HAL
open science

Extreme value distributions of observation recurrences

Th Caby, Davide Faranda, S. Vaienti, Pascal Yiou

To cite this version:

Th Caby, Davide Faranda, S. Vaienti, Pascal Yiou. Extreme value distributions of observation recurrences. Nonlinearity, 2021, 34 (1), pp.118-163. 10.1088/1361-6544/abaff1 . hal-02904585

HAL Id: hal-02904585

https://hal.science/hal-02904585v1

Submitted on 22 Jul 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.

EXTREME VALUE DISTRIBUTIONS OF OBSERVATION RECURRENCES

Abstract. We study analytically and numerically the extreme value distribution of observables defined along the temporal evolution of a dynamical system. The convergence to the Gumbel law of observable recurrences gives information on the fractal structure of the image of the invariant measure by the observable. We provide illustrations on idealized and physical systems.

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

Contents

1. Introduction 2
1.1. A general overview 2
1.2. Salient results of the paper 2
2. The formal approach 4
2.1. Basics on Extreme Value Theory 4
2.2. The perturbative spectral approach 6
3. The extremal index 10
4. Phenomenology of the image measure 12
4.1. The baker map 15
4.2. The product of two Cantor sets 18
4.3. The Lorenz system 20
4.4. Conclusions 20
5. Statistics of visits for the observable 20
6. Randomly perturbed systems 24
6.1. Perturbing the map 25
6.2. Perturbing the observable 28
7. Open systems 31
7.1. EVT on fractals I 33
8. Hitting time statistics in the neighborhood of sets 35
8.1. Smooth sets 35
8.2. EVT of fractals II 36
9. Large deviations 36
9.1. Examples 38
10. Applications 39
10.1. From scalar to vector-valued observables 40
10.2. Vector-valued observables 41

Date: July 22, 2020.

[1]Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France 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`.

11. Conclusions 44
12. Acknowledgements 44
References 44

1. Introduction

1.1. **A general overview.** Extreme value theory (EVT) has been used in dynamical systems in the last years to quantify the probability of visiting a small set in the phase space, which constitutes a *rare event*. With this approach, the asymptotic statistics of hitting times and of the number of visits [15] in small sets can be described. Methods based on EVT and more generally on the recurrence properties of chaotic systems have found applications in climate science [27, 14, 24, 23]. Quantifying the recurrence properties of weather patterns via dynamical indicators has proven useful to solve a number of issues in climate and atmospheric sciences. In [27] the recurrence properties of the North-Atlantic sea-level pressure fields have been studied. A number of instantaneous metrics that track rarity, predictability and persistence of atmospheric jet states and circulation patterns have been derived starting from quantities defined in the framework of EVT for dynamical systems, e.g. the local dimensions and the extremal index. In [12, 59] the same metrics have been used to classify and evaluate the dynamical consistence of state-of-the art climate models in representing the atmospheric dynamics. The impact of climate change of the atmospheric dynamical features was identified through shifts of the local dimensions between 1850 and 2100, in various datasets (observations, ensembles of scenario climate model simulation) [23]. A critical discussion of the methods used in these studies is available in [16, 15]. To justify them, one needs to work with data sampled from the original high dimensional system, while experimentalists often have access to a lower dimensional representation of the underlying attractor through measurements. A first approach to recover information on the underlying system from observations is to use embedding techniques, which is allowed by Takens' theorem [66]. Thanks to the theory of extreme value distribution applied to observables developed in this paper, we are able to propose an alternative technique and we will propose an application to atmospheric sciences. On a more general ground, the aim of our work is to study the statistics of recurrences of smooth observables in chaotic dynamical systems. We will state some general results that could be applied in a wide range of situations. Our basic inspirations were the works of [60, 41, 11], where the authors developed different theoretical ideas and tools to derive, among others, recurrence rates for observations and compute them for various dynamical systems.

1.2. **Salient results of the paper.**

(1) **Section 2** puts the basis of EVT for observations. We look at the distribution of the maximum of a sequence of random variables obtained by evaluating a vector valued observable along the orbit of a dynamical system and approaching a limiting value of the observable itself (the target set). We obtain rigorously a limit distribution of Gumbel type by using a perturbation theory applied to dynamical systems of hyperbolic type.

(2) An extremal index (EI) modulates the limit distribution, by adding a factor to the Gumbel law. This EI is related to the frequency of the occurrences (visits to the target sets), which is interpreted as a clustering of the orbits. The EI becomes smaller than one when the target set exhibits periodic patterns. In **Section 3** we first provide general formulas for the EI for a large class of one-dimensional

expanding maps and non-invertible observables. Then we show that the observable could generate several coexisting clusters and we explicitly compute the EI in a few cases.

(3) The numerical approach to the limit distribution via the Generalized Extreme Value (GEV) distribution, allows us to estimate the local properties of the image measure. **Section 4** is partially devoted to a brief exposition of the Hunt and Kaloshin theory of prevalent spaces in relation with the point-wise dimension of image measures. We therefore study in details two examples, the baker map and the product of two Cantor sets, and show that a few quite simple observables are not prevalent. This means that the dimension of the image measure is not integer (that of the ambient space), but smaller or larger and coinciding with that of the underlying attractor for the dynamics. The theoretical results were supported by numerical computations using the EVT techniques. This, combined with a suitable choice of the observable, is therefore a very efficient tool to describe the fine geometric structure of the limit sets of the dynamics.

(4) We go beyond the Gumbel law in **Section 5**, by studying the statistics of the number of visits of the observable in the neighborhood of a value of interest. This is the point process associated to the distribution of the first hitting time, and we show that it is either purely Poisson distributed or it deviates from the usual Polyà-Aeppli distribution, which characterizes the point process when the rare set is around a periodic point. A particular example is studied in detail and a limit compound Poisson distribution is exhibited *via* its generating function and a recursive formula for the probability mass function. Application to climate data shows a compound Poisson distribution, despite the relative modest length of the time series and the unavoidable approximations in their detection.

(5) In **Section 6** we consider what happens when the dynamical system and the observable are randomly perturbed. We show with analytical and numerical arguments, that if the perturbation of the map produces a smooth stationary measure or the observable changes randomly but staying prevalent, then the dimension of the image measure becomes integer. Stability behaviors are also discussed.

(6) We then move to open systems in **Section 7** by considering in the phase space the presence of absorbing regions (*holes*), where the orbits could be trapped and disappear forever. Nevertheless and under general conditions, a fractal repeller survives and it is possible to study the recurrence properties of observables defined in a neighborhood of such a repeller.

(7) **Section 8** gives a geometrical interpretation to our results and shows that our approach can be used to compute the hitting time statistics in the neighborhood of hypersurfaces embedded in the phase space of the system. Applications to fractal sets are also given.

(8) The experimental and numerical computation of the local dimension by EVT shows a discrete variability of such dimensions, even if they are constant almost everywhere (at least when they exist almost surely with respect to ergodic measures). The presence of those (large) deviations, is revealed by the non-linearity of the so-called spectrum of generalized dimensions (the *free-energy function* of the process), which are accessible to analytic and numerical computations. In **Section 9** we treat the large deviations of the dimensions of the image measure and discuss how those deviations are influenced by the choice of observable.

(9) We quoted in section 1.1 the embedding technique as a tool to reconstruct the attractor by considering the iterates of a unidimensional projection of the dynamics.

When considering enough delay coordinates, the dimension of the attractor becomes accessible. In **Section 10** we propose an alternative approach that allows us to have access to the dimension of the attractor by using directly observational data. In particular, this is possible when the dimensionality of the observations is larger than the information dimension of the underlying system. To achieve this either we dispose of a vector-valued observable, or we could use a scalar observable to construct several images just by composing with the dynamics. In some sense, the delay coordinate observable used in embedding techniques is a particular case of the smooth observables that we consider.

2. The formal approach

We now introduce the basic concepts on Extreme Value Theory and apply them to a sequence of observations. The stationary random process that arises is then studied with a perturbative spectral technique, which allows us to prove directly the convergence to the Gumbel law.

2.1. Basics on Extreme Value Theory.

Let us consider a dynamical system (X, T, μ), where T acts on the measurable space X and preserves the invariant probability measure μ. In the following we will consider X as a compact subset of $\mathbb{R}^n$, $(n \geq 1)$ and we put the Borel σ-algebra on it. We take $f : X \to \mathbb{R}^l$ a measurable function, called the *observable*; it will play a fundamental role in this paper, and additional assumptions on its regularity will be progressively added.

Let us now construct the new measurable function

$$\phi(x) = -\log(\mathrm{dist}(f(x) - f(z))), \tag{1}$$

where z is given in X. This function has values in $\mathbb{R} \cup \{+\infty\}$ and achieves a global maximum at the pre-images of $f(z)$, where it is precisely infinite. With dist we take a distance defining the metric on $\mathbb{R}^l$. Consider the maximum of the process $\{\phi \circ T^k\}_{k>0}$, namely

$$M_n(x) = \max\{\phi(x), \ldots, \phi(T^{n-1}(x)\} \tag{2}$$

and the distribution

$$\mu(M_n \leq u_n) = \mu(\phi \leq u_n, \ldots, \phi \circ T^{n-1} \leq u_n). \tag{3}$$

Définition 1. *We say that we have an extreme value law for M_n if there is a non-degenerate distribution function $H : \mathbb{R}^+ \to [0, 1]$ and for every $\tau > 0$ there exists a sequence of levels $u_n = u_n(\tau), n \in \mathbb{N}$, such that*

$$n\mu(\phi > u_n) \to \tau, \ \text{as } n \to \infty. \tag{4}$$

and for which the following holds:

$$\mu(M_n \leq u_n) \to H(\tau), \ \text{as } n \to \infty.$$

Remark 1. *We name Eq. (4) the* **Assumption F***: it allows us to avoid a degenerate limit for the distribution of M_n. We will see later on that the perturbative spectral technique prescribes* **Assumption F** *in a very natural way.*

Notice that Eq. (4) is equivalent to

$$n\mu(x \in X, f(x) \in B(f(z), e^{-u_n})) \to \tau, \tag{5}$$

where $B(a, r)$ denotes the ball of radius r centered at the point a in the metric given by the chosen distance.

By introducing the *image measure* $f_*\mu$ defined as

$$f_*\mu(A) = \mu(f^{-1}A), \tag{6}$$

where A is any Borel set in $\mathbb{R}^l$, we can equivalently rewrite Eq. (5) as

$$n\ f_*\mu(B_{n,z}) \to \tau, \tag{7}$$

where we set

$$B_{n,z} := B(f(z), e^{-u_n}) \text{ and } C_{n,z} := B(f(z), e^{-u_n})^c. \tag{8}$$

The superscript A^c is the complementary set of A in X.

Remark 2. *The presence of the observable imposes some natural conditions on the combined choice of f and T if we want to satisfy Eq. (5). For instance if f is locally constant in the neighborhood of the target point z and μ is not atomic in z, we see immediately that Eq. (5) cannot hold for large n. A less trivial example is given by the direct product map T on the unit square defined by*

$$T(x,y) = \begin{cases} 2x, \ x \in [0,1/2]; 1-2x, \ x \in [1/2,1], \\ ay, \ 0 < a < 1, y \in [0,1]. \end{cases}$$

This map preserves the product of the Lebesgue measure on the x-axis times the Dirac mass at 0 on the y-axis. If we now take the observable $f(x,y) = y$ and the target point in $(0,0)$, we see that the set $T^{-1}[B(f(0,0), e^{-u_n})]$ is a strip of length 1 and of width e^{-u_n} on the square and the measure of this strip will be 1 for any n.

