to\infty} \frac{\int d\mathbb{G}(\xi)\rho_s(H_n(0) \cap f_\xi^{-1} H_n(0))}{\rho_s(H_n(0))} = \lim_{n\to\infty} \left[\frac{1}{2}\frac{\rho_s(H_n(0) \cap f_0^{-1} H_n(0))}{\rho_s(H_n(0))} + \frac{1}{2}\frac{\rho_s(H_n(0) \cap f_1^{-1} H_n(0))}{\rho_s(H_n(0))}\right] \tag{23}$$

Since 0 is not a fixed point of f_1, a standard argument gives 0 to the limit in the square bracket on the right hand side. Instead by the mean value theorem applied to the term in the left square bracket we get in the limit of large n: $q_0 = \frac{1}{4}$.

Remark 4. *The preceding example could be interpreted as a perturbation of the map $f_0 = 2x - \text{mod}1$ around its fixed point 0. Instead we change the point of view and consider it as a perturbation of the map $f_1 = 2x + b - \text{mod}1$ around 0. Supposing the map f_0 is chosen with probability p_0, we see that when $p_0 = 0$ (absence of perturbation), the $EI = 1$, since 0 is not a fixed point for f_1, but it jumps to $EI \leq 1 - 0.5\, p_0$ as soon as $p_0 > 0$ and regardless of the value of b. This behavior is specular of what happened for the additive noise described above: there it was enough to switch on the noise, no matter of its magnitude, to make the EI equal to one. Here, the EI changes in term of the probability of appearance of the perturbed map, no matter of its topological distance from the unperturbed one, the value of $|b|$ in this case. However, it is easy to make the EI depending on the* distance *between the maps. For instance, let us take f_0 as $f_0 = (2+j)x - \text{mod}1$, $j \in \mathbb{N}$, which could be seen as a strong perturbation. By repeating the argument above we get that the $EI \leq 1 - p_0\, \frac{1}{2+j}$.*

The preceding computation of q_0 is valid for all values of $b \in (0,1)$. The question we may ask now is whether other q_k are non zero. By a similar reasoning as for the computation of q_0, this problem is equivalent to ask for the existence of a concatenation $f_{\xi_{k+1}} \circ \cdots \circ f_{\xi_1}$ returning the point 0 to itself after $k+1$ iterations and not before. Indeed, in that case we have that

$$q_k \geq \lim_{n\to\infty} \frac{\mathbb{G}(\xi_1, \dots, \xi_{k+1})\rho_s(H_n(0) \cap (f_{\xi_1}^{-1} H_n(0))^c \cap \cdots \cap (f_{\xi_k}^{-k} H_n(0))^c \cap f_{\xi_{k+1}}^{-k-1} H_n(0))}{\rho_s(H_n(0))} = \mathbb{G}(\xi_1, \dots, \xi_{k+1}))/2^k. \tag{24}$$

If the random orbit will not return to 0, a standard argument gives the value 0 for all q_k when $k \neq 0$.

To investigate the existence of such a concatenation, let us distinguish between the cases when b is irrational or rational.

If b is irrational and $\xi_1 \neq 0$, any concatenation of any length maps 0 into an irrational point, and so the orbit cannot come back to 0 after having left it. For this reason, all the q_k but q_0 are equal to 0.

We investigate the case where $0 < b = p/q < 1$ is rational, where $p, q \in \mathbb{N}$ are mutually prime.

We start by showing by induction that for $n \geq 1$ and for all $i \in [2^n, 2^{n+1}-1]$, the random orbit starting from p/q can attain the points $\frac{ip}{q} - \text{mod}1$ (and only them) in exactly n steps.
This proposition is true at rank $n = 1$, since $f_0(p/q) = 2p/q$, $f_1(p/q) = 3p/q - \text{mod}1$. Suppose now that it holds at rank n. Then we have that

- $f_0(2^n p/q - mod1) = 2^{n+1}p/q - \text{mod}1$
- $f_1(2^n p/q - mod1) = (2^{n+1}+1)p/q - \text{mod}1$

  $\vdots$
- $f_0((2^{n+1}-1)p/q - mod1) = (2^{n+2}-2)p/q - \text{mod}1$
- $f_1((2^{n+1}-1)p/q - mod1) = (2^{n+2}-1)p/q - \text{mod}1,$

and the proposition is true at rank $n+1$. This result implies that any point of the form $jp/q - \text{mod}1$, with $j \in \mathbb{N}, j \geq 2$ can be attained by the random orbit starting from the point p/q.
Let us consider $f_{\xi_m} \circ \cdots \circ f_{\xi_2} \circ f_1$ the concatenation of minimal length $m \geq 2$ such that $f_{\xi_m} \circ \cdots \circ f_{\xi_2} \circ f_1(0) = qp/q - \text{mod}1 = 0$. Note that the first map involved in this concatenation is f_1, so that the first iterate gives p/q. By the preceding argument, there exists such a concatenation and as we consider the smallest of the concatenations having this property, the random orbit leaving 0 does not come back to 0 before m iterations. As discussed earlier, this implies a non zero value of q_{m-1}.
In fact, there exists infinitely many non zero q_k: consider the concatenation of minimal length $m' \geq 2$, $f_{\xi'_{m'}} \circ \cdots \circ f_{\xi'_2} \circ f_1$ such that $f_{\xi'_{m'}} \circ \cdots \circ f_{\xi'_2} \circ f_1(0) = (q+1)p/q - \text{mod}1 = p/q - \text{mod}1$. By the argument in the induction, we have that either $m' = m$ or $m' = m+1$. Then we see that

$$f_{\xi_m} \circ \cdots \circ f_{\xi_2} \circ f_{\xi'_{m'}} \circ \cdots \circ f_{\xi'_2} \circ f_1(0) = f_{\xi_m} \circ \cdots \circ f_{\xi_2}(p/q) = 0.$$

This proves the existence of a concatenation returning the point 0 to itself after exactly either $2m$ or $2m+1$ iterations (and not before). Again we have proved that either q_{2m} or q_{2m-1} is non zero. Applying the same reasoning, we can prove that infinitely many q_k have positive values.
An interesting situation happens when $b = 1/2$. In this case, for all k, 0 can come back to itself after exactly $k+1$ iterations, and this happens only for the concatenation $f_1 \circ \cdots \circ f_1 \circ f_0$, where the map f_1 is applied succesively k times. This sequence having probability $(1-p)^k p = 1/2^{k+1}$, it is easily seen (by the very definition of q_k in Eq. (7) and analogously to the proof for the computation of q_0), that the q_k are non zero and are equal to:

$$q_k = \frac{1}{2^{k+1}} \lim_{n\to\infty} \frac{\rho_s(H_n(0) \cap (f_1^{-1}H_n(0))^c \cap \ldots \cap (f_1^{-k}H_n(0))^c \cap f_0^{-k-1}H_n(0))}{\rho_s(H_n(0))} = \frac{1}{4^{k+1}}.$$

In this situation, θ can be computed analytically and we find:

$$\theta = 1 - \sum_{k=0}^{\infty} (\frac{1}{4})^{k+1} = 2/3.$$

We computed in table 3 some estimates $\hat{q}_k$ and we find a very good agreement between theoretical and numerical estimates.

	Theoretical values	Estimate $\hat{q}_k$	Uncertainty
q_0	0.25	0.2500	0.0020
q_1	0.0625	0.0624	0.0016
q_2	0.015625	0.0156	$7.9.10^{-4}$
q_3	0.00390625	0.0039	$3.94.10^{-4}$
q_4	$9.765625.10^{-4}$	$9.321.10^{-4}$	$1.69.10^{-4}$

TABLE 3. Estimations of the q_k using $\hat{q}_k$ for the precedent example with $b = 1/2$. These results compare favorably with theoretical values to a precision of the order of 10^{-4}. Results are averaged over 20 trajectories of 5.10^6 points and we took the 0.999-quantile of the observable distribution as a threshold. The uncertainty is the standard deviation of the results.

4.2. **Moving target.** Another source of disturbance comes from the uncertainty of fixing the target ball for the hitting times of the orbit. A more general theory of random transformations and moving balls will be presented elsewhere. Here we simply consider the case of a deterministic dynamics and a target set shifted by a random displacement and we assume an annealed approach.