Notice that if the observable f is not locally constant in the neighborhood of the target point and the image measure is not atomic we can always choose a sequence u_n verifying for each n: $n\ f_*\mu(B_{n,z}) = \tau$. We will see in the next section, in particular the scaling (26), that u_n is an affine function of the variable $y := -\log\tau$ which can be written as:

$$u_n = \frac{y}{a_n} + b_n, \ a_n > 0. \tag{9}$$

When the sequence $\mu(M_n \leq u_n) = \mu(a_n(M_n - b_n) \leq y)$ converges to a non-degenerate distribution function $G(y)$, in the point of continuity of the latter, then we have an extreme value law. The starting point of EVT, related to the affine choice for the sequence u_n, is that such a $G(y)$ could be only of three types, called Gumbel, Fréchet and Weibull (see [47] for a general account of the theory). One of the main goal of this paper is to show that for the particular observable Eq. (1), we will get the Gumbel law, see **Proposition 1**. The scaling (26) shows that the parameters a_n and b_n are expressed in terms of the local dimension of the image measure, in Eq. (24). It would be therefore useful to have access to those parameters. In this regard, we begin to notice that the distribution function of the form $\mu(M_n \leq y)$ is modeled for n sufficiently large, by the so-called *generalized extreme value (GEV)* distribution [58], which is a function depending upon three parameters $\xi \in \mathbb{R}$, (the tail index), $\kappa \in \mathbb{R}$, (the location parameter) and $\hat{\sigma} > 0$, (the scale parameter):

$$\mathrm{GEV}(y; \kappa, \hat{\sigma}, \xi) = \exp\left\{ -\left[1 + \xi\left(\frac{y-\kappa}{\hat{\sigma}}\right)\right]^{-1/\xi} \right\}.$$

The location parameter κ and the scale parameter $\hat{\sigma}$ are scaling constants in place of b_n and a_n. The idea is now to use a block-maxima approach (see Section 4.1) and fit our unnormalised data to a GEV distribution; for that it will be necessary to find a linkage among a_n, b_n, κ and $\hat{\sigma}$. For observables ϕ producing the Gumbel law, it has been shown in [25], that for n large we have

$$a_n \sim \frac{1}{\hat{\sigma}};\ b_n \sim \kappa; \tag{10}$$

moreover the shape parameter ξ tends to zero. The systematic use of this approach from Section 4, will allow us to compute the local dimensions of the image measures and give therefore a numerical and experimental support to the theoretical results: this is another relevant aspect of our work. We finish this section by giving another definition.

Définition 2. *We say that the process $\{\phi \circ T^k\}_{k>0}$, for the observable (1), has an Extremal Index (EI) $0 \leq \theta \leq 1$, if we have a Gumbel distribution as*

$$\mu(M_n \leq u_n) \to e^{-\theta\tau},\ n \to \infty,$$

with the sequence u_n verifying **Assumption F.**

As we anticipated above, Gumbel's law is the limiting distribution for the maxima. The next sections will be devoted to the analytic computation of the EI. Besides rigorous estimates, we will also proceed to numerical computations. The EI is less than one when clusters of successive recurrences happen, which is the case, for instance, when the target point z is periodic. In our paper [15] we showed that the usual algorithms to compute the EI have strong limitations when clusters of higher order are present, and a new technique was proposed which consists in computing the first five q_k terms in the expansion of θ, see formula (20). We used this technique for the numerical estimates of the EI all along the paper.

2.2. **The perturbative spectral approach.** In order to apply the aforementioned perturbative spectral technique, we suppose that the system (X, T, μ) is REPFO (*Rare events Perron-Frobenius operators*) according to the terminology introduced by G. Keller [44, 45]. The definition of a REPFO dynamical system is quite technical even if its assumptions are verified in several situations when the system is uniformly hyperbolic or expanding. We must detail those assumptions because they impose new constraints on the choice of the observable f. The basic object is the transfer (Perron-Fröbenius) operator P associated to the map T. This operator acts on a suitable Banach space $(\mathcal{B}, ||\cdot||)$, equipped with a second (weak) norm $|\cdot|$ for which the closed unit ball of $(\mathcal{B}, ||\cdot||)$ is $|\cdot|$-compact.

The Banach space is a space of functions or of distributions. We will mostly treat non-invertible maps and in this case $\mathcal{B}$ will be the space of bounded variation (BV) functions and the weak norm will be the L^1 norm with respect to the Lebesgue measure. We will also consider invertible maps and, in this case, $\mathcal{B}$ is a space of distribution and we defer to [21] for a nice presentation of those spaces or to [4] for an easy description in the context of EVT. To make the exposition simpler we will suppose that $\mathcal{B}$ is the space of BV functions and the weak norm is the space of integrable functions with respect to the Lebesgue measure Leb.[5]

We will see below that the operator P is slightly perturbed to get a sequence of operators $\tilde{P}_n$ which converge to P in a sense that we are going to precise: for the moment we retain that $\tilde{P}_n$ is defined as $\tilde{P}_n(g) = P(\mathbf{1}_{W_n} g)$, g and $\mathbf{1}_{W_n} g \in \mathcal{B}$, where the Lebesgue measure of the measurable set W_n goes to one when $n \to \infty$, (see below for the explicit

[5]Sometimes, especially in the integral, we will write dx instead of dLeb.

construction of such a W_n; its complementary set W_n^c should be interpreted as a *hole* with vanishing measure, not necessarily with vanishing diameter). The following four items define precisely what a REPFO system is: they are taken from [44] and slightly modified to our situations:

- **A1** The unperturbed operator P is quasi-compact: this means, in particular, that 1 is a simple isolated eigenvalue and there are no other eigenvalues on the unit circle. This implies the existence of a unique mixing invariant measure μ for T which is absolutely continuous with respect to Leb with the density h.
- **A2** There are constants $\beta, D > 0$, such that $\forall n$ sufficiently large, $\forall g \in \mathcal{B}$ and $\forall k \in \mathbb{N}$ we have (*Lasota-Yorke inequalities*):

$$|\tilde{P}_n^k g| \leq D|g|, \tag{11}$$

$$||\tilde{P}_n^k g|| \leq D\beta^k ||g|| + D|g|. \tag{12}$$

- **A3** We can bound the weak norm of $(P - \tilde{P}_n)g$, with $g \in \mathcal{B}$, in terms of the norm of g as:

$$|(P - \tilde{P}_n)g| \leq \chi_n ||g||,$$

  where χ_n is a monotone upper semi-continuous sequence converging to zero; this is called the *triple norm* estimate.
- **A4** If we put for $g \in \mathcal{B}$

$$\eta_n := \sup_{||g|| \leq 1} |\int P(g\ \mathbf{1}_{W_n^c}) d\text{Leb}|, \tag{13}$$

  we must show that

$$\lim_{n \to \infty} \eta_n = 0, \tag{14}$$

$$\eta_n ||P(\mathbf{1}_{W_n^c} h)|| \leq \text{const}\ \mu(W_n^c). \tag{15}$$

We now associate to the space of BV functions a uniformly expanding endomorphism T of the unit interval and preserving the absolutely continuous invariant mixing measure μ with density h. The transfer operator has now a simple definition; for $v \in L^1(\text{Leb})$ and $w \in L^\infty(\text{Leb})$, we have

$$\begin{cases} \int P(v) w dx = \int v w \circ T dx, \\ P(h) = h. \end{cases}$$

Using this duality relation, the distribution in Eq. (2) reads

$$\mu(M_n \leq u_n) = \int (\mathbf{1}_{C_{n,z}} \circ f)(x) \cdots (\mathbf{1}_{C_{n,z}} \circ f)(T^{n-1}x) h(x) dx = \int (\tilde{P}_n^n h)(x) dx. \tag{16}$$

where

$$\tilde{P}_n g := P(\mathbf{1}_{C_{n,z}} \circ f\ g),\ g \in \text{BV}. \tag{17}$$

In the case of hyperbolic diffeomorphisms, we have a slightly different formula, since the operator acts on measures, not on functions. When $n \to \infty$, the preceding assumptions allow us to express the largest eigenvalue of $\tilde{P}_n$, say χ_n, in terms of the largest eigenvalue of the unperturbed operator, which is 1, as: $\chi_n = 1 - (\theta \Delta_n + o(\Delta_n))$, where $\Delta_n =$

$\mu(f^{-1}(B_{n,z}))$. The quantity θ is formally defined as $\theta = 1 - \sum_{k=0}^{\infty} q_k$, and the q_k are given by the following limits, *when they exist*:

$$q_k = \lim_{n\to\infty} q_{k,n}, \text{ where } q_{k,n} := \frac{\int (P - \tilde{P}_n)\tilde{P}_n^k (P - \tilde{P}_n)(h)dx}{\Delta_n}. \tag{18}$$

The operator $\tilde{P}_n$ now decomposes as the sum of a projection along the one dimensional eigenspace associated to the eigenvalue χ_n and an operator with a spectral radius exponentially decreasing to zero and which can be neglected in the limit of large n.[6] Remembering this, writing $\tilde{P}_n h_n = \chi_n h_n$, with h_n converging to h in the L^1(Leb) norm and replacing into the right hand side of Eq. (16) and after a few manipulations we get, by neglecting higher order terms:

$$\mu(M_n \leq u_n) \approx e^{-\theta n \Delta_n}.$$

The product $n\Delta_n = nf_*\mu(B_{n,z})$ is now controlled by **Assumption F**, which allows us to get a limiting distribution. The justification of the previous statements is a direct application of Keller's theory [44] which gives the following

Proposition 1. *Let us suppose that (X, T, μ) is a REPFO system. Suppose moreover that the* **Assumption F** *holds. Then*

$$\mu(M_n \leq u_n) \underset{n\to\infty}{\to} e^{-\theta\tau}, \tag{19}$$

where the extremal index θ is defined as

$$\theta = 1 - \sum_{k=0}^{\infty} q_k, \tag{20}$$

where

$$q_k := \lim_{n\to\infty} q_{k,n} \tag{21}$$

The quantities q_k are given by the limit (18) where the quantities $q_{k,n}$ are equivalently expressed as:

$$q_{k,n} = \frac{1}{f_*\mu(B_{n,z})}\mu(f^{-1}B_{n,z} \cap T^{-1}(f^{-1}B_{n,z})^c \cdots \cap T^{-k}(f^{-1}B_{n,z})^c \cap T^{-(k+1)}f^{-1}B_{n,z}). \tag{22}$$

Comments. As we said above the proof of this proposition follows immediately from Keller's theory, see also our previous works [22, 26, 16, 15, 4]. Three issues deserve to be discussed. The first two deal with the possibility to give examples which fit with the REPFO assumptions. Whenever f is the identity function, the aforementioned references give a large class of examples. The problem now is the presence of the observable f which could affect the hypothesis **A2**-**A4**. The third issue concerns the computation of the extremal index.

(1) A particular attention must be drawn to the Lasota-Yorke inequalities **A2** which has to do with the characteristic function of sets of the type $f^{-1}(C_{n,z})$, which could have a geometric shape quite different from balls. We should guarantee that the Banach norm of these sets is computable and allows to get the desired inequalities. This will be the case for all the systems with associated observables which we will be treated analytically in this paper. We will in fact consider the observables f as continuous and local C^1 functions and in this case the Lasota-Yorke inequalities for the perturbed operators can be proved using the arguments

[6] This is precisely what quasi-compactness means.

in [7], Lemma 2.6, or [48], Lemma 7.4. For the baker map (see section 4), we defer to the paper [4], section 3.1.

(2) The second isuue concerns the Assumptions **A3** and **A4**. They will follow if we could prove that the $L^1(\text{Leb})$ norm of $(\tilde{P}_n - P)g$, with $g \in BV$, is bounded by the BV norm of g and the image-Lebesgue ($f_*\text{Leb}$) measure of $B_{n,z}$.
We have

$$\begin{aligned} \int |(\tilde{P}_n - P)g|h(x)dx &= \int |P(\mathbf{1}_{B_{n,z}}(f(x))g(x))|h(x)dx \\ &\leq ||h||_\infty \int P(\mathbf{1}_{B_{n,z}}(f(x))|g(x)|)dx \\ &\leq ||h||_\infty \int \mathbf{1}_{B_{n,z}}(f(x))|g(x)|dx \\ &\leq ||h||_\infty ||g||_{BV} f_*\text{Leb}(B_{n,z}), \end{aligned} \tag{23}$$

since both h and g are in BV and the infinity norm is bounded by the BV norm $||\cdot||_{BV}$. The perturbative theorem requires finally that $f_*\text{Leb}(B_{n,z}) \leq$ constant $f_*\mu(B_{n,z})$, which is surely true if the density h is bounded from below: we will tacitly assume it if necessary.

(3) Finally we should check the existence of the limits (21) to give the fundamental expression of Eq. (20) for the extremal index. Note that the Poincaré recurrence theorem implies that $\sum_{k=0}^{\infty} q_{k,n} = 1$; therefore whenever q_k exists, the extremal index is at most 1. The quantities $q_{k,n}$ have a simple geometrical interpretation: they give the conditional measure of the points that are at the beginning in the set $f^{-1}B_{n,z}$, are iterated outside it for the next k times, and finally return to it at the $k+1$ iteration. As we will argue below, in particular in section 4, this structure of the $q_{k,n}$ allows us to compute them explicitly in several situations, or guess their possible behavior.

We now define the local dimensions of the image measure. We put

$$\underline{d}^f_\mu(x) := \liminf_{r\to 0} \frac{\log f_*\mu(B(f(x),r))}{\log r}, \tag{24}$$

$$\overline{d}^f_\mu(x) := \limsup_{r\to 0} \frac{\log f_*\mu(B(f(x),r))}{\log r}. \tag{25}$$

Whenever $\underline{d}^f_\mu(x) = \overline{d}^f_\mu(x) = d^f_\mu(x)$, we will say that the image measure $f_*\mu$ is *exact dimensional.*

Notations: Sometimes instead of $d^f_\mu(x)$ we will use the notation $d^f_\mu(f_0)$, meaning that the pointwise dimension is computed in the point $f_0 = f(z)$ without specifying the value of z. When the measure is exact dimensional we will simply write d^f_μ as the almost sure value. We will also use the symbol d^f_μ to denote what we presume to be the almost sure value of the image measure in a few numerical computations for which the measure μ could only be reconstructed numerically. This especially concerns the last chapter.

We now suppose that $d^f_\mu(x)$ exists and express Eq. (7) as the scaling:

$$f_*\mu(B_{n,z}) \sim e^{-u_n\ d^f_\mu(z)} \sim \tau/n. \tag{26}$$

The result on the Gumbel law given by Eq. (19) could be reformulated by: if f is an observable on the space X, and $f(z)$ is the value at a given point z, then *the probability that the distance between $f(T^n x)$ and $f(z)$ after n iterations is less than $(\frac{\tau}{n})^{d_\mu^f}$ for the first time, is approximately $e^{-\theta\tau}$.*

Let us come back to the results in [60]. The equality between the recurrence rate for the observable defined as

$$R_f(z) = \lim_{r\to 0} \frac{\log \inf\{k \in \mathbb{N}^* : f(T^k z) \in B(f(z), r)\}}{\log r}, \tag{27}$$

and $d_\mu^f(z)$ is proven for a class of systems with superpolynomial decay of correlations and such that the image measure $f_*\mu$ is exact dimensional. In the spectral theory, the property of superpolynomial decay of correlation is strengthened by the presence of the spectral gap for the transfer operator, which implies exponential decay of correlations. We point out that our approach is slightly different from the one of [60], in the sense that we get a recurrence rate for hitting times instead of return times and its distribution for shrinking target sets. For this reason we will not further elaborate on the connections with the quantity in Eq. (27).

3. The extremal index

The extremal index θ is usually considered as a measure of *clustering*, whenever several and repeated occurrences take place in the ball $B_{n,z}$. For the usual observable $-\log \mathrm{dist}(x, z)$, this happens around periodic points for the map T. When it comes to recurrence of observables, some clustering can also occur when z is a periodic point. We now show that the extremal index for observables reveals new interesting features.

We start with a simple example.

Take a real observable f defined on the unit interval such that in any point where it is defined, the derivative is bounded below away from zero and above from infinity.

Let us first consider the case on an *invertible* f and take T as a uniformly expanding map of the interval which has $z = f^{-1}(f_0)$, where $f_0 = f(z)$, as a fixed point and is continuous in such a point together with the density of the absolutely continuous invariant measure h. Then we have

$$q_{0,n} = \frac{1}{f_*\mu(B_{n,z})} \mu(f^{-1}B_{n,z} \cap T^{-1}f^{-1}B_{n,z}). \tag{28}$$

At this point we can repeat the standard argument (see, for example, section 4.2 of [36]) to get immediately that

$$\theta = 1 - \frac{1}{|T'(z)|}.$$

We now take a *non-invertible* f. In particular we suppose f has two branches: f_1, f_2. Suppose the ball $B_{n,z}$ is again centered at a point $f_0 = f(z)$ and the point $z_1 := f_1^{-1}(z)$ is the inverse point of f such that

$$Tz_1 = z_1.$$

Moreover suppose that the other pre-image $z := f_2^{-1}(z)$ is not periodic for T.

In Eq. (28) for the $q_{0,n}$ above, only the pre-images by T of the set $f_1^{-1}(B_{n,z})$ matter in the computation of the EI, but we have to take into account the relative ratio of the measure of $f_1^{-1}(B_{n,z}), f_2^{-1}(B_{n,z})$ in the denominator. These measures are obtained by

pulling back the Lebesgue measure of $B_{n,z}$ with the reciprocal images of f, which amounts to multiply the length of $B_{n,z}$ with the reciprocal of the derivative of f in the pre-images of f_0, and multiply what we get by the density h in such pre-images. In conclusion we have

$$\theta = 1 - \frac{1}{|T'(z_1)|} \frac{1}{1 + \frac{h(z)|f'(z_1)|}{h(z_1)|f'(z)|}}. \tag{29}$$

The preceding argument can be generalized to give an exact formula for the q_k. As we will see, the existence of several pre-images of the ball $B_{n,z}$ could generate multiple clusters coexisting with different degrees of periodicity.

Proposition 2. *Let us suppose that T is a uniformly expanding map as above and the observable f is differentiable with a derivative bounded away from zero and infinity. Fix z in the unit interval M and put $f_0 = f(z)$; suppose also that f is a finite-to-one map. Consider the set of the pre-images w of f_0, one of them being z, and suppose that they do not belong to the countable union of the pre-images of the boundary points of the domains of local injectivity of T and that the invariant density h is continuous in such points. Consider the set*

$$A_k = \{w \in M : f(w) = f_0, f(Tw) \neq f_0, \dots, f(T^k w) \neq f_0, f(T^{k+1}w) = f_0\}.$$

When $A_k = \emptyset$, then $q_k = 0$. Conversely, whenever A_k is finite and non-empty, we have

$$q_k = \sum_{w \in A_k} \frac{1}{|T^{(k+1)}(w)'|} \frac{1}{1 + \frac{|f'(w)|}{h(w)} \sum_{y \in B_k^\omega} \frac{h(y)}{|f'(y)|}}, \tag{30}$$

where $B_k^\omega = \{y \in M : f(y) = f_0\} \setminus \{w\}$.

The extremal index is obtained by

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

We point out that having fixed the center z of the ball $B_{n,z}$ and having f a finite number of pre-images, there are only finitely many points in A_k and consequently finitely many terms in the sum $\sum_{k=0}^{\infty} q_k$. Moreover we could relax the global assumption on f by asking that f be C^1 in z and the pre-images of z as the next example will require.

Let us give two examples. In the first consider the map $T(x) = 3x$-mod 1. Then take a point $a > \frac{1}{2}$ which is not periodic for T (these points yield a full Lebesgue measure), and consider a piece-wise continuous straight line with two branches (f_1, f_2), f_1 passing through the points $(0,0)$ and $(a,1)$, and f_2 through the points $(a,1)$ and $(1,0)$. The equations are

$$\begin{cases} f_1(x) = x/a, \\ f_2(x) = \frac{1}{a-1}\, x - \frac{1}{a-1}. \end{cases}$$

We choose a point z_1 that is a fixed point of T and z_2 that have the same image by f but is not periodic by T. We take

$$\begin{cases} z_1 = \frac{1}{2}, \\ z_2 = f_2^{-1}(f(z_1)) = (a-1)/2a + 1. \end{cases}$$

We can choose a so that z_2 is irrational. In this case, z_2 is not periodic for T and the trajectory starting from z_2 will not pass through z_1 which is rational. Therefore, we see

easily from formula (30) that $q_k = 0$ for $k > 0$. We have $h(z_1) = h(z_2) = 1$ and

$$\begin{cases} |f'(z_1)| = 1/a, \\ |f'(z_2)| = \frac{1}{|a-1|}. \end{cases}$$

Therefore

$$\theta = 1 - q_0 = 1 - \frac{1}{3\left(1 + \frac{|a-1|}{a}\right)}.$$

We checked this formula numerically for various values of a. For example, taking $a = 2/\pi$, we find a numerical value of 0.788 against a theoretical value of 0.7878. We used the estimate $\hat{\theta}_5$ introduced in [15], which consists in estimating the q_k terms up to the order 5 and subtracting them from 1.

We notice that in Eq. (30), the extremal index depends explicitly on the density of the invariant measure, which was constant in the example above.

We give another example where $h(z_1)$ and $h(z_2)$ are different. We take the Hemmer map defined in $[-1, 1]$ by $T(x) = 1-2\sqrt{|x|}$. Its density is $h(x) = \frac{1}{2}(1-x)$, [40]. The point $z_1 = 3 - 2\sqrt{2}$ is a fixed point of the map. We choose the point $z_2 = -1/2$, which is not periodic and we take f piecewise linear with different slopes: $f(x) = x$ for $-1 \leq x \leq 0$ and $f(x) = -2x + 11/2 - 4\sqrt{2}$ otherwise, so that $f(z_1) = f(z_2) = -1/2$. Eq. (3) gives

$$\theta = 1 - q_0 = \frac{\sqrt{3 - 2\sqrt{2}}}{1 + \frac{3}{4(\sqrt{2}-1)}} \approx 0.9104.$$

Our numerical computations confirm this result to the fourth digit with the estimate $\hat{\theta}_5$.

We have given a quite general formula for the one dimensional case, and it is apparent from it that the clustering structure can be quite complicated if the observable and the dynamics have some kind of compatibility. For this reason, giving a general formula for the extremal index in higher dimensional systems is out of the scope of this paper. We however believe that for large class of observables, no clustering is detected and the extremal index should be equal to 1. This is confirmed by several numerical simulations that will be described in the next chapter.

4. Phenomenology of the image measure

We are now interested in estimating the quantity d_μ^f that appears in the distribution of maxima. This question has been partially answered by Rousseau and Saussol in [60], in the case of smooth observables and measures μ that are absolutely continuous with respect to Lebesgue. In particular, Theorem 9 in [60] states that $d_\mu^f(z)$ exists μ almost everywhere, is integer valued and is equal to the rank of $Df(z)$ almost everywhere. For example, if f has values in $\mathbb{R}$ and $\mu(\partial_x f(x) = 0) = 0$, $d_\mu^f(z)$ is equal to 1 for μ-almost any z. It implies also that if f is constant on some regions of the phase space of positive measure, $d_\mu^f(z)$ will be 0 in that region. A first observation is that this result does not hold at some special points of the attractor. We now give an example where d_μ^f is not an integer.

Consider the map $Tx = 2x \mod 1$ defined on the circle, $z = 0$ and the observable $f(x) = x^a$, with $a > 0$. Then we have:

$$\mu(-\log|f(x)-f(0)| > u_n) = \mu(-\log|x^a| > u_n) = \mu(-\log|x| > \frac{u_n}{a}) = \mu(B(0, e^{-\frac{u_n}{a}})) = 2e^{-\frac{u_n}{a}}.$$

Therefore,

$$d_\mu^f(0) = \lim_{n\to\infty} \frac{\log 2e^{-\frac{u_n}{a}}}{\log e^{-u_n}} = 1/a.$$

Depending on the value of a, this quantity can be non integer and either smaller or larger than 1.

In many physical applications, the measure is not smooth, but has a (multi)fractal structure. This happens for chaotic dynamics in neuroscience and climate science [65, 16]. Being able to compute the value of d_μ^f in such situations is of crucial importance to describe the statistics of recurrence of the observable. The simplest case to consider is when the observable f is a diffeomorphism from $\mathbb{R}^k$ to $\mathbb{R}^k$ (where k is the dimension of the ambient space): the image of the invariant set by f is then a deformation of the original attractor which preserves its local structure. We therefore expect that $d_\mu^f(z) = \gamma_\mu(z)$, the pointwise dimension at the point z. Let us now remind the definition of these local dimensions, since they will be used later on. Consider the limits

$$\gamma_\mu^-(z) = \liminf_{r\to 0} \frac{\log \mu(B(z,r))}{\log r} \tag{31}$$

$$\gamma_\mu^+(z) = \limsup_{r\to 0} \frac{\log \mu(B(z,r))}{\log r}. \tag{32}$$

They are called respectively the lower and upper pointwise dimensions of μ at z. If $\gamma_\mu^-(z) = \gamma_\mu^+(z)$, the common value $\gamma_\mu(z)$ is called the pointwise dimension of μ at z. We defer to our paper [16] for a discussion of these pointwise dimensions with the associated references.