4.2.1. *Discrete noise.* Let us take a point $z \in X$; then we consider σ the one-sided shift on m symbols and set $\Omega = \{1, \cdots, m\}^{\mathbb{Z}}$. To each symbols, $j, j = 1, \dots, m$, there corresponds a $z(j) \in [z - \epsilon, z + \epsilon]$. As a probability $\mathbb{G}(\omega), \omega = (\omega_0, \omega_1, \dots) \in \Omega$, we put the Bernoulli measure of weights $p_i = \frac{1}{m}, i = 1, \dots, m$. It is invariant under the shift. Note that ϵ is fixed and is the strength of the uncertainty. At each temporal step the map T moves x and the target point moves as well around z under the action of σ. We now construct the direct product $S(x, \omega) := (T(x), \sigma\omega)$. The probability measure $\mathbb{P} := \mu \times \mathbb{G}$ is invariant under S. As in section 2 we are interested in the distribution of the random variable

$$H_n^r(x) := \{\text{First time the iterate } T^n(x) \text{ enters the ball } B(z(\sigma^n\omega)', e^{-u_n})\}, \tag{25}$$

where $\omega' = \omega_0$ is the first coordinate of ω. We define the *usual* observable:

$$\phi(x, \omega) = -\log|x - z(\omega')| = -\log|x - z(\omega_0)|,$$

and notice that for all $k \geq 1$:

$$\phi(S^k(x, \omega)) = -\log|T^k(x) - z((\sigma^k\omega)')| = -\log|T^k(x) - z(\omega_k)|.$$

The distribution of $H_n^r(x)$ is given by $\mathbb{P}(M_n \leq u_n)$, where

$$M_n(x, \omega) = \max\{\phi(x, \omega'), \phi(S(x, \omega)), \dots, \phi(S^{n-1}(x, \omega))\},$$

and the boundary levels u_n verify

$$n\mathbb{P}(\phi(x, \omega) > u_n) = n \int\int d\mathbb{G}(\omega)\mu(B(z(\omega'), e^{-u_n}) \to \tau$$

for some positive number τ. We now choose a map T on the unit interval admitting a spectral gap for the transfer operator and preserving an absolutely continuous invariant measure μ. That operator acts, for instance, on the space of bounded variation function on $[0, 1]$, which can be enlarged to a Banach space by adding to the total variation, $|\cdot|_{TV}$, the L^1 norm with respect to the Lebesgue measure m, $||\cdot||_{L^1(m)}$. We endow the space of functions $g(x, \omega)$ defined on $X \times \Omega$ with the Banach norm defined by:

$$||g||_{\mathcal{B}} := \int |g(\cdot, \omega)|_{TV} d\mathbb{G}(\omega) + ||g||_{L^1(\mathbb{G}\times m)}.$$

One can show that with respect to this norm the transfer operator $\mathcal{L}_S$ for the direct product S has a spectral gap on the largest eigenvalue 1. We introduce now the perturbed operator $\tilde{\mathcal{L}}_S(g(x, \omega) = \mathcal{L}_S(g(x, \omega)\mathbf{1}_{B^c(z(\omega'), e^{-u_n})}(x))$, where $g \in \mathcal{B}$ is a function defined on the Banach space just introduced. We first notice that

$$\mathbb{P}(M_n \leq u_n) = \int\int \tilde{\mathcal{L}}_S^n(h) dm d\mathbb{G},$$

where h is the density of μ, which, with our assumptions, is bounded from below by the constant C. Since $||(\tilde{\mathcal{L}}_S - \mathcal{L}_S)h||_{L^1(\mathbb{G}\times m)} \leq Ce^{-u_n}||h||_{\mathcal{B}}$, we can apply the perturbation theory mentioned in section 2, (see [32], [15], [13] ch. 7), and show that $\mathbb{P}(M_n \leq u_n)$ converges to the Gumbel

law $e^{-\theta\tau}$. The extremal index θ is given by the adaptation to the actual case of Eq. (7), with the q_0 which reads:

$$q_0 = \lim_{n\to\infty} \frac{\int d\mathbb{G}\ \mu\left(B(z(\omega'),n)\cap T^{-1}B(z(\sigma^{-1}\omega)',n)\right)}{\int d\mathbb{G}\ \mu(B(z(\omega'),n)}. \tag{26}$$

We now give a simple example for which $q_0 > 0$, which implies that the extremal index is strictly less than 1. Take as T the map $T(x) = 2x - \mathrm{mod}1$, and an alphabet of 4 letters: $\{0,1,2,3\}$ with equal weights 1/4. Moreover μ is the Lebesgue measure. Then we set the associations:

- $0 \to z_0$
- $1 \to z_1$
- $2 \to z_2$
- $3 \to z_3$,

where $z_i, i = 0,1,2,3$ are points in the unit interval verifying the following assumptions:

- $T(z_1) = T(z_2) = z_0;\ T(z_0) = z_3$.
- $T(z_3) \neq z_i, i = 0,1,2,3$.

The numerator of q_0 is:

$$\int_{[\omega_0=1,\omega_{-1}=0]} \mu\left(B(z_1,n)\cap T^{-1}B(z_0,n)\right) d\mathbb{G}+$$

$$\int_{[\omega_0=2,\omega_{-1}=0]} \mu\left(B(z_2,n)\cap T^{-1}B(z_0,n)\right) d\mathbb{G} + \int_{[\omega_0=0,\omega_{-1}=3]} \mu\left(B(z_0,n)\cap T^{-1}B(z_3,n)\right) d\mathbb{G}+$$

$$[\text{sum over all the other cylinders of length 2}],$$

where $[\omega_0,\omega_{-1}]$ denotes a cylinder with fixed coordinates ω_0 and ω_{-1}.
For the denominator we get the value:

$$\int d\mathbb{G}\mu(B(z(\omega'),n) = \sum_{l=1}^{n}\int_{[\omega_0=l]} d\mathbb{G}\mu(B(z(l),n) = \sum_{l=1}^{4} p_l\ \mu(B(z(l),n),$$

By using the same arguments as in Eq. (23), we see that for all the cylinders in the numerator different from the three explicitly given above, the integrals give zero in the limit of large n. For the first of the three cases above (the others being similar), we get

$$\int_{[\omega_0=1,\omega_{-1}=0]} \mu\left(B(z_1,n)\cap T^{-1}B(z_0,n)\right) d\mathbb{G} = \frac{1}{2}\int_{[\omega_0=1,\omega_{-1}=0]} \mu\left(B(z_0,n)\right) d\mathbb{G} = \frac{1}{2}(\frac{1}{4})^2\mu\left(B(z_0,n)\right).$$

The numerator therefore contributes with $\frac{3}{2}(\frac{1}{4})^2\mu\left(B(z_0,n)\right)$. The denominator is (by the Lebesgue translation invariance): $\mu\left(B(z_0,n)\right)$. Therefore we get $q_0 = \frac{3}{2}(\frac{1}{4})^2 = 0.09375$.
To get q_1, we need to compute the following quantity:

$$q_1 = \lim_{n\to\infty} \frac{\int d\mathbb{G}\ \mu\left(B(z(\omega'),n)\cap T^{-1}B(z(\sigma^{-1}\omega)',n)^c\cap T^{-2}B(z(\sigma^{-2}\omega)',n)\right)}{\int d\mathbb{G}\ \mu(B(z(\omega'),n)}. \tag{27}$$

We start, as before, by summing the integral in the numerator over all the possible cylinders of length 3. Among those cylinders, it is easy to check that only 6 of them contribute to the mass, namely ($[3,1,1],[3,2,1],[3,3,1],[3,1,2],[3,2,2]$ and $[3,3,2]$). For the integral associated to the first cylinder, we have (the five others are similar):

$$\int_{[3,1,1]} d\mathbb{G}\ \mu\left(B(z(\omega'),n)\cap T^{-1}B(z(\sigma^{-1}\omega)',n)^c\cap T^{-2}B(z(\sigma^{-2}\omega)',n)\right) = (\frac{1}{2})^2\int_{[3,1,1]} d\mathbb{G}\ \mu(B(z_3,n)$$

$$= (\frac{1}{4})^3(\frac{1}{2})^2\mu(B(z_3,n))$$

By summing over the six cylinders and dividing by the denominator, we get the result:

$$q_1 = 6(\frac{1}{4})^3(\frac{1}{2})^2.$$

We tested this result numerically with different sets of target points following the aforementioned assumptions. We find good agreements of $\hat{q}_0$ with the theoretical value of $q_0 = 0.09375$ and of $\hat{q}_1$ with the theoretical value of $q_1 = 0.0234375$ with a precision of order 10^{-3} (see table 4).