Most observables used in practice are not diffeomorphisms. The most general result concerning the local dimension of image measures is due to Hunt and Kaloshin, in particular Theorem 4.1 in [41]. Before stating their theorem, we must recall the important notion of *prevalence* used in the aforementioned paper, see also [42], [55] and [68]. We consider a real topological vector space V and a Borel-measurable subset S of V. S is said to be *prevalent* if there exists a finite-dimensional subspace P of V, called the *probe* set, such that for all $v \in V$ we have $v + p \in S$ for Leb_P-almost all $p \in P$, where Leb_P denotes the P-dimensional Lebesgue measure on P. In the case of interest for us, V is the space of C^1 functions $f : \mathbb{R}^n \to \mathbb{R}^m$. The notion of prevalence could be thought as the analogue of almost everywhere in infinite dimensional spaces. We give a few properties and examples of prevalence to point out its significance. All prevalent subsets S of V are dense in V. Then, if we declare that *almost every* means that the stated property holds for a prevalent subset of the space in question, we have:

- almost every continuous function from the interval $[0, 1]$ into $\mathbb{R}$ is nowhere differentiable. Here, V is the space of continuous functions on the unit interval with the supremum norm topology;
- take now $V = L^1(dx)$ the space of Lebesgue summable functions on the unit interval. Then almost every function $f \in V$ has the property that $\int_0^1 f(x)dx \neq 0$.
- if A is a compact subset of $\mathbb{R}^n$ with Hausdorff dimension d_H, $m \geq d_H$, and $1 \leq k \leq \infty$, then, for almost every C^k function $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(A)$ also has Hausdorff dimension d_H.

Other examples will now be stated in terms of dimension of measures. We summarize them in the following theorem:

Theorem 1. *(Hunt and Kaloshin* [41]*)*

- *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support. For a prevalent set of C^1 functions (also, for almost every linear transformation) $f : \mathbb{R}^n \to \mathbb{R}^m$,*

$$\underline{d}^f_\mu(x) = \min(m, \gamma^-_\mu(x))$$

for almost every x with respect to μ. If in addition $d^f_\mu(x)$ exists for almost every x, then for almost every f the pointwise dimension of f at $f(x)$ exists and is given by

$$d^f_\mu(x) = \min(m, \gamma_\mu(x))$$

for almost every x.

- *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support. If the pointwise dimension $\gamma_\mu(x)$ exists and does not exceed m for almost every x with respect to μ, then for a prevalent set of C^1 functions (also, for almost every linear transformation) $f : \mathbb{R}^n \to \mathbb{R}^m$, the information dimension of f, $D_1(f(\mu))$ exists and is given by the information dimension $D_1(\mu)$ of μ:*

$$D_1(f(\mu)) = D_1(\mu).$$

The information dimension of a measure μ is defined as the following limit, when it exists [7]:

$$D_1(\mu) := \lim_{r\to 0} \frac{\int \log \mu(B(x,r))d\mu(x)}{\log r}.$$

We note that with the assumptions of the second item of the preceding theorem, we have

$$D_1(f(\mu) = \int d^f_\mu(x) df_*\mu(x);\ D_1(\mu) = \int \gamma_\mu(x) d\mu(x).$$

An important class of measures are those called *exact dimensional*: they enjoy the property that

$$\gamma_\mu(x) = D_1(\mu),\ x - \mu\ a.e.$$

Notations. We will call *typical* a point x that belongs to the set of full measure giving $D_1(\mu)$. Sometimes we will simply write D_1 instead of $D_1(\mu)$ if the measure μ is clear from the context; moreover and still for exact dimensional measures we will use the short $\mu(B(x,r)) \approx r^d$ in place of $\lim_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} = d$.

Several dynamical systems with hyperbolic properties have an invariant measure that is exact dimensional. It is enough that the limit defining the local dimensions exists almost everywhere and that the measure is ergodic to have exact dimensionality [69]. In these cases the information dimension can be expressed in terms of the Lyapunov exponents and of the metric entropy.

Remark 3. *In the rest of the section, we will consider a few cases where we compute d^f_μ and compare it with the conclusions of the Hunt and Kaloshin Theorem. We will see that non-prevalent observables arise very easily in simple examples. With abuse of language we will say that an observable is prevalent if it belongs to the prevalent space of the Hunt-Kaloshin Theorem. We declare that an observable is not prevalent whenever it does not satisfy the Theorem above and for almost all choices of the target point x (typical points). Later on (example of the product of two Cantor sets), we will show an example of observable that violates the Hunt-Kaloshin Theorem for a given point x. Even in that case we will say that the observable is not-prevalent.*

[7] Otherwise one should turn to the $\liminf$ and $\limsup$. We defer to [41] and [16] for the details.

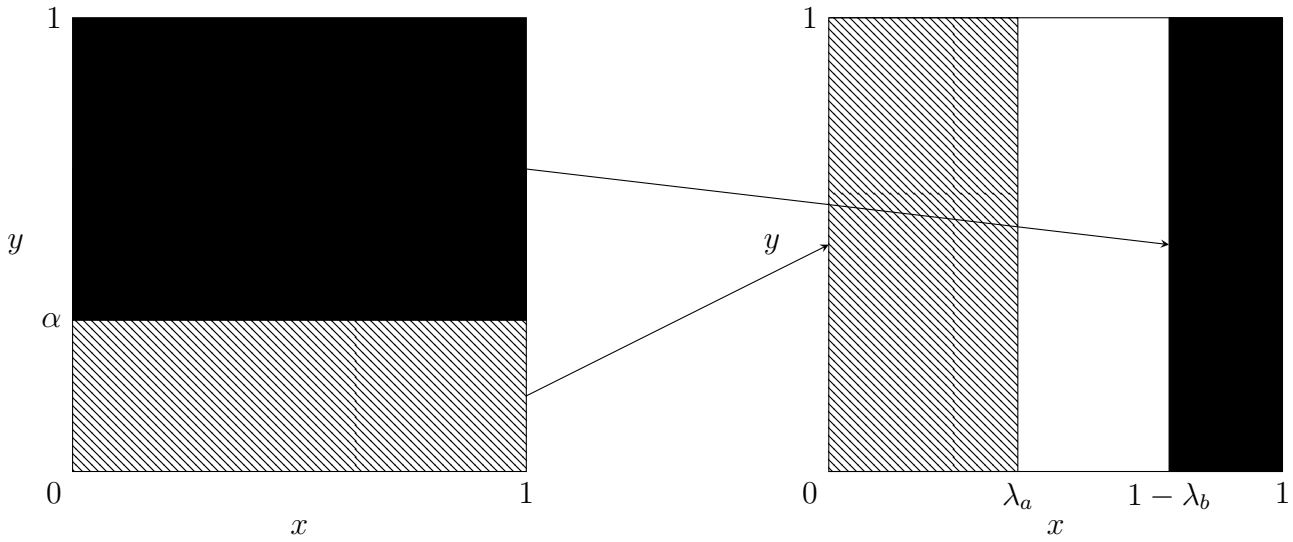


FIGURE 1. Action of baker's map on the unit square. The lower part of the square is mapped in the left part and the upper part in the right part.

4.1. **The baker map.** We start with the two dimensional dynamics defined by the baker map, whose fractal SRB measure has been extensively studied, [56], [10]. It is defined on the unit square $Q = [0,1] \times [0,1]$ by the equations

$$x_{n+1} = \begin{cases} \lambda_a x_n, \ y_n < \alpha, \\ (1-\lambda_b) + \lambda_b x_n, y_n > \alpha, \end{cases} \tag{33}$$

and

$$y_{n+1} = \begin{cases} \dfrac{y_n}{\alpha}, y_n < \alpha, \\ \dfrac{y_n - \alpha}{1-\alpha}, y_n > \alpha, \end{cases} \tag{34}$$

where $\alpha \in (0, 1/2]$ and $\lambda_a + \lambda_b \leq 1$. The action of the map on the unit square is shown in figure 1. The SRB measure is exact dimensional, and its information dimension is given by [56]:

$$D_1 = 1 + D_{1,s}, \tag{35}$$

with

$$D_{1,s} := \frac{\alpha \log(\alpha^{-1}) + (1-\alpha)\log((1-\alpha)^{-1})}{\alpha \log(\lambda_a^{-1}) + (1-\alpha)\log(\lambda_b^{-1})}.$$

The spectral approach to EVT used in section 2, applies to baker's map [4]. Let us first consider the *mean value* observable defined as

$$f(x,y) = \frac{x+y}{2}.$$

To compute numerically the quantity d_μ^f, we generate a trajectory of $M = 10^8$ points starting from a point x chosen at random on the square and compute at each iteration the value of $\phi_z(T^i x) = -\log|f_0 - f(T^i x)|$ [8]. We then compute the empirical distribution of the maximum taken by ϕ_z over blocks of size $n = 5 \cdot 10^4$. The scale parameter $\hat{\sigma}$ of the GEV distribution is computed with a maximum likelihood estimate, using the Matlab

[8]The orbit of x will approach quickly the attractor and it will give the the right statistical information by definition of SRB (physical) measure.

function *gevfit* [54]. An estimate for d_μ^f is then given by $1/\hat{\sigma}$. The estimates of d_μ^f are then averaged over 10 different trajectories. The results are displayed in table 1 (the error is the standard deviation of the results over the 10 trajectories). Although the measure has a fractal structure, we found, for different values of f_0 and α, estimates for d_μ^f that are very close to 1, as expected from the result of Hunt and Kaloshin. Since the proof of their theorem does not allow a clear geometrical understanding of what happens, we provide now an illustration and a heuristic explanation for that result.

Let us take a typical point $z = (z_1, z_2)$, not lying on the border of the square, such that $f(z) = c$, $0 < c < 1$ and let $\varepsilon > 0$. The points verifying $|f(x, y) - c| \leq \varepsilon$, are those on the straight lines $\frac{x+y}{2} = s,\ c - \varepsilon < s < c + \varepsilon$. This defines a strip where each couple (x, y) will meet infinitely many vertical unstable leaves foliating the attractor. Then the ball $B(c, \varepsilon)$ is completely filled and it could be reasonable to argue that $d_\mu^f = 1$. This would be true if the measure $f_*\mu$ is absolutely continuous, as prescribed in [60]. But there is no reason that $f_*\mu$ has such a property, if μ is not absolutely continuous. As Fig. 2 shows, we really found $d_\mu^f = 1$, which fits with Theorem 1 and suggests that f is prevalent. Before giving a rigorous direct proof of this fact, we point out that the above example can be easily modified with a drastic change in the dimension of the image measure, which therefore exhibits a non-prevalent observable. We defer to the end of this section for such an example.

We therefore consider the observable $f(x, y) = \frac{x+y}{2}$ and the strip Σ'_ε defined through $\frac{x+y}{2} = s,\ c-\varepsilon < s < c+\varepsilon$. We begin to remind that the SRB measure μ disintegrates along the vertical unstable leaves with absolutely continuous conditional measures (actually Lebesgue measures normalized to 1), and with singular measures along the horizontal stable leaves, we will use them later on. The unstable leaves $W_{u,\iota}$ are indexed by ι and counted by the counting measure ζ'; as we said, the conditional measure $\mu_{u,\iota}$ is the linear Lebesgue measure of mass 1. Then the SRB measure of the strip Σ'_ε reads:

$$\mu(\Sigma'_\varepsilon) = \int \mu_{u,\iota}(\Sigma'_\varepsilon \cap W_{u,\iota}) d\zeta'(\iota).$$

But $\mu_{u,\iota}(\Sigma'_\varepsilon \cap W_{u,\iota}) = 2\varepsilon$ and what are left in the integral above are an ensemble of unstable leaves of finite ζ' measure due to the affine term c cutting the y-axis. In conclusion $\mu(\Sigma'_\varepsilon) \approx \varepsilon$ in agreement with theorem 1. We notice that the previous proof adapts easily to all affine observables of type $f(x, y) = ax + by + c$, provided that b is different from zero.

As an interesting example of violation of prevalence, we take f as a multivariate Gaussian function maximized at the typical point $z = (x_0, y_0)$, with a covariance matrix equal to the identity:

$$f(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left((x - x_0)^2 + (y - y_0)^2\right)\right). \tag{36}$$

The set of points on Q for which $|f(x, y) - f(x_0, y_0)| \leq \varepsilon$ are the points belonging to the ball $B(z, 2\sqrt{\pi\varepsilon})$:

$$0 \leq (x - x_0)^2 + (y - y_0)^2 \leq 4\pi\varepsilon.$$

Since the point (x_0, y_0) is typical, we have $d_\mu^f = \frac{D_1}{2}$.

	$\alpha = 1/5$	$\alpha = 1/4$	$\alpha = 1/3$
$f_0 = 0.1$	1.00 ± 0.01	1.00 ± 0.02	1.00 ± 0.01
$f_0 = 0.3$	1.00 ± 0.01	1.00 ± 0.01	1.00 ± 0.01
$f_0 = 0.8$	1.00 ± 0.01	1.00 ± 0.01	1.00 ± 0.02

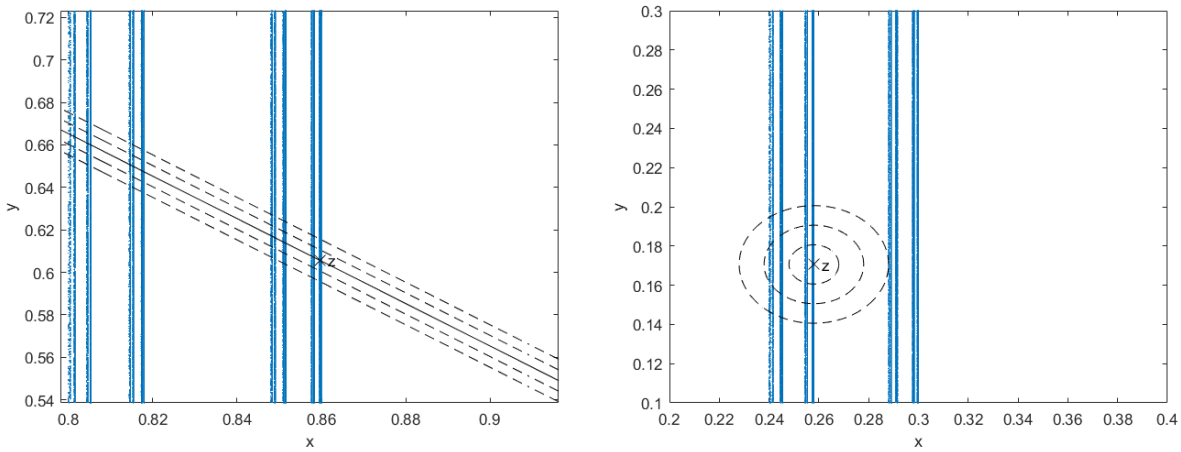


FIGURE 2. Pictorial representation of the situation described in the main text for the observable $\frac{x+y}{2}$ (left) and a Gaussian centered at z (right), in different regions of Q. The baker attractor is depicted in blue, and the graphs $\{(x,y) : f(x,y) = f(z) \pm \varepsilon\}$ are dotted lines. In both situations, these manifolds intersect the attractor an infinite number of times.

TABLE 1. Values of $d_\mu^f(f_0)$ computed for the mean value observable, for different values of α and f_0. We took $\lambda_a = 0.3$, $\lambda_b = 0.2$.

As mentioned earlier, we expect to detect no clustering of high values for such generic observables and non-periodic z. The extremal index is computed using the estimate $\hat{\theta}_5$, using as a threshold the 0.999-quantile of the observable distribution. Results are averaged over 10 trajectories and are presented in table 2.

	$\alpha = 1/5$	$\alpha = 1/4$	$\alpha = 1/3$
$f_0 = 0.1$	1 ± 0	1 ± 0	1 ± 0
$f_0 = 0.3$	1 ± 0	1 ± 0	1 ± 0
$f_0 = 0.8$	1 ± 0	1 ± 0	1 ± 0

TABLE 2. Values of θ computed for the mean value observable, for different values of α and f_0. We took $\lambda_a = 0.3$, $\lambda_b = 0.2$. The error of 0 is the standard deviation of the estimates.

To get the quantity d_μ^f different from 1 on a set of full measure, we should take an observable with range at least in $\mathbb{R}^2$.

We performed numerical computations using the baker map with parameters $\alpha = 1/3$, $\lambda_a = 1/3$, $\lambda_b = 1/4$ and by taking the observable $f(x,y) = (x, x^2 + y^2)$. For different points z not lying on the $x-$axis, we find indeed a value of $d_\mu^f(z)$ that is close to the information dimension $= 1.2682$, as it is computed from Eq. (35). For the point $z = (0.9581, 0.0612)$ for example, we find a local dimension equal to 1.26 ± 0.03. We used again the parameters $M = 10^8$ and $n = 2 \cdot 10^5$.

As promised above, we now give another example of a non-prevalent observable. Let us take the function $f(x,y) = x$ and $f_0 = f(z_1, z_2) = z_1 = c$, where (z_1, z_2) is a typical point. We need to compute the scaling of the SRB measure of the vertical strip $\Sigma_\varepsilon := \{(x,y) \in Q; |x - c| \leq \varepsilon\}$. To this end, we disintegrate the SRB measure μ along the horizontal

stable leaves. These measures can be seen as generated by an Iterated Function System (IFS) with two scales λ_a, λ_b and two weights $\alpha, 1-\alpha$, [10]. We now evaluate the SRB measure of the strip Σ_ε, as:

$$\mu(\Sigma_\varepsilon) = \int \mu_{s,\nu}(\Sigma_\varepsilon \cap W_{s,\nu}) d\zeta(\nu), \tag{37}$$

where $\mu_{s,\nu}$ is the conditional measure along the stable leaf $W_{s,\nu}$, indexed by ν and counted by the counting measure ζ. These conditional measures are the same on each $W_{s,\nu}$ and for almost all choices of z_1 they behave as exact dimensional fractal measure with the exponent given by the term $D_{1,s}$ in Eq. (35):

$$\mu_{s,\nu}(\Sigma \cap W_{s,\nu}) \approx \varepsilon^{D_{1,s}}.$$

Since the counting measure $\zeta(Q) = 1$, we finally get $\mu(\Sigma_\varepsilon) \approx \varepsilon^{D_{1,s}}$, which violates theorem 1 because the exponent should be equal to 1.

4.2. **The product of two Cantor sets.** As a second example which can be worked out analytically, we consider the cartesian product of two ternary Cantor sets $K \times K$ on the unit interval I. The dynamics is generated by two independent iterated function system (see [56]), each of them defined by two linear contractive maps g_1, g_2 with slope 1/3. On each factor K we take a measurable map, our dynamical system, $T : K \to K$, with $T(x) = g_i^{-1}(x)$, for $x \in g_i(K)$. The Cantor set K will be the invariant set for the transformation T. The invariant measure $\mu^{(2)} = \mu \times \mu$ is the product of the two invariant measures on the factor spaces. Each factor measure is a balanced measure with two equal weights 1/2, which means that for any Borel set B on the unit interval we have $\mu(B) = \sum_{i=1}^{2} \frac{1}{2}\mu(g_i(B))$. All these measures are exact dimensional and the information dimension of $\mu^{(2)}$ is $D_1 = \frac{2\log 2}{\log 3} \approx 1.26$.
The spectral approach to EVT used in section 2, applies to this systems, see [26] and [36]. As a first observable, we take the standard multivariate Gaussian function (36). If we take a typical point $z := (x_0, y_0)$, we can repeat the argument given above for the baker's map and found $d_\mu^f = D_1/2$, which shows that (36) is not prevalent. This result is confirmed by the numerical simulations for which we used the same algorithm as described earlier for the baker map, using the parameters $M = 5 \cdot 10^8$ and $n = 5 \cdot 10^4$. We used more data and a different size of blocks because stable estimates are difficult to get. We detected no clustering, as for the baker's map. The results for the estimates of d_μ^f are shown in table 3. We found for different points z a value for d_μ^f close to 0.61, which is comparable with $D_1/2 \approx 0.63$.

We now take $f(x, y) = x$ and look at the strip $\Sigma_\varepsilon = \{|x - c| \leq \varepsilon\}$, where c is chosen on a typical line (c, y). Instead of disintegrating, we can now use Fubini's theorem since $\mu^{(2)}$ is a product measure. We have

$$\mu^{(2)}(\Sigma_\varepsilon) = \int \mu_x(\Sigma_\varepsilon \cap K) d\mu_y$$

where we write μ_x (resp. μ_y) for the factor measure on the Cantor set K on the x-axis (resp. on the y-axis). The sectional measure $\mu_x(\Sigma_\varepsilon \cap K)$ is independent of y and since it is exact dimensional it yields $\mu_x(\Sigma_\varepsilon \cap K) \approx \varepsilon^{D_1/2}$, which finally gives $d_\mu^f = D_1/2$, showing that the observable is not prevalent. The same results holds for the observable $f(x, y) = y$.

A less trivial observable that violates prevalence is given by $f(x, y) = y - x$ and $f_0 = f(0, 0) = 0$. We warn the reader that we are not sure that the point $(0, 0)$ is typical,

but we discuss this situation since it could arise in concrete applications, in the same way the periodic points are negligible in measure but play an important role in recurrence. We therefore have to compute the logarithm of the $\mu^{(2)}$ measure of the ε neighborhood Σ_ε of the diagonal $\{x = y\}$ and compare it with the logarithm of ε. It is easy to check that we can restrict ourselves to countable sequences like $\varepsilon_n = \beta^n$, with $\beta > 1$ and $n \to \infty$. In particular we now take $\varepsilon_n = 3\ 3^{-n}$. Then we have as above $\mu^{(2)}(\Sigma_\varepsilon) = \int \mu_x(\Sigma_\varepsilon \cap K) d\mu_y$. Each time $x \in K$ on the x-axis, Σ_ε will meet the Cantor set $K \times K$ in the point (x, y). We therefore evaluate the μ_y measure of the section $\Sigma_\varepsilon \cap K$ by splitting it over the 2^n cylinders of the n-th generation in the construction of the Cantor set along the y-axis. There will be at least one of these cylinders of μ_y-measure 2^{-n} inside that section. Then

$$\mu^{(2)}(\Sigma_{\varepsilon_n}) = \int \mu_x(\Sigma_{\varepsilon_n} \cap K) d\mu_y \geq 2^{-n}.^{9}$$

With the prescribed choice of ε_n we finally get

$$\limsup_{n\to\infty} \frac{\log \mu^{(2)}(\Sigma_{\varepsilon_n})}{\log \varepsilon_n} \leq \frac{\log 2}{\log 3},$$

which immediately implies $\overline{d}^f_{\mu^{(2)}}(0,0) \leq \frac{\log 2}{\log 3}$.

It is difficult to prove directly that $d^f_\mu = 1$ as prescribed by the Hunt-Kaloshin theorem for the *almost sure prevalence affine functions* $f(x, y) = ax + by + c$ (see figure 3 for a pictorial representation). Numerical experiments confirm such a behavior, although stable estimates are difficult to get. We chose the parameters a, b and c at random in the unit interval. For the point $z = (0.893, 0.307)$, $a = 0.557$, $b = 0.6596$, $c = 0.0046$, we find that $d^f_\mu = 0.993 \pm 0.02$. We used the same parameters as for the Gaussian observable.

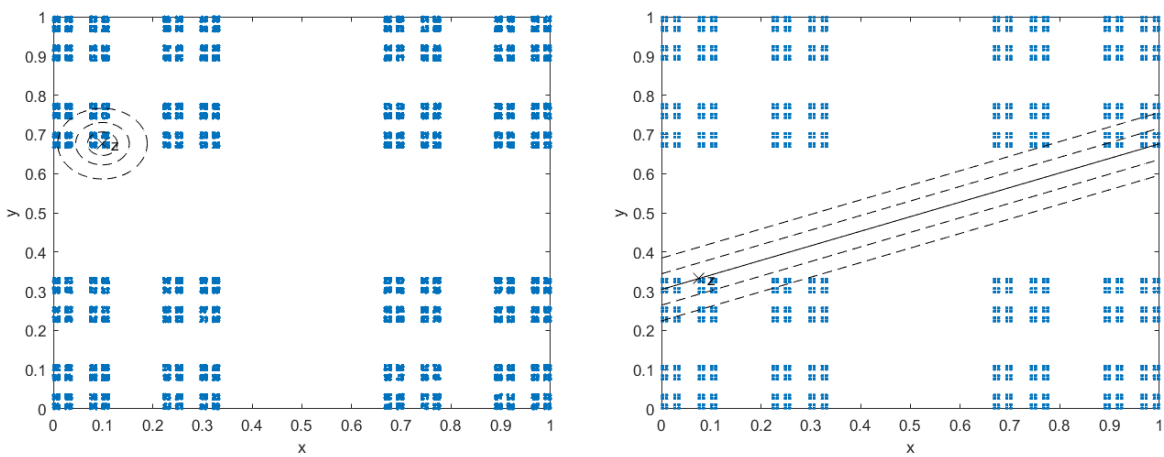


FIGURE 3. For the product of Cantor sets, pictorial representation of the situation described in the main text for the gaussian observable (left) and a generic linear observable (right). The attractor is depicted in blue, and the graphs $\{(x, y) : f(x, y) = f(z) \pm \varepsilon\}$ are dotted lines.