In the case where none of the points z_i is the kth iterate of another (in particular none of the points z_i is k-periodic), it can be shown that q_{k-1} is equal to 0 for $k > 2$. This is indeed what we find in our numerical simulations, which were performed up to the order 5.

z_0	$\hat{\theta}_5$	uncertainty	$\hat{q}_0$	$\hat{q}_1$
2/11	0.883	0.0017	0.0937	0.0231
10/13	0.883	0.0016	0.0936	0.0234
$1/\pi$	0.883	0.0022	0.0933	0.0234

TABLE 4. 5-order estimates of θ found for different z_0 (the other points z_1,z_2 and z_3 are computed to satisfy the assumptions of the presented example). We used as a threshold the 0.999-quantile of the observable distribution and the uncertainty is the standard deviation of the results. The values for $\hat{q}_0$ and $\hat{q}_1$ are averaged over the 20 trajectories and match with the theoretical results. All the other $\hat{q}_k$ computed are equal to 0.

This example is in some sense atypical since the four points $z_0, \ldots, z_3$ could be far away from each other. For instance, z_1 and z_2 are the two predecessors of z_0 and therefore they are on the opposite sides of $\frac{1}{2}$. By constraining the points to be in a small neighborhood of a given *privileged* center, the previous effects should be absent and the EI should be one, or very close to it. However our example shows that in the presence of moving target, it is not the periodicity which makes the EI eventually less than one. This gives another concrete example where the series in the spectral formula has at least two terms different from zero.

4.2.2. *Continuous noise.* We claimed above that when the center of the target ball can take any value in a small neighborhood of a given point z_0, the EI collapses to 1, even for periodic points z_0. To model this more physically realistic situation, let us fix $\epsilon > 0$ and consider the set $Z_\epsilon = [z_0 - \epsilon, z_0 + \epsilon]$. We define a map f acting on Z_ϵ with an associated invariant probability measure ν that drives the dynamics of the target point $z \in Z_\epsilon$. We suppose that ν is not atomic. The observable considered is now $\phi(x, z) = -\log|x - z|$ on the product space $\{X \times Z_\epsilon, \mu \times \nu\}$. By similar arguments to the ones we described for the discrete perturbation of the target point, we can show the existence of an extreme value law for the process $\phi \circ (T^k, f^k)$, with an EI given by Eq. (7), with

$$q_k = \lim_{n\to\infty} \frac{\int_{Z_\epsilon} d\nu(z)\mu\left(B(z,n) \cap T^{-1}B(f(z),n)^c \cdots \cap T^{-k-1}B(f^{k+1}(z),n)\right)}{\int_{Z_\epsilon} d\nu(z)\mu(B(z,n)}, \tag{28}$$

where $B(y, n)$ denotes a ball around y of radius e^{-u_n}. We have:

Proposition 3. *Suppose that for all $k \in \mathbb{N}$, $\nu(\{z \in Z_\epsilon | T^{k+1}(z) = f^{k+1}(z)\}) = 0$ and that μ is absolutely continuous with respect to Lebesgue with a bounded density h such that $h \geq \iota > 0$. Then the extremal index is 1.*

Proof. We compute the term q_k given by Eq. (28) and denote

$$Z_1^{n,k} = \{z \in Z_\epsilon | \mathrm{dist}(z, T^{-k-1}B(f^{k+1}(z), n)) > r_n\},$$

where $r_n = e^{-u_n}$. We can write the numerator in Eq. (28) as a sum of integrals over $Z_1^{n,k}$ and its complementary set $Z_2^{n,k} = Z_\epsilon \backslash Z_1^{n,k}$.

Since for $z \in Z_1^{n,k}$, z and z^* are at a distance larger than r_n, for all $z^* \in T^{-k-1}B(f^{k+1}(z), n)$, we have that $B(z, n) \cap T^{-k-1}B(f^{k+1}(z), n) = \emptyset$, and the integral over $Z_1^{n,k}$ is zero for all n.

It remains now to treat the integral over $Z_2^{n,k}$. We have that:

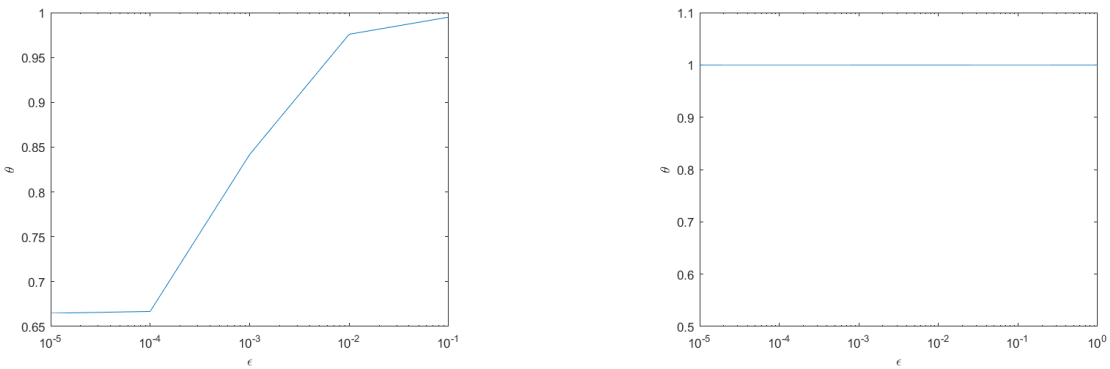
$$\begin{aligned} q_k &\leq \lim_{n\to\infty} \frac{\int_{Z_2^{n,k}} d\nu(z)\mu(B(z,n))}{\int_{Z_\epsilon} d\nu(z)\mu(B(z,n))} \leq \lim_{n\to\infty} \nu(Z_2^{n,k}) \frac{\sup_{x\in M}\mu(B(x,n))}{\inf_{x\in M}\mu(B(x,n))} \\ &= \lim_{n\to\infty} \nu(Z_2^{n,k}) \frac{\sup_{x\in M} h(x)}{\inf_{x\in M} h(x)}. \end{aligned} \tag{29}$$

With our assumptions on h, the fraction in the last term is finite. Let us consider the limit set $Z_2^k = \cap_n Z_2^{n,k} = \{z \in Z_\epsilon | T^{k+1}(z) = f^{k+1}(z)\}$. We have that $\lim_{n\to\infty} \nu(Z_2^{n,k}) = \nu(Z_2^k) = 0$, by hypothesis and so $q_k = 0$ for all k and therefore $\theta = 1$. □

Remark 5. *In practice, the requirement that $\nu(\{z \in Z_\epsilon | T^{k+1}(z) = f^{k+1}(z)\}) = 0$ for all k is satisfied for a large class of situations. We take two maps T and f intersecting at a countable number of points Z_2^1. We suppose that the preimage of any point by the applications T and f is a countable set. Then the set $Z_2^k = T^{-k-1}Z_2^1 \cup f^{-k-1}Z_2^1$ is countable as well for all $k \in \mathbb{N}$. Since ν is non-atomic, we have that $\nu(\{z \in Z_\epsilon | T^{k+1}(z) = f^{k+1}(z)\}) = 0$ is verified and the proposition holds.*

4.2.3. *Observational noise.* Instead of considering a perturbation on the target set driven by the map f, we consider the center of the target ball as a random variable uniformly distributed in the neighborhood of size ϵ of the point z_0. This situation models data with uncertainty or disturbances in their detection, and is equivalent to the case *observational noise* which we considered in [21] in the framework of extreme value theory. In this approach, the dynamics of a point x is given by $T^k x + \xi_k$, where ξ_k is a random variable uniformly distributed in a ball centered at 0 of radius ϵ. The process is therefore given by $-\log(dist(T^k x + \xi_k, z_0))$. When this is transposed in the moving target case, it becomes $-\log(dist(T^k x, z_0 + \xi_k))$ and the two approaches define the same process, which give the same stationary distribution. In [21] we proved rigorously in some cases and numerically in others, that for a large class of chaotic maps, an extreme value law holds with an extremal index equal to 1. By equivalence of the two processes, an extreme value law also holds in the moving target scenario and the EI is equal to 1.

We tested numerically the scenario of a target point following a uniform distribution in an interval of length ϵ centered in a point z_0, for several maps of the circle (Gauss map, $3x - \text{mod}1$, rotation of the circle). For all of them, we find that for periodic z_0, the extremal index which is less than 1 when unperturbed, converges to 1 as the noise intensity ϵ increases (see an example for the $3x - \text{mod}1$ map in figure 4-A). When z_0 is not periodic (we chose it at random on the circle), the EI is equal to 1 for all ϵ (see again figure 4-B).