[9]Actually we should remove from it a small contribution $o(2^{-n})$ due to integration on the corners $(0, 0), (1, 1)$, which will not affect the final result.

z	$(0.994, 0.0029)$	$(0.6679, 0.9914)$	$(0.0861, 0.2565)$
d_μ^f	0.61 ± 0.002	0.60 ± 0.002	0.62 ± 0.002

TABLE 3. Values of d_μ^f computed for the Gaussian observable, for different points z. The error is the standard deviation of the results.

4.3. **The Lorenz system.** Let us now turn to a higher dimensional situation and consider the Lorenz 1963 system [49] that we reconstruct with the Euler method with step $h = 0.01$. With this iterative procedure, the system can be seen as a discrete mapping, for which the developed theory is applicable. We chose an observable with image in $\mathbb{R}^5$. We tested several observables but the results are displayed for the observable $f(x, y, z) = (x^2 + y^2, z, y + z, \pi yz, 1/x)$. We find that the values of $d_\mu^f(x, y, z)$ are all close to $D_1 = 2.06$. For the point $(-1.7323, 8.9400, 32.6818)$ for example, we have that $d_\mu^f(x, y, z) = 2.05 \pm 0.02$ using the parameters $M = 10^8$ and $n = 2 \cdot 10^5$ (the results are averaged over 20 trajectories, and the error is the standard deviation of the results).

Instead, if we take a scalar observable, we find values very close to 1, indicating again that when the observable decreases the dimensionality, the fractal structure of the attractor is smoothed in the image measure (we are *supposing* here that the invariant measure is exact dimensional). These numerical results are in perfect agreement with the Hunt-Kaloshin Theorem.

4.4. **Conclusions.** We conclude this section by pointing out the few examples which we found and do not verify the conclusions of the Hunt-Kaloshin Theorem. It happens when $f : \mathbb{R}^n \to \mathbb{R}^m$ with $m \leq n$. If the dimension of the attractor in $\mathbb{R}^n$ is larger than m, one expects to find $d_\mu^f = m$ for a prevalent observable. We exhibit several examples where, in the same circumstances, $d_\mu^f < m$. This shows that we are in presence of a non prevalent observable.

Another example of observable that does not belong to the prevalent set of the measure is a function whose Jacobian does not have a full rank on a set of positive measure, for absolutely continuous measures. This is a consequence of theorem 9 in [60]. We emphasize that the image measure can have counter-intuitive properties. For instance, Rousseau [62] gives the example of an image measure that is non atomic and yet d_μ^f is 0 on a set of positive measure. This example is built upon a Cantor set and the C^∞ observable is defined as the limit of an iterative process.

5. STATISTICS OF VISITS FOR THE OBSERVABLE

It can be interesting from a physical point of view to study the number of visits of the observable f near a certain value $f_0 = f(z)$. This problem is well understood in the framework of EVT. Let us consider the following counting function:

$$N_n(t) = \sum_{l=1}^{\lfloor \frac{t}{f*\mu(B(f_0, r_n))} \rfloor} \mathbf{1}_{B(f_0, r_n)}(f(T^l x)), \tag{38}$$

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

$$\mu(N_n(t) = k), \ k \in \mathbb{N} \tag{39}$$

when $n \to \infty$. It has been proved (see for instance [37, 33, 34]) that for $f = Id$ and 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, $\mu(N_n(t) = k)$ converges to the Polyà-Aeppli distribution, which is a particular kind of compound Poisson distribution. Before continuing, we remind that a probability measure $\tilde{\nu}$ on $\mathbb{N}_0$ is compound Poisson distributed with parameters $t\lambda_\ell$, $\ell = 1, 2, \ldots$, if its generating function $\varphi_{\tilde{\nu}}$ is given by $\varphi_{\tilde{\nu}}(z) = \exp \int_0^\infty (z^x - 1)\, d\rho(x)$, where ρ is the measure on $\mathbb{N}$ defined by $\rho = \sum_\ell t\lambda_\ell \delta_\ell$, with δ_ℓ being the point mass at ℓ.

An important non-trivial compound Poisson distribution is the Pólya-Aeppli distribution which holds when the random variables given by the hitting times of the ball $B(f_0, r_n)$ is geometrically distributed, which implies that $\lambda_\ell = (1-p)p^{\ell-1}$ for $\ell = 1, 2, \ldots$, for some $p \in (0, 1)$. In this case

$$\mathbb{P}(N_n(t) = k) \to e^{-pt} \sum_{j=1}^{k} p^{k-j}(1-p)^j \frac{(pt)^j}{j!} \binom{k-1}{j-1}, \ n \to \infty, \tag{40}$$

where p is the extremal index. In particular $\mathbb{P}(W = 0) = e^{-t}$. In the case of $p = 1$ this reverts to the usual Poisson distribution. For more general target sets, the limit law of $N_n(t)$ is given by a compound Poisson distribution when the extremal index is different from 1, and by a pure Poisson distribution if no clustering occurs, [37], [38]. We refer also to our paper [15] for a discussion of this matter and related references.

We now show that in presence of non-invertible observables f, we get compound Poisson distributions which are not Pòlya-Aeppli.

Proposition 3. *With the assumptions of Proposition 2, suppose the ball $B(f_0, r_n)$, $f_0 = f(z)$, has two pre-images $B_{1,n}, B_{2,n}$, the first containing the periodic point $w_1 = z$ of period p_1, the second the periodic point w_2 of period p_2. Then the distribution $N_n(t)$ is compound Poisson, but not Pòlya-Aeppli.*

Proof. We notice that Eq. (38) can be rewritten as

$$N_n(t) = \sum_{l=1}^{\lfloor \frac{t}{f_*\mu(B(z,r_n))} \rfloor} \mathbf{1}_{f^{-1}(B(f_0,r_n))}(T^l x)$$

We can therefore apply the theory recently developed by [38], where entry times are considered for sets whose measure goes to zero. In our case those sets are the pre-images of the ball $B(f_0, r_n)$ and they are located around the points $w_i \in A_k, k \in \mathbb{N}$, where the set A_k has been defined in Proposition 2; actually there are now only two pre-images.

If we now refer to the theory in [38] and in particular to Section 8.3 therein, we can easily compute the quantity $\tilde{\alpha}_l = \sum_i b_i^l$, where $b_i^l := \lim_{n\to\infty} \mu_{B_n}(\tau_{B_n}^{l-1} = i)$, being $B_n = B_{1,n} \cup B_{2,n}$ and $\tau_{B_n}^{l-1}$ is the $l-1$th return time into the set B_n; with μ_A we intend the conditional measure to the set A. For a given l, only the terms $(l-1)b_1$ and $(l-1)b_2$ count in the sum defining $\tilde{\alpha}_l$. By repeating the computation in Lemma 4 in [38] we have

$$\tilde{\alpha}_l = b_1^{l-1} \lim_{n\to\infty} \left[\frac{\mu(B_{1,n})}{\mu(B_n)} \right] + b_2^{l-1} \lim_{n\to\infty} \left[\frac{\mu(B_{2,n})}{\mu(B_n)} \right],$$

where

$$b_1 = \lim_{n\to\infty} \left[\frac{\mu(B_{1,n} \cap T^{-p_1} B_{1,n})}{\mu(B_{1,n})} \right], \ b_2 = \lim_{n\to\infty} \left[\frac{\mu(B_{2,n} \cap T^{-p_2} B_{2,n})}{\mu(B_{2,n})} \right]$$

Notice that in the particular case we are considering and by repeating the computation in section 3 we have

$$b_1 = \frac{1}{|(T^{p_1})'(w_1)|}, \; b_2 = \frac{1}{|(T^{p_2})'(w_2)|}, \tag{41}$$

$$\mu_1 := \lim_{n\to\infty} \left[\frac{\mu(B_{1,n})}{\mu(B_n)}\right] = \frac{1}{1+\frac{h(w_2)|f'(w_1)|}{h(w_1)|f'(w_2)|}}, \tag{42}$$

$$\mu_2 := \lim_{n\to\infty} \left[\frac{\mu(B_{2,n})}{\mu(B_n)}\right] = \frac{1}{1+\frac{h(w_1)|f'(w_2)|}{h(w_2)|f'(w_1)|}}. \tag{43}$$

Moreover by recalling the definition of the quantities q_k introduced in section 3 we have

$$q_{p_1} = b_1 \; \mu_1; \; q_{p_2} = b_2 \; \mu_2.$$

According to the theory developed in [38], the parameter λ_l which we introduced before Eq. (40) to define the compound Poisson distribution is given by

$$\lambda_l = \frac{\alpha_l - \alpha_{l+1}}{\alpha_1}, \text{ where } \alpha_k = \tilde{\alpha}_k - \tilde{\alpha}_{k+1},$$

and α_1 is the extremal index defined as the *reciprocal of the expected length of the clusters*:

$$\sum_{k=0}^{\infty} k\lambda_k = \frac{1}{\alpha_1}.$$

In our case and using the expression for the quantities introduced above we have:

$$\alpha_1 = 1 - (q_{p_1} + q_{p_2}).$$

The latter is an alternative way to define the extremal index, which in the present situation is consistent with the formula found in section 3 for the extremal index θ using the spectral technique. We defer to our article [15] for a critical discussion of these equivalent definitions. Moreover, putting $1 - b_1\mu_1 - b_2\mu_2 = \theta$, we have

$$\lambda_l = \frac{b_1^{l-1}\mu_1(1-b_1)^2 + b_2^{l-1}\mu_2(1-b_1)^2}{\theta}$$

and for the generating function of the random variable given by the number of visits

$$\begin{aligned} \phi(z) &= \exp\left[\sum_{l=1}^{\infty} \theta t \lambda_l (z^l - 1)\right] \\ &= e^{-\theta t} e^{t\mu_1(1-b_1)^2 \frac{z}{1-zb_1}} e^{t\mu_2(1-b_2)^2 \frac{z}{1-zb_2}}, \end{aligned} \tag{44}$$

which gives a compound distribution different from the Polyà-Aeppli distribution.

□

We remind that deviations from the Pòlya-Aeppli distribution were exhibited in other situations, for instance when the target set is a neighborhood of the diagonal in [37, 15] or a neighborhood of periodic points where the map is not continuous in [1].

We now give a recursive formula that produces the distribution of $N_n(t)$. Let us denote

$$a_1 = t\mu_1(1-b_1)^2 \text{ and } a_2 = t\mu_2(1-b_2)^2.$$

We first notice that

$$\phi'(z) = \phi(z)\pi(z)', \tag{45}$$

where

$$\pi(z) = \frac{a_1 z}{1 - b_1 z} + \frac{a_2 z}{1 - b_2 z}.$$

We easily see that the k derivatives of π (for $k > 0$) are given by

$$\pi^{(k)}(z) = \frac{k! a_1 b_1^{k-1}}{(1 - b_1 z)^{k+1}} + \frac{k! a_2 b_2^{k-1}}{(1 - b_2 z)^{k+1}}. \tag{46}$$

Using the Leibniz formula for derivations, we have from Eq. (45):

$$\phi^{(n)}(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} \phi^{(k)}(z) \pi(z)^{(n-k)}(z).$$

We now use this last formula and combine it with Eq. (46) to obtain:

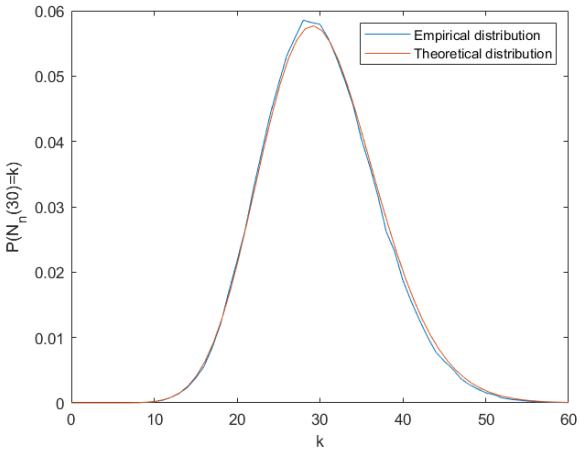
$$\phi^{(n)}(0) = \sum_{k=0}^{n-1} \binom{n-1}{k} \phi^{(k)}(0)(n-k)!(a_1 b_1^{n-k-1} + a_2 b_2^{n-k-1}). \tag{47}$$

Keeping in mind that from Eq. (44), $\phi(0) = e^{-\theta t}$, we can use formula (47) to determine the probability that $N_n(t) = k$ by computing recursively the derivatives of the generating function ϕ at 0 and dividing by $k!$.