(A) $z_0 = 1/2$ (fixed point) (B) Generic z_0

FIGURE 4. Influence of the intensity of noise ϵ on the EI for the $3x - \text{mod}1$ map perturbed by uniform noise. We simulated trajectories of 5.10^7 points and took $p = 0.999$. The computations are made using the estimate $\hat{\theta}_5$.

5. The dynamical extremal index

Up to now the target set was a ball around a point. In the attempt to describe the synchronisation of coupled map lattices, Faranda et al. [15] introduced the neighborhood of the diagonal in the n-dimensional hypercube and defined accordingly the extremal index related to the first time the maps on the lattices become close together. Faranda et al. [15] then used that approach to define a new type of observable, first in [20] and successively generalized in [8], which brought to a different interpretation of the extremal index, in terms of Lyapunov exponents instead of periodic behaviors.

Let us consider the $k-$fold ($k > 1$) direct product $(X, \mu, T)^{\otimes k}$ with the direct product map $T_k = T \circ \cdots \circ T$ acting on the product space X^k and the product measure $\mu_k = \mu \circ \cdots \circ \mu$. Let us now define the observable on X^k

$$\phi(x_1, x_2, \ldots, x_k) = -\log(\max_{i=2,\ldots,k} d(x_1, x_i)), \tag{30}$$

where each $x_i \in X$. We also write $\overline{x}_k := (x_1, x_2, \ldots, x_k)$ and $T_k(\overline{x}_k) = (Tx_1, \ldots, Tx_k)$. As we explained in [20], if we put $M_n(\overline{x}_k) = \max\{\phi(\overline{x}_k), \ldots, \phi(T_k^{n-1}(\overline{x}_k)\}$ and look at the distribution of the maximum $\mu_k(M_n \leq u_n)$, this distribution is non-degenerate and converges to the Gumbel law for $n \to \infty$ $e^{-\theta_k \tau}$, provided we can find a sequence $u_n \to \infty$ and verifying $\mu_k(\phi > u_n) \to \frac{\tau}{n}$, where τ is a positive number. The quantity θ_k is our new extremal index and it will be described later on. Notice first that

$$\mu_k(\phi > u_n) = \int_{M^k} d\overline{x}_k \mathbf{1}_{B(x_1, e^{-u_n})}(\overline{x}_k) \ldots \mathbf{1}_{B(x_1, e^{-u_n})}(\overline{x}_k) = \int_M dx_1 \mu(B(x_1, e^{-u_n}))^{k-1}. \tag{31}$$

Our first motivation to investigate the observable in Eq. (30) was the fact that the integral on the right hand side of Eq. (31) scales like $e^{-u_n D_k(k-1)}$, where D_k denotes the generalized dimension of order k of the measure μ. We now describe θ_k. We first define:

$$\Delta_n^k = \{(\overline{x}_k), d(x_1, x_2) < e^{-u_n}, \ldots, d(x_1, x_k) < e^{-u_n}\}. \tag{32}$$

By using the spectral technique in [20], and the analytical results of [15], it is possible to show that:

$$\theta_k = 1 - \lim_{n \to \infty} \frac{\mu_k(\Delta_n^k \cap T_k^{-1} \Delta_n^k)}{\mu_k(\Delta_n^k)}. \tag{33}$$

The quantity θ_k is the *dynamical extremal index* (DEI) appearing at the exponent of the Gumbel law.

For C^2 expanding maps of the interval, which preserve an absolutely continuous invariant measure $\mu = hdx$ with strictly positive density h of bounded variation, it is possible to compute the right hand side of Eq. (33) and get:

$$\mu_k(\Delta_n^k \cap T_k^{-1} \Delta_n^k) = \int dx_1 h(x_1) \int dx_2 h(x_2)\ \chi_{B(x_1, e^{-u_n})}(x_2) \chi_{B(Tx_1, e^{-u_n})}(Tx_2) \cdots$$
$$\cdots \int dx_k h(x_k) \chi_{B(x_1, e^{-u_n})}(x_k) \chi_{B(Tx_1, e^{-u_n})}(Tx_k). \tag{34}$$

All the $k-1$ integrals above factorize, and depend on the parameter x_1. Therefore they can be treated as in the proof of Proposition 5.5 in [15], yielding the rigorous result:

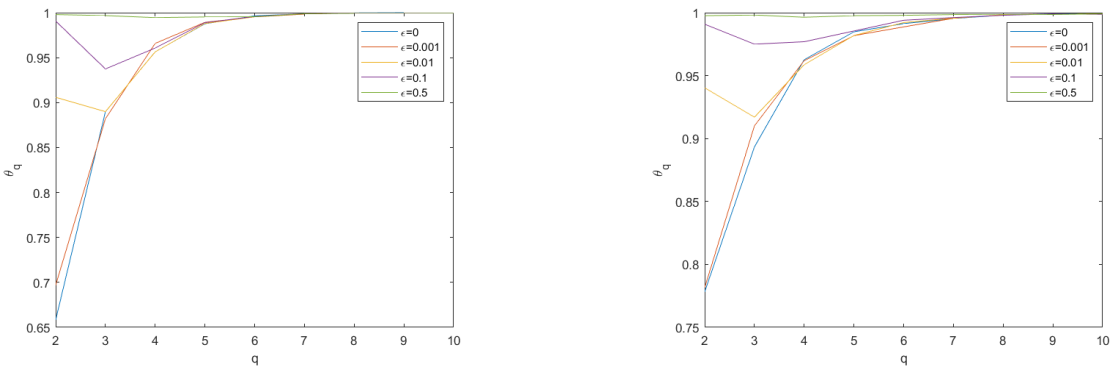
Proposition 4. *Suppose that: the map T belongs to C^2; it preserves an absolutely continuous invariant measure $\mu = hdx$, with strictly positive density h of bounded variation; it verifies conditions $P1 - P5$ and $P8$ in* [15][8]. *Then*

$$\theta_k = 1 - \frac{\int \frac{h(x)^k}{|DT(x)|^{k-1}} dx}{\int h(x)^k dx}. \tag{35}$$

[8]These conditions essentially ensure that the transfer operator associated with the map T has a spectral gap and that the density h has finite oscillation in the neighborhood of the diagonal.

This formula uses the translational invariance of the Lebesgue measure: we refer to sections II-B and II-C in [20] for analogous extensions to more general invariant measures and to SRB measures for attractors. As remarked in [20], whenever the density does not vary too much, or alternatively the derivative (or the determinant of the Jacobian in higher dimensions) are almost constant, we expect a scaling of the kind: $\theta_k \sim 1 - e^{-(k-1)h_m}$, where h_m is the metric entropy (the sum of the positive Lyapunov exponents).
In section D of [20] and for the case $k = 2$, we replaced the iterations of a single map with concatenations of i.i.d. maps chosen with additive noise, see above, section 4.1. Although the dynamical extremal index is related to the Lyapunov exponents, it is also influenced by the fact that the set $\Delta = \cup_{i=2}^{k}\{x_i = x_1\}$ is invariant in the deterministic situation. By looking at Eq. (33), we see that we estimate the proportion of the neighborhood of the invariant set Δ returning to itself. As argued in [20], that estimate gives information on the rate of backward volume contraction in the unstable direction. Since the noise generally destructs these invariants sets, we expect the extremal index be equal to 1 or quickly approaching 1 for all k when the noise increases. The situation depends again on the characteristics of the noise.

5.1. **Additive noise.** The prediction at the end of the preceding section is confirmed by concatenating maps perturbed with additive noise and smooth probability density function (see the numerical computations shown in figure 9). As in section 4.1.2 the perturbation is of annealed type. For each map, trajectories of 10^7 points were simulated and the 0.999−quantile of the distribution of the observable ϕ was selected as a threshold. As only the q_0 term is non-zero, we choose the Süveges estimate for our computations, to be assured that no error coming from $\hat{q}_k$ terms of higher order are added.



(A) $3x - mod1$ (B) $\frac{1}{x} - mod1$ (Gauss map)

FIGURE 5. Influence of the intensity of noise ϵ on the θ_k spectrum for maps of the circle perturbed by uniform noise. We used the Süveges estimate and the 0.999-quantile of the observable distribution as a threshold, as mentioned in the text.

5.2. **Discrete noise.** A different scenario happens if we look at sequences of finitely maps chosen in an i.i.d. way according to a Bernoulli process as we did in section 4.1.2. We choose now the two maps $f_0(x) = 3x - \text{mod}1$ and $f_1(x) = 3x + b - \text{mod}1$ and we put the same distribution $\mathbb{G}$ defined in 4.1.2. A combination of Eqs. (23) and (33) gives for the term q_0 entering the infinite sum defining the DEI:

$$q_0 = \lim_{n\to\infty} \frac{\int_\Omega \rho_k(\Delta_n^k \cap f_{\xi^k}^{-1}\Delta_n^k) d\mathbb{G}^k(\xi^k)}{\rho_k(\Delta_n^k)}, \tag{36}$$

where: ξ^k is the vector $\xi^k = (\xi_1, \ldots, \xi_k)$, $f_{\xi^k}^{-1} = f_{\xi_1}^{-1} \cdots f_{\xi_k}^{-1}$ and ρ_k is the stationary measure constructed as in [4]. We now split the integral over the cylinders $\xi^{k,0} = [\xi_1 = 0, \ldots, \xi_k =$

$0], \xi^{k,1} = [\xi_1 = 1, \dots, \xi_k = 1]$ and their complement. Therefore we have:

$$\frac{\int_\Omega \rho_k(\Delta_n^k \cap f_{\xi^k}^{-1}\Delta_n^k) d\mathbb{G}^k(\xi)}{\rho_k(\Delta_n^k)} = \mathbb{G}(\xi^{k,0})\frac{\rho_k(\Delta_n^k \cap f_0^{-n}\Delta_n^k)}{\rho_k(\Delta_n^k)} + \mathbb{G}(\xi^{k,1})\frac{\rho_k(\Delta_n^k \cap f_1^{-n}\Delta_n^k)}{\rho_k(\Delta_n^k)} + \tag{37}$$

$$\{\text{the integrals are computed over all the other cylinders of length } k\}. \tag{38}$$

The two fractions in Eq. (37) are equal to the $q_0^{(\text{unp})}$ of the unperturbed systems, which are the same by the particular choices of the maps and by Eq. (35).
All the other terms in Eq. (38) are zero, we now explain why. Let us in fact consider a vector $\tilde{\xi}^k$ different from $(0, 0, \dots, 0)$ and $(1, 1, \dots, 1)$. Suppose that the first coordinate $\tilde{\xi}_1^k$ of $\tilde{\xi}^k$ is zero and let $\tilde{\xi}_j^k \neq 0$ for some $1 < j \leq k$. If $x_k \in \Delta_n^k$, and since $f_0 x_1 = 3x_1 - \text{mod}1$, $f_1 x_j = 3x_j + b - \text{mod}1$, we have that $d(f_0 x_1, f_1 x_i) \to_{n\to\infty} \min\{b, 1-b\}$ as $d(x_1, x_i) < e^{-u_n}$.
Therefore when n is large enough, $\forall x_k \in \Delta_n^k$, $f^{\tilde{\xi}^k} x_k := f^{\tilde{\xi}_k^k} \cdots f^{\tilde{\xi}_1^k} x_k \notin \Delta_n^k$ and $\Delta_n^k \cap f_{\tilde{\xi}^k}^{-1}\Delta_n^k = \emptyset$.
In conclusion we have

$$\theta_k = 1 - q_0^{(\text{unp})}(p^k + (1-p)^k), \tag{39}$$

where p is the weight of the Bernoulli measure. This last formula has been tested numerically with the maps f_0 and f_1, for different values of b and with $p = 0.5$. Good agreement is found for different values of b (see figure 8). We use trajectories of 5.10^7 points and the $0.995-$ quantile of the observable distribution as a threshold. Estimations are made using the Süveges estimate, for the same reasons as described earlier.

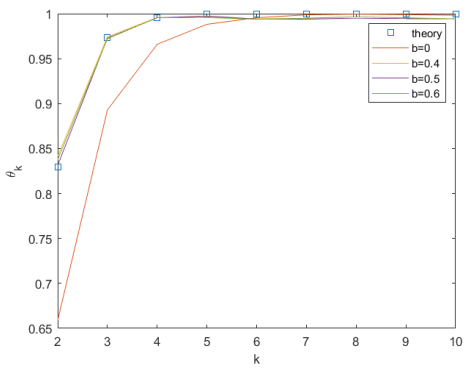


FIGURE 6. θ_k spectrum for the sequence of two maps described in the example, found using the Süveges estimate. The red curve is the value for $b = 0$, and for different choices of $b \neq 0$, the values of θ_k follow the predictions of Eq. (39) (blue squares).

6. Point processes: statistics of persistence

We said in section 2 that the extremal index is the inverse of the average cluster size when we look at exceedances in the shrinking neighborhood of a given target point z. To be more precise let us start with the deterministic case.

6.1. The deterministic case.

We consider the following counting function

$$N_n(t) = \sum_{l=1}^{\lfloor \frac{t}{\mu(B(z,r_n))} \rfloor} \mathbf{1}_{B(z,r_n)}(T^l(x)), \tag{40}$$

where the radius r_n goes to 0 when n tends to infinity. We are interested in the distribution

$$\mu(N_n(t) = k), \ k \in \mathbb{N}. \tag{41}$$

for $n \to \infty$. It has been proved (see for instance [30, 24]) that when z is not a periodic point, $\mu(N_n(t) = k)$ converge to the Poisson distribution $\frac{t^k e^{-t}}{k!}$, while for a periodic point of minimal period p we have the Polyà-Aeppli distribution

$$\mu(N_n(t) = k) \to e^{-\theta t} \sum_{j=1}^{k} (1-\theta)^{k-j} \theta^j \frac{(\theta t)^j}{j!} \binom{k-1}{j-1}, \tag{42}$$

where θ is the extremal index.
We now invoke a very recent theory developed by [31], which allows us to put the statistics of the number of visits in a wider context and to make explicit the connection with the EI, which is the main object of this note.
We first replace the ball $B(z, r_n)$ with a sequence of sets $B_n(\Gamma)$ whose intersection over $n \in \mathbb{N}$ gives the zero measure set Γ. For instance in section 5 such sets are the neighborhoods Δ_n^k. We keep in mind this example for the next application. We then define the *cluster size distribution* π_l as

$$\pi_l := \lim_{n\to\infty} \lim_{K\to\infty} \frac{\mu(\sum_{i=0}^{2K} \mathbf{1}_{B_n(\Gamma)} \circ T^i = l)}{\mu(\sum_{i=0}^{2K} \mathbf{1}_{B_n(\Gamma)} \circ T^i \geq 1)}. \tag{43}$$

Under a few smoothness assumptions for the map T, which are verified for the majority of what we called chaotic systems in this note, it is possible to prove ([31], Th. 1) that

$$\mu\left(\sum_{l=1}^{\lfloor \frac{t}{\mu(B_n(\Gamma))} \rfloor} \mathbf{1}_{B_n(\Gamma)}(T^l(x)) = k \right) \to \nu(\{k\}), \tag{44}$$

as $n \to \infty$, where ν is the compound Poisson distribution for the parameter $t\pi_l$. This distribution can be recovered by its generating function $g_\nu(z) = \exp \int_0^\infty (z^x - 1) dH(x)$, where H is the counting measure on $\mathbb{N}$ defined as $H = \sum_l t\pi_l \delta_l$, with δ_l the point mass at l. The Polyà-Aeppli distribution is a particular case of compound Poisson distribution with the π_l having a geometric distribution $\pi_l = \theta(1-\theta)^l$. The interesting feature of the theory in [31] is that the π_l can be expressed in terms of more accessible quantities, in particular of the *distribution function* $\hat{\alpha}_l$ *of the* $(l-1)$ *return into the set* $B_n(\Gamma)$ *for* n *tending to infinity*, namely $\pi_l = \frac{\hat{\alpha}_l - 2\hat{\alpha}_{l+1} + \hat{\alpha}_{l+2}}{\hat{\alpha}_1 - \hat{\alpha}_2}$, where

$$\hat{\alpha}_\ell = \lim_{K\to\infty} \lim_{n\to\infty} \mu_{B_n(\Gamma)}(\tau_{B_n(\Gamma)}^{\ell-1} \leq K), \tag{45}$$

being $\tau_{B_n(\Gamma)}^{\ell}(x) = \tau_{B_n(\Gamma)}^{\ell-1}(x) + \tau_{B_n(\Gamma)}(T^{\tau_U^{\ell-1}}(x))$, with $\tau_{B_n(\Gamma)}^1 = \tau_{B_n(\Gamma)} = \min\{j \geq 1 : T^j(x) \in B_n(\Gamma), x \in B_n(\Gamma)\}$, the first return time. Moreover if θ denotes the extremal index (as defined in section 5 for general target sets), we have the identity:

$$\sum_{k=1}^{\infty} k\pi_k = \frac{1}{\theta}. \tag{46}$$

This result is a generalization of Eq. (10) stated in section 2 and which targeted neighborhoods of periodic points.
We now apply Eq. (46) to the dynamical extremal index introduced in section 5 in the case of direct product of two maps. For the class of maps verifying Proposition 35 we have the general result for the $\hat{\alpha}_l$ [31]:

$$\hat{\alpha}_{l+1} = \frac{1}{\int_I h^2(x)\, dx} \int_I \frac{h^2(x)}{|DT^l(x)|}\, dx.$$