We now give an example. We take the map $Tx = 3x \mod 1$, the observable $f(x) = (x - 1/2)(x - 1/4)$ and $f_0 = 0$. The two pre-images of f_0 are $1/2$ and $1/4$, of periods 1 and 2 respectively. From proposition 2, $\theta = 7/9$. Then we have: $b_1 = 1/3$, $b_2 = 1/9$, $\mu_1 = \mu_2 = 1/2$, $a_1 = 2t/9$ and $a_2 = 32t/81$. In figure 4 we show the empirical distribution of the number of visits of 10^5 different trajectories of length 10^6 of the observable f in the interval $I_r = (f_0 - r, f_0 + r)$, where $r = e^{-u}$, u being the 0.995-quantile of the distribution of ϕ. We notice very good agreement with the theory.

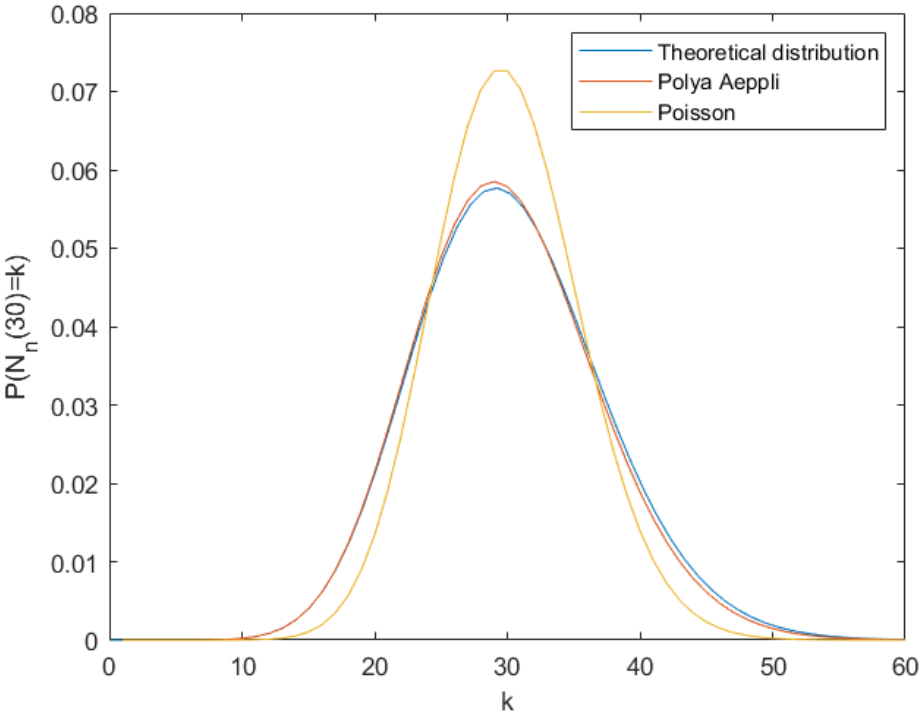
It is interesting to observe that if we take $\mu_1 \neq \mu_2$, but $b_1 = b_2 = b$, which means we take the same periodicity for the two points w_1, w_2, we recover the Pòlya-Aeppli distribution since $\lambda_l = b^{l-1}(1 - b)$, with the extremal index $\theta = 1 - b$. In fact, even when $b_1 \neq b_2$, numerical experiments suggest that the distribution stays close to a Pòlya-Aepply distribution. In figure 5, we show this effect by comparing the distribution associated with the example described in the text to a Pólya-Aeppli distribution of parameters given by $\theta = 7/9$ and $t = 30$. The vicinity between the two distributions is striking and is found in a whole variety of examples. In [15], we also observed this phenomenon in cases when the clustering structure is even more complex and for systems perturbed with discrete noise.

As we mentioned earlier, for a whole variety of observables f, no clustering is detected and the EI is 1. We therefore expect to have a Poisson distribution for the statistics of visits. This is indeed what we observed for the baker map, with parameter $\alpha = 1/3$, and the observable $f(x, y) = \frac{x+y}{2}$ (see figure 6). We took a point z at random in the attractor (actually we iterated a point in the basin several time to get it very close to the attractor), and computed the empirical distribution of the number of visits of 10^5 different trajectories of length 10^6 for the observable f in the interval $I_r = (f_0 - r, f_0 + r)$, where $r = e^{-u}$, u being the 0.995-quantile of the distribution.



0.5

FIGURE 4. Comparison between the empirical distributions of the number of visits of the observable $f(x) = (x - 1/2)(x - 1/4)$ in a ball centered at 0 and the theoretical distribution described in the text for the map $3x$ mod 1.



0.5

FIGURE 5. Comparison between the distributions of the number of visits for the example in the text, a pure Poissonian distribution and a Polya Aeppli distribution with parameters given by $t = 30$ and the extremal index $\theta = 7/9$.

6. Randomly perturbed systems

One could wonder what happens to the theory developed above when the dynamical system is randomly perturbed; this has of course important physical applications when

0.5

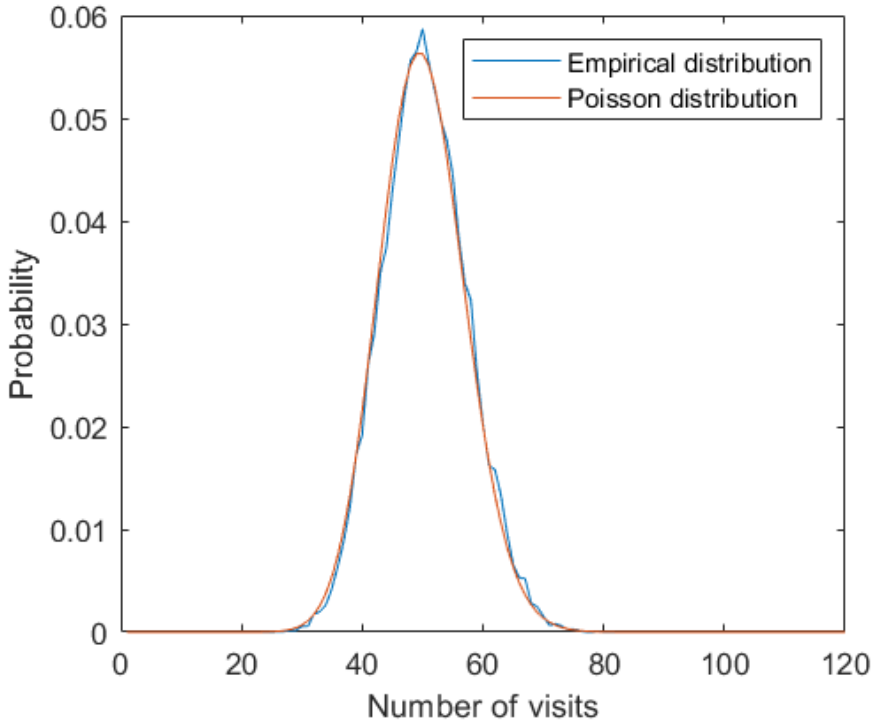


FIGURE 6. Comparison between the empirical distributions of the number of visits of the observable "mean value" in a ball centered at $f(z)$ and a Poisson distribution for the baker map.

the system or its environment are affected by noise or when the available time series give only a partial description of the evolution of the system variables. As we anticipated in the Introduction, we will show that with suitable but very general choices of the perturbations on the map or on the observable f with values in R^m, the dimension of the image measure will increase to m if less than the dimension m before perturbation, and drops to m otherwise.

6.1. **Perturbing the map.** We defer to our paper [15] for an exhaustive presentation of different random perturbations in connection with recurrence properties. For the purposes of this paper, we will consider random transformations, where the iteration of the single map T is replaced by the concatenation $T^n_{\underline{\omega}} := T_{\omega_n} \circ \cdots \circ T_{\omega_1}$, where the $\omega_k \in \underline{\omega} := (\omega_1, \cdots, \omega_k, \cdots)$ are i.i.d. random variables with (common) distribution $\mathbb{G}$. Sometimes it is possible to show the existence of the so-called *stationary measure* ρ_s, verifying for any real bounded function q: $\int q d\rho_s = \int q \circ T_\omega d\rho_s d\mathbb{G}$, see [50] Chap. 7, for a general introduction to the matter. The product $\mathbb{P} := \rho_s \times \mathbb{G}^{\mathbb{N}}$ will give a stationary measure for the random process $q(T^n_{\underline{\omega}}(x), \sigma^n(\underline{w}))$, where σ denotes the shift. The measure $\mathbb{P}$ will allow us to consider the limit theorems for such random processes in the so-called *annealed setting*; it will also weight the sets $B_{n,z}, C_{n,z}$ entering in the definition of the quantities $q_{n,k}$ expressing the extremal index. We defer to our papers [15] and [1] for the analytic derivation of the extremal index in the annealed setting. We showed there in several examples, that whenever the distribution $\mathbb{G}$ has a density, the EI becomes equal to 1, while it could be less than one for discrete distributions. The same happens in the present situation as the following two relatively simple situations indicate.

- *Continuous noise.* We consider a map T verifying the assumptions in Proposition 2 and in particular we define it on the circle; we will say later how to generalize the result to the interval. We perturb T with additive noise, namely we put $T_\omega(x) = T(x) + \omega$- mod 1 and we choose ω with some smooth distribution $\mathbb{G}$ with density q bounded from below. It is therefore possible to prove the existence of a stationary measure $\rho_s = h_s d\text{Leb}$ absolutely continuous with respect to Lebesgue with density h_s. The computation of the extremal index follows now exactly the proof of Proposition 5.3 in [1] with one difference: the connected ball U_m there is now replaced by the set $f^{-1}B_{n,z}$ which is, in general, the disjoint union of a finite number of preimages. These sets are "centered" at the pre-images $\{z_l\}_{l\geq 1}$ of the target point $f(z)$. The key idea in [1] was to show that for the majority of the realizations, with respect to $\mathbb{G}^{\mathbb{N}}$, the numerator in the quantities $q_{k,n}$ was zero. The rest was of higher order with respect to the denominator and vanished in the limit of large n. The control in the numerator was based on the possibility to achieve, for a big portion of realizations $\underline{\omega}$, that $|T^j_{\underline{\omega}}(z) - z| > 2(\max|T'|)^j|U_m|$, where $|U_m|$ denotes the diameter of $|U_m|$ and the latter is centered at z. It easy to see that the same lower bound persists when the random orbit $T^j_{\underline{\omega}}$ is computed starting from, say, z_{l_1} and the right-hand side of the bound is replaced by the set $f^{-1}B_{n,z_{l_2}}$ around another point of the sequence $\{z_l\}_{l\geq 1}$. This is possible since the diameters of the $f^{-1}B_{n,z_l}$ are comparable, since f is piece-wise C^1. We left the details to the reader. At the end we get that all the $q_l = 0$, and therefore the extremal index $\theta = 1$. As we said above the proof extends easily to the additive perturbation of a piece-wise expanding map with finitely many branches verifying the other assumptions of Proposition 2.
- *Discrete noise* The purpose here is to give paradigmatic examples of the applicability of our theory with observable, leaving specific cases to other occasions. For the discrete noise we could adapt to our first example described at the beginning of section 3 with an invertible f, the example studied in section 4.1.2 in our paper [15]. We considered there two maps on the circle $T_0 = 2x$- mod 1 and $T_1 = 2x + b$- mod 1, $0 < b < 1$. If we now take the observable f which is zero in 0, $f(0) = 0$, we can repeat the argument in [15] with the set $B(0, e^{-u_n})$ there replaced by our $f^{-1}B(0, e^{-u_n})$. The conclusion is that $q_0 > 0$, and that $\theta < 1$.

We argued in section 3 that in presence of observables the EI is difficult to compute; we believe that if in addition the system is randomly perturbed the EI is even more complicated and in general it should be 1 or close to it.

The computation of d^f_μ in presence of noise is also interesting. We first point out that our Proposition 1 easily generalizes to the annealed situation as we proved in [15] for discrete distributions and in [1] for distributions with density. Moreover, we suppose that the target set is fixed and the parameter τ defining the boundary level in Eq. (7) is independent of the noise, so that what we estimate *via* the convergence to the Gumbel law is the stationary measure of sets of type $\rho_s(B(f(z), \epsilon)$. It is therefore interesting to evaluate that stationary measures; there are several ways to determine the existence of a stationary measures in connection with a given random perturbation, see for instance [3, 67, 6]. Usually one needs a precise description of the stationary measures in order to establish stochastic stability, namely to recover the statistical properties of the unperturbed system when the noise is sent to zero. We are not interested in it; instead we are interested in getting an experimental way to construct a stationary measure and check its general properties. A useful result by Alves and Araujo [3] will provide us with what we need.