This formula was also previously obtained by [11]. In order to make rigorous computations, we consider a Markov map of the interval T for which the density h is piecewise constant (see for

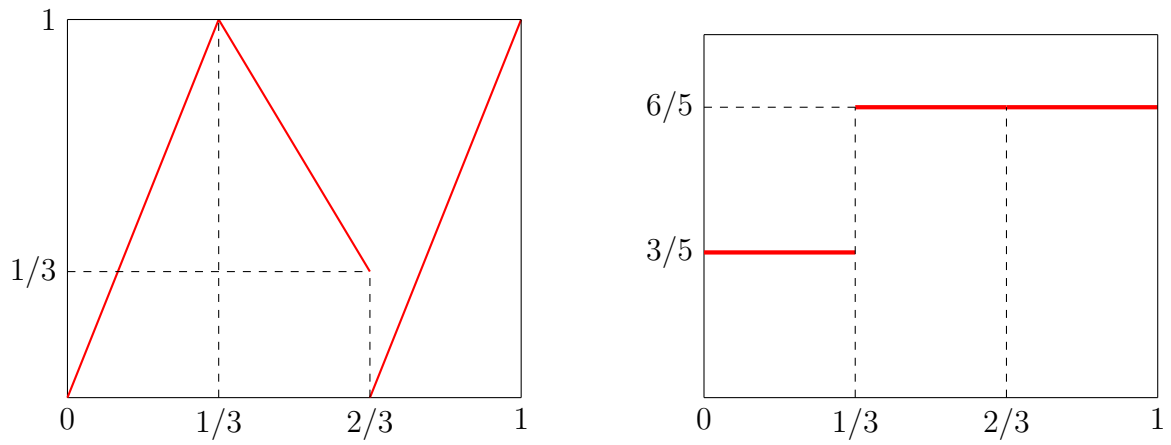


FIGURE 7. Graphical representation of the Markovian map (left) and its associated invariant density (right).

instance [28]):

$$\begin{aligned} T(x) &= 3x \ , \quad x \in I_1 := [0, 1/3) \\ &= 5/3 - 2x \ , \quad x \in I_2 := [1/3, 2/3) \\ &= -2 + 3x \ , \quad x \in I_3 := [2/3, 1). \end{aligned}$$

The density reads

$$\begin{aligned} h(x) &= 3/5 =: h_1 \ , \quad x \in I_1 \\ &= 6/5 =: h_2 \ , \quad x \in I_2 \\ &= 6/5 =: h_3 \ , \quad x \in I_3. \end{aligned}$$

The extremal index is given by the integral in Eq. (35) which has already been computed in [8] and gives $\theta \approx 0.5926$. The quantities $\hat{\alpha}_l$ can also be computed explicitly. We defer to [26] for the details, and obtain:

$$\hat{\alpha}_{k+1} = \frac{\left[(3\sqrt{145} - 37)(17 - \sqrt{145})^k + (37 + 3\sqrt{145})(17 + \sqrt{145})^k\right]}{2 \times 72^k \sqrt{145}}.$$

The interesting point is that the π_k do not follow a geometric distribution. For the statistics of the number of visits, we get a compound Poisson distribution which is not Polyà-Aeppli. To illustrate that, we show the distributions $\mu(N_n(t)) = k$ for $t = 50$ and different values of k for the real process associated to the markov map T, and the distribution obtained by supposing a Polyà-Aeppli with parameter equal to 0.5926. In this context, the sets $B_n(\Gamma)$ are the tubular neighborhoods of the diagonal Δ_n^2 defined in Eq. (32). For our numerical simulation, we fix a small set Δ_n^2 by taking u_n in Eq. (32) equal to the 0.99−quantile of the distribution of the observable computed along a pre-runned trajectory of 10^4 points. We obtain that for a small enough neighborhood of the diagonal, $\mu_2(\Delta_n^2) \approx (h_1^2 + h_2^2 + h_3^2)\sqrt{2} \exp(-u_n)$.

6.2. **The random case.** There are only few results on the statistics of the number of visits in the case of random systems. In the forthcoming paper [23], we show that the quenched Lasota-Yorke example studied in section 3.2.1 gives a Poisson distribution for the visits around almost all target point z. A Polyà-Aeppli distribution is seen to hold for some particular random subshits.

We are not aware of a rigorous treatment of these point processes for stationary (annealed) random systems. We believe that the theory developed in [31] could be applied as well to those systems. In particular we expect to get the extremal index as the inverse of the average cluster size. To corroborate this claim, we computed the statistics of the number of visits for two

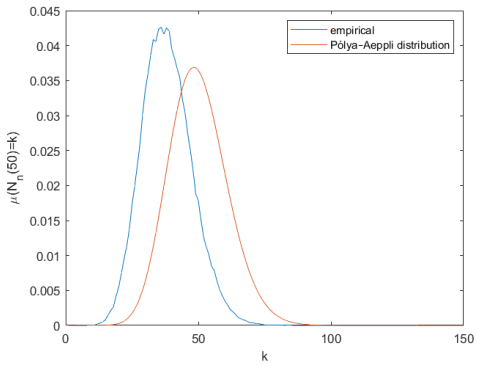
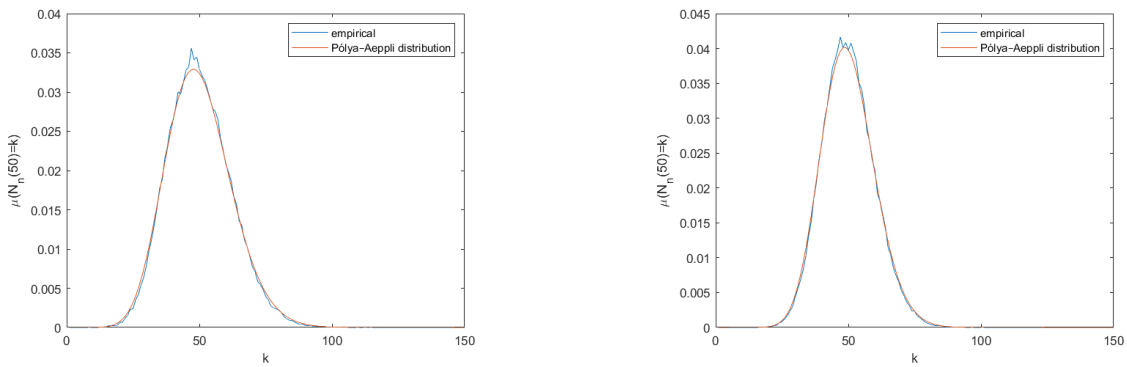


FIGURE 8. Comparison between the empirical distribution $\mu(N_n(50) = k)$ and the Polyà-Aeppli distribution of parameters $t = 50$ and $\theta = 0.5926$. The empirical distribution is obtained by considering 10^5 runs for which is recorded the number of entrances in the set Δ_n^2 ,as defined in the text, after $t/\mu_2(\Delta_n^2)$ iterations of the map.

examples: the case in section 4.1.2 for the composition of two maps with $b = 1/2$, for which there are infinitely many non zero q_k, and the moving target example in section 4.2. In both cases we found that the real distribution coincide with Polyà-Aeppli with parameter given by $t = 50$ and the extremal index of respectively is $2/3$ and $1 - \frac{3}{2}(\frac{1}{4})^2 - 6(\frac{1}{4})^3(\frac{1}{2})^2 \approx 0.8828$. These results are stable against different values of t.



(A) (B)

FIGURE 9. For the composition of two maps with $b = 1/2$ (left) and the moving target example (right), comparison between the empirical distribution $\mu(N_n(50) = k)$ and the Polyà-Aeppli distribution of parameters $t = 50$ and respectively $\theta = 2/3$ and $\theta = 0.8828$. The empirical distribution is obtained as for the Markov maps, by considering 10^5 runs. For both examples, we considered target balls of radius $r = \exp(-u)$, where u is the 0.99-quantile of the empirical distribution of the observable computed along a pre-runned trajectory of size 10^4.

6.3. **The rotational case.** We now compute the statistics of the number of visits for the process described in section 3.2.2. The statistics are computed in the same way as for the preceding examples and the theoretical value of 1 for the extremal index suggests that the number of visits is purely Poissonian. This is indeed what is observed in figure 10. Again the results are stable against different values of t.

6.4. **Climate data.** We study now the statistics of the number of visit for the atmospheric data presented in section 2.1. The target set considered in the analysis is a ball centered at a point z that corresponds to the mean sea-level pressure recorded on Jan 5th 1948. The radius

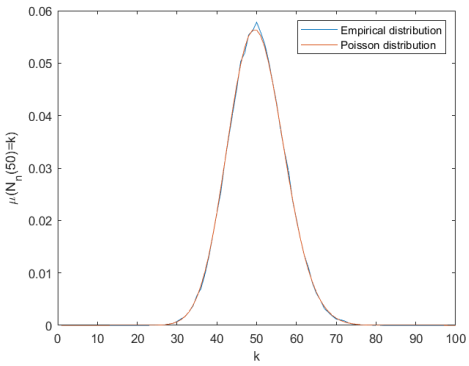


FIGURE 10. Comparison between the empirical distribution $\mu(N_n(50) = k)$ and the Poisson distribution of parameter $t = 50$, computed with a trajectory of length 5.10^7 of the rotational system described in section 3.2.2. We used the same parameters and procedure as the one in figure 9.

$r = e^{-\tau}$, is fixed using as τ the 0.95−quantile of the observable distribution. We considered the number of visits in an interval of length 365 days, that corresponds to the annual cycle[9] so that the parameter t in Eq. (40) is taken equal to $t = \lfloor 365\mu(B(z,r))\rfloor$. For the target set considered, we get $t = 18$, that is close to the mean value that we find for the number of visits of $18,13$. In figure 11, we compare the obtained empirical distribution with the Polyà-Aeppli distributions of parameters t and the EI obtained using different estimates: Süveges, $\hat{\theta}_5$ and $\hat{\theta}_{10}$. The distribution using the Süveges estimate seems to fit the empirical data better than the other two. In fact if we increase the order of the estimate $\hat{\theta}_m$, the value it attributes to the EI decreases and the associated Polyà-Aeppli distribution flattens. We believe that this is due to finite effects; indeed we observe that if the fixed threshold u in formula 19 is too small, the numerator approaches 0 as m increases, yielding bad estimates for the EI. Qualitatively similar results are found for different atmospheric states and different quantiles considered (although taking higher quantiles gives too few data for the statistics), suggesting that a Polyà-Aeppli distribution for the number of visits holds for this type of high dimensional observational data. This analysis, applied systematically to a state of a complex system, can provide important information on the recurrence properties of the target state, namely how many visit to expect in a given time interval. In this sense, the extremal index only provides an averaged information. The drawback is in the number of data required to provide a reliable estimates of the visits distribution (see Figure 11). This problem could be however overcome by averaging the distribution of visits for class of events (several states z) instead of considering just one state.

7. CONCLUSIONS

In this paper we have reviewed the meaning of the extremal index in dynamical systems and provided numerical and theoretical strategies for its estimation. We have detailed how, depending on the nature (stochastic or deterministic) of the underlying system the estimators may provide different insights on the dynamics of the system. Our main findings can be summarized as follows:

- As soon as a deterministic system is perturbed with noise having a smooth distribution, the EI becomes equal to 1. With discrete (point masses) distributions, the EI could be 1 for almost choices of the target and of the realisations, (see sections 3.2.1, 3.2.2, 5.1), but it could also be less than 1, (see sections 4.1.2, 5.2).
- What we described in the previous item happens either with annealed and quenched stochastic perturbations. In the annealed case with discrete probability distributions,

[9]This time scale is relevant for the underlying dynamics of the atmospheric circulation.

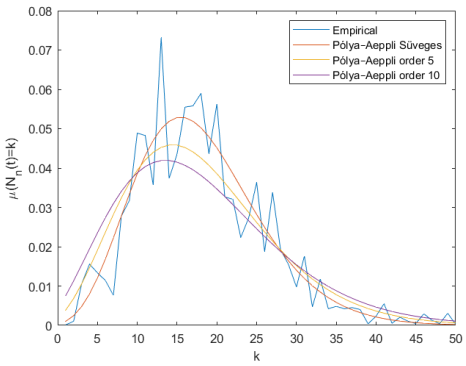


FIGURE 11. Comparison between the empirical distribution $\mu(N_n(t) = k)$ and the Polyà-Aeppli distributions of parameters $t = 18$ and the EI found with different estimates. The empirical distribution is obtained by considering 24396 intervals of length 365 and counting the number of visits inside the target ball described in the text.

the value of the EI could depend either on the choice of the masses and on the closeness of the maps (see Remark 4). This suggests that in the situations where the EI is generically 1, if one finds a lesser value, this could be the effect of a discrete perturbation.

- The situation where the dynamics is deterministic, but the target set is known with some probability with a continuous distribution function is interesting, as discussed in section 4.2.2. This describes physical situations where the localisation and observation of the extreme event are perturbed. No matter of the intensity of those perturbations, the extremal index is everywhere equal to one.
- The statistics of the number of visits (*persistence*) in the shrinking target set gives rise to point processes of Poisson type. We have shown that the geometric nature of the target set produces different compound Poisson distributions. Moreover these compound Poisson distributions persist even when the system is randomly perturbed. Indeed the fact that the extremal index is the inverse of the average cluster size seems to be a robust property of chaotic systems in the deterministic and random settings.
- In presence of periodicity one expects asymptotic distributions of Polyà-Aeppli type. Nevertheless for target sets composed by more general invariant sets, different compound Poisson distribution could emerge up as we showed above. This suggest that for the statistics of the number of visits around sets with a more complicated geometry, and this could be relevant in applications, the computation of the EI is probably not enough, and one should go to higher moments to get the real asymptotic distribution.

8. APPENDIX

As mentioned in section 3, results proving the existence of extreme value laws for non-stationary sequential systems are lacking. Nevertheless, the methods of estimation described in the paper can still be used to evaluate a quantity θ that would correspond to the extremal index in the eventuality that an extreme value law holds for this kind of systems, which by the way was the case in [25]. We could also alternatively define the extremal index by computing the statistics of the number of visit and evaluating the expectation in Eq. (46). We choose this second option and consider the motion given by the concatenation

$$f_{\bar{\xi}}^n(x) = f_{\xi_n} \circ \cdots \circ f_{\xi_1}(x),$$

where the probability law of the ξ_i changes over time. In particular, we consider the 10 maps

$$f_i(x) = 2x + b_i \mod 1,$$

where b_i is the ith component of a vector $\bar{b}$ of size 10, with entries equally spaced between 0 and 1/2. We consider sequences of time intervals $[\tau k+1,(k+1)\tau]$, with $\tau = 10$, for $k = 0, 1, 2, \ldots$ in which the weights associated to ξ_i are equal to p_i^k. For every τ iterations, the weight associated to each ξ_i changes randomly, with the only constraint that they sum to 1. $\hat{\theta}_5$ is computed considering a trajectory of 5.10^7 points, and a threshold u corresponding to the $0.995-$quantile of the observable distribution. Using the same trajectory, we computed the empirical distribution of the number of visits in a ball centered at the origin and of radius $r = e^{-u}$ in intervals of time of length $\lfloor 2tr \rfloor$, with $t = 50$. This is indeed t times the Lebesgue measure (which is the invariant measure associated to all the maps f_i) of a ball of radius r centered in the origin. Choosing the same trajectory for the computation of θ and for the statistics of visits is of crucial importance here, because different probability laws imply large variations of θ depending on the trajectory considered. In figure 12, we observe again a perfect agreement between the empirical distribution of the number of visits and the Polyà-Aeppli distribution of parameters $t = 50$ and $\hat{\theta}_5$, which is equal to 0.92 for the trajectory presented in the figure 12. Although different trajectories give variations for θ, this agreement is stable against different trajectories and different values of t and τ, suggesting that a geometric Poisson distribution for the number of visits given by θ is a universal feature, even for non-stationary scenarii.

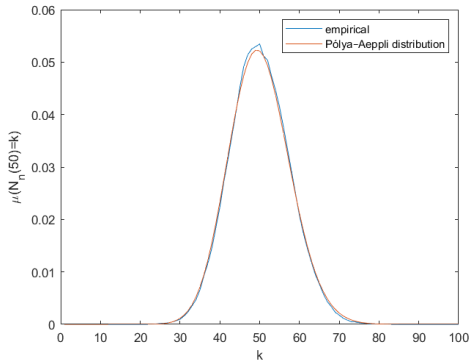


FIGURE 12. Comparison between the empirical distribution $\mu(N_n(50) = k)$ and the Polyà-Aeppli distribution of parameters $t = 50$ and $\hat{\theta}_5 \approx 0.92$, computed with a trajectory of length 5.10^7 of the sequential system described in the text. The procedure used to compute the empirical distribution is described in the text.

9. Acknowledgement

We would like to thank Jorge Freitas, Michele Gianfelice, Nicolai Haydn and Giorgio Mantica for several discussions related to different parts of this work. PY was supported by ERC grant No. 338965-A2C2.

References

[1] M. Abadi, A.C.M. Freitas, J.M. Freitas, Dynamical counterexamples regarding the Extremal Index and the mean of the limiting cluster size distribution, arxiv 1808.02970
[2] L. Arnold, Random Dynamics, Springer Monographs in Mathematics. Springer- Verlag, Berlin, 1998.
[3] V.I. Arnold, A. Avez, Ergodic Problems of Classical Mechanics, The Mathematical Physics Monograph Series, W. A. Benjamin Inc., New York/Amsterdam, 1968.
[4] H. Aytac, J.M. Freitas, S. Vaienti, Laws of rare events for deterministic and random dynamical systems, Trans. Amer Math. Soc., 367, no. 11, 8229-8278, 2015.
[5] J. Beirlant, Y. Goegebeur, J. Segers, J. L. Teugels, D. De Waal (Contributions by), C. Ferro (Contributions by), *Statistics of Extremes: Theory and Applications.*
[6] R.C. Bradley, Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions, Probab. Surveys, Volume 2, (2005), 107-144.

[7] R. Caballero, D. Faranda, G. Messori, A dynamical systems approach to studying midlatitude weather extremes, Geophys. Res. Lett. 44(7) (2017) 3346–3354.
[8] T. Caby, D. Faranda, G. Mantica, S. Vaienti, P. Yiou, Generalized dimensions, large deviations and the distribution of rare events, submitted, arxiv 1812.00036
[9] M. Carvalho, A.C. Moreira Freitas, J. Milhazes Freitas, M. Holland, M. Nicol, *Extremal dichotomy for uniformly hyperbolic systems.* Dyn. Syst. 30(4) (2015)
[10] M.R. Chernick, T. Hsing, W.P. McCormick, Calculating the extremal index for a class of stationary sequences. Adv. in Appl. Probab., 23 (4), 835–850, (1981).
[11] Z Coelho and P Collet: Asymptotic limit law for the close approach of two trajectories in expanding maps of the circle; Prob. Theory Relat. Fields, 99, (1994), 237–250.
[12] J.P. Conze, A. Raugi, Limit theorems for sequential expanding dynamical systems on $[0, 1]$, Ergod. Theory Relat. Fields. Contemp. Math. 430, (2007), 89-121.
[13] D. Faranda, A.C. Moreira Freitas, J. Milhazes Freitas, M. Holland, T. Kuna, V. Lucarini, M. Nicol, M. Todd, S. Vaienti, Extremes and Recurrence in Dynamical Systems, Wiley, New York, 2016.
[14] D. Faranda, J.-M. Freitas, V. Lucarini, G. Turchetti, S. Vaienti, Extreme Value Statistics for Dynamical Systems with Noise, Nonlinearity, 26, 2597-2622, (2013).
[15] D. Faranda, H. Ghoudi, P. Guiraud, S. Vaienti, Extreme Value Theory for synchronization of Coupled Map Lattices, Nonlinearity 31 (2018) 3326–3358.
[16] Faranda, D., Lucarini, V., Manneville, P. and Wouters, J. On using extreme values to detect global stability thresholds in multi-stable systems: the case of transitional plane Couette flow. Chaos Solitons and Fractals, 64. 26 - 35, (2014)
[17] D. Faranda, G. Messori, P. Yiou, Dynamical proxies of North Atlantic predictability and extremes, Sci. rep. 7 (2017) 41278.
[18] D. Faranda, M. C. Alvarez-Castro, G. Messori, D. Rodrigues, P. Yiou. The hammam effect or how a warm ocean enhances large scale atmospheric predictability. Nature communications, 10(1), 1316 (2019).
[19] D. Faranda, Y. Sato, G. Messori, N. R. Moloney,P. Yiou, Minimal dynamical systems model of the northern hemisphere jet stream via embedding of climate data, Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2018-80
[20] D. Faranda, S. Vaienti, Correlation dimension and phase space contraction via extreme value theory, Chaos 28 (2018) 041103.
[21] D. Faranda, S. Vaienti, Extreme value laws for dynamical systems under observational noise, Phys. D, 2014.
[22] M. Ferreira, Heuristic tools for the estimation of the extremal index: a comparison of methods, REVSTAT-Statistical Journal, 16, 115-136, (2018).
[23] A.C.M. Freitas, J.M. Freitas, M.Magalhaes, S. Vaienti, Point processes of non stationary sequences generated by sequential and random fynamical systems, in preparation.
[24] A. C. M. Freitas, J. M. Freitas and M. Todd. The extremal index, hitting time statistics and periodicity. Adv. Math. 231 (5) (2012) 2626–2665.
[25] A.C.M. Freitas, J.M. Freitas, S. Vaienti, Extreme value laws for nonstationary processes generated by sequential and random dynamical systems, Annales de l'Institut Henri Poincaré, 53, 1341–1370, (2017).
[26] S. Gallo, N. Haydn, S. Vaienti, in preparation.
[27] C. Gonzalez-Tokman, Multiplicative ergodic theorems for transfer operators: towards the identification and analysis of coherent structures in non-autonomous dynamical systems, Contributions of Mexican mathematicians abroad in pure and applied mathematics. Guanajuato, Mexico: American Mathematical Society. Contemporary Mathematics, 709, 31-52 (2018).
[28] P. Gora, A. Boyarsky, Laws of Chaos, Birkhäuser, (1997)
[29] N. Haydn, M. Nicol, A. Torok, S. Vaienti, Almost sure invariance principle for sequential and non-stationary dynamical systems, Trans. Amer Math. Soc., 36, (2017), Pages 5293-5316.
[30] N. Haydn, S. Vaienti,The compound Poisson distribution and return times in dynamical systems. Probab. Theory Related Fields, 144 (3-4), 517–542, (2009).
[31] N. Haydn, S. Vaienti, Limiting entry times distribution for arbitrary sets, preprint
[32] G. Keller, Rare events, exponential hitting times and extremal indices via spectral perturbation, Dyn. Syst. 27 (2012) 11–27.
[33] G. Keller and C. Liverani, Rare events, escape rates and quasistationarity: some exact formulae, J. Stat. Phys. 135 (2009), 519-534.
[34] M.R. Leadbetter, M.R., On extreme values in stationary sequences, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 28, 289–303,(1973/74).
[35] Lucarini, V., Blender, R., Herbert, C., Ragone, F., Pascale, S. and Wouters, J. (2014) Mathematical and physical ideas for climate science. Reviews of Geophysics, 52 (4). pp. 809-859, (2014)
[36] Lucarini, V., Faranda, D., Wouters, J. and Kuna, T. Towards a general theory of extremes for observables of chaotic dynamical systems. Journal of Statistical Physics, 154 (3). pp. 723-750, (2014)

[37] N. Moloney, D. Faranda, Y. Sato, *An overview of the Extremal Index*, Submitted to Chaos, 2018
[38] G.L. O'Brien, Extreme values for stationary and Markov sequences, Annals of Probability, 15, 281–291, (1987).
[39] M. Süveges, Likelihood estimation of the extremal index, Extremes 10 (2017) 41–55.