The idea is to look for a composition of maps close to a given one T and assume that the noise will verify two *nondegeneracy conditions*, namely:

- (N1) The measure $\mathbb{G}^{\mathbb{N}}$ will be supported on a small set $S_\epsilon := (\text{supp}\mathbb{G})^{\mathbb{N}}$ such that there is $\chi = \chi(\epsilon)$, for which each random orbit $T^n_{\underline{\omega}}(x)$ contains the ball of radius χ around $T^n(x)$ for all $x \in X$ and n sufficiently large. As is written in [3], this "condition means that perturbed iterates cover a full neighborhood of the unperturbed ones after a threshold for all sufficiently small noise level."
- (N2) We require that the measure $\int \mathbf{1}_A(T^n_{\underline{\omega}}(x))d\mathbb{G}^{\mathbb{N}}(\underline{\omega})$ for any Borel set A be absolutely continuous with respect to the Lebesgue measure Leb on X, for all $x \in X$ and n sufficiently large. This means that "sets of perturbation vectors of positive $\mathbb{G}^{\mathbb{N}}$ measure must send any point $x \in X$ onto subsets of X with positive Lebesgue measure after a finite number of iterates", [3].

We now fix $x \in X$ and consider the measure, for any Borel set $A \subset X$:

$$\rho_n(A) := \frac{1}{n}\sum_{j=0}^{n-1}\int \mathbf{1}_A(T^j_{\underline{\omega}}(x))d\mathbb{G}^{\mathbb{N}}(\underline{\omega}). \tag{48}$$

It has been proved in [3], Lemma 3.5, that every weak* accumulation point of the sequence ρ_n is stationary and absolutely continuous with respect to the Lebesgue measure whenever (N2) holds. Notice that the Cesaro mean in Eq. (48) is exactly the numerical procedure to get the measure of a set by averaging over different realizations $\underline{\omega}$, so that we expect that with noise verifying the assumptions (N1) and (N2), the stationary measure ρ_s is absolutely continuous with respect to Lebesgue [10]. This has an interesting consequence for the computation of $d^f_{\rho_s}$, since in presence of smooth observable f and absolutely continuous measure ρ_s, Theorem 9 in [60] states that the dimension of the observable exists ρ_s-almost everywhere, is integer and is equal to the rank of Df. We therefore expect that for such noises, the non-integer dimensions computed in the preceding examples for non-prevalent observable, become integer. For prevalent observable with large dimensionality, $d^f_\mu = D_1 < m$, where m is the dimension of the range of f, we expect that d^f_μ drops to $\min(m, n)$, (n being the dimension of the ambient space of the original system) in presence of a smooth stationary measure.

We tested this result by considering the dynamics on the product of two Cantor sets, with the non prevalent observable $f(x, y) = x - y$, and $f(z) = f(0, 0) = 0$. The original dynamics given by an iterative function system is perturbed by an additive noise drawn with a uniform distribution in $B(0, \eta)$, for a small $\eta > 0$. To avoid that the dynamics leaves the unit square, we apply the mod-1 folding after having applied the additive perturbation. We observe in figure 7 that d^f_μ, which is about 0.63 when $\eta = 0$ goes to 1 as η increases. To compute d^f_μ, we simulated trajectories of 10^7 points and considered blocks of size 10^3.

Of course, if the noise does not verify assumptions (N1) and (N2), we do not know anymore if any weak limit of Eq. (48) is absolutely continuous. This is in particular true if the unperturbed map T will not preserve an absolutely continuous invariant measure. Otherwise and for uniformly expanding maps, it is always possible to get stationary measure which are absolutely continuous and that independently of the nature of the noise, [6].

[10] We notice that by general results on random perturbations, see for instance [3, 6], if the map T preserves a unique absolutely continuous invariant measure, the absolutely continuous stationary measure is also unique.

If the stationary measure exists, the Hunt-Kaloshin Theorem still applies for the perturbed system, whatever the perturbation is. Indeed, this Theorem concerns measures and not the underlying dynamics.

When the perturbation is discrete, and the original measure has a fractal structure, the shape of the stationary measure is not yet completely understood. We therefore choose to study it numerically. We considered the successive iterations of a baker map with $\lambda_a = \lambda_b = 0.4$ and the parameter α equal to $1/4$ and $1/3$, each one with probability $1/2$. For the observable $f(x, y) = (x + y, y^2)$, we found values for d_μ^f around 1.70 (we averaged the results over 20 points of the attractor), which we interpreted as the local dimensions of the stationary measure. In fact, when we compute directly the local dimensions of this system, we also find a value of around 1.70. If we now we take a scalar observable $f(x, y) = x^2 - y$, we find as expected values close to 1 for d_μ^f.

6.2. **Perturbing the observable.** We now suppose that the map T does not change, but the observable does. In particular we assume that it changes in an i.i.d. way at each iteration. This could have physical importance since it models the influence of a random environment on the deterministic dynamics, or the uncertainty associated with the measurement process. By using the notations of section 1, we now consider the maximum of the random variables, for $k = 0, \cdots, n-1$:

$$\phi(T^k(x), \omega_k) = -\log(\mathrm{dist}(f_{\omega_k}(T^k(x)) - f(z))),$$

where the ω_k are i.i.d. random variable with common distribution $\mathbb{G}$ and $f(z)$ is the value of a fixed observable at the point z. The probability will now be $\mu \times \mathbb{G}^{\mathbb{N}}$ and we indicate it with $\mathbb{P}$. We write again $\underline{\omega}$ for the vector with components $\{\omega_k\}_{k=0\cdots,\infty}$. The maximum will therefore be a function of x and $\underline{\omega}$, $M_n := M_n(x, \underline{\omega})$. By setting ourselves in the framework of the uniformly expanding maps considered in section 1, we immediately have

$$\mathbb{P}(M_n \leq u_n) = \int (\mathbf{1}_{C_{n,z}} \circ f_{\omega_0})(x) \cdots (\mathbf{1}_{C_{n,z}} \circ f_{\omega_{n-1}})(T^{n-1}x) h(x) dx d\mathbb{G}^{\mathbb{N}}(\underline{\omega}),$$

where the $C_{n,z}$ have the same meaning as in section 1. Since the $\{\omega_k\}_{k=0\cdots,\infty}$ are independent and performing first the integration with respect to $\mathbb{G}^{\mathbb{N}}$, we have

$$\mathbb{P}(M_n \leq u_n) = \int U_n(x) U(Tx) \cdots U_n(T^{n-1}x) h(x) dx,$$

where

$$U_n(x) := \int (\mathbf{1}_{C_{n,z}} \circ f_\omega)(x) d\mathbb{G}(\omega).$$

For instance, if we keep an initial f with value in $\mathbb{R}$ and add to it a random term η with uniform distribution in $[-a, a]$ we have

$$U_n(x) = \frac{1}{2a} \int_{-a}^{a} \mathbf{1}_{C_{n,z}}(f(x) + \eta) d\eta = \frac{1}{2a} \mathrm{Leb}\{[-a, a] \cap [C_{n,z} - f(x)]\}.$$

Another choice is to add to an unperturbed observable f two quantities η_1, η_2 taken with respective probabilities p_1, p_2. In this case we have

$$U_n(x) = \mathbf{1}_{f^{-1}[C_{n,z} - \eta_1]}(x) p_1 + \mathbf{1}_{f^{-1}[C_{n,z} - \eta_2]}(x) p_2.$$

Then

$$\mathbb{P}(M_n \leq u_n) = \int \tilde{P}_n^n(h)(x) dx,$$

where $\tilde{P}_n(h) := P(U_n h)$.

We are now in position to apply the spectral theory since we have just constructed a REPFO system: we leave the details to the reader in order to check the necessary requirements. What is important for us now, is to give an expression for the extremal index and for the boundary level u_n, which will reflect on the dimension of the image of the observable. Let us begin with the extremal index. The quantities $q_{k,n}$ are now defined as [44, 45]:

$$q_{k,n} = \frac{\int (P - \tilde{P}_n)\tilde{P}_n^k (P - \tilde{P}_n)(h)dx}{\int (P - \tilde{P}_n)(h)dx}.$$

By posing

$$V_n(x) := \int (\mathbf{1}_{B_{n,z}} \circ f_\omega)(x) d\mathbb{G}(\omega),$$

we immediately have

$$q_{k,n} = \frac{\int V_n(T^{k+1}(x))U_n(T^k(x)) \cdots U_n(T(x))V_n(x)d\mu}{\int V_n d\mu},$$

which allows us to construct the EI θ.
As in the previous section, we now give the computation of the EI in two situations, with continuous and discrete noise.

- *Continuous noise*

  We put ourselves in the setting of Proposition 2 plus other assumptions which we will add during the proof. Let us consider the additive noise $f(x)+\eta$ described above with η much smaller than 1. The first and the last terms in the integral in the numerator of the $q_{k,n}$ are:

$$\frac{1}{2a}\int_{-a}^{a} \mathbf{1}_{B_{n,z}}(f(x)+\eta)d\eta \text{ and } \frac{1}{2a}\int_{-a}^{a} \mathbf{1}_{B_{n,z}}(f(T^{k+1}x)+\eta)d\eta. \tag{49}$$

  In particular both quantities are bounded by

$$\frac{1}{2a}\text{Leb}[(B_{n,z} - f(T^j x)) \cap [-a,a]] \leq \frac{1}{2a}\text{Leb}(B_{n,z}),\ j = 0, k+1,$$

  and therefore the numerator in $q_{k,n}$ is bounded from above by $\frac{1}{4a^2}(\text{Leb}(B_{n,z}))^2$. We now rewrite the denominator as

$$\int \frac{1}{2a}\int_{-a}^{a} \mathbf{1}_{B_{n,z}}(f(x)+\eta)d\eta d\mu = \frac{1}{2a}\int_{-a}^{a} \left[\mu(f^{-1}(B_{n,z}-\eta))\right] d\eta.$$

  We now suppose that the preimages of the set $B_{n,z} - \eta$ are at most L for any η and set $\max|f'| = M_D < \infty$; moreover we suppose that the density h of μ is bounded from below by h_m. Then

$$\mu(f^{-1}(B_{n,z}-\eta)) \geq L h_m M_D^{-1}\text{Leb}(B_{n,z}-\eta)$$

  which implies after integration with respect to η:

$$\int \frac{1}{2a}\int_{-a}^{a} \mathbf{1}_{B_{n,z}}(f(x)+\eta)d\eta d\mu \geq \frac{1}{2}L h_m M_D^{-1}\text{Leb}(B_{n,z}).$$

  If we now divide the numerator with the denominator, we will find the ratio going to zero for $n \to \infty$, which shows that all the $q_{k,n}$ are zero and the EI is one.

- *Discrete noise*

  We give this example again in the setting of Proposition 2. Take the discrete noise with distributions $\{(\eta_1, p_1), (\eta_2, p_2)\}$, and the ball $B_{n,z}$ around the point $f(z)$. Put $B_{n,z,1} = B_{n,z} - \eta_1, B_{n,z,2} = B_{n,z} - \eta_2$ and $\eta_1 < 0 < \eta_2$. If z is a point where f is monotone and we choose n sufficiently large, the sets $B_{n,z,1}, B_{n,z,2}$ will be disjoint and the same for the four sets $f_i^{-1} B_{n,z,j}, i, j = 1, 2$. Call $z_{i,j}$ the point such that $f(z_{i,j}) \in B_{n,z,j}, f(z_{i,j}) + \eta_j = f(z)$. Suppose now that the point $z_{1,1}$ is a fixed point for T, but the remaining points $z_{i,j}$ are not periodic for T. Then the only term which could give a non zero contribution is $q_{0,n}$, which reads

$$q_{0,n} = \frac{\int [\mathbf{1}_{f^{-1}[B_{n,z,1}]}(x)p_1 + \mathbf{1}_{f^{-1}[B_{n,z,2}]}(x)p_2][\mathbf{1}_{f^{-1}[B_{n,z,1}]}(Tx)p_1 + \mathbf{1}_{f^{-1}[B_{n,z,2}]}(Tx)p_2]h(x)dx}{\int [\mathbf{1}_{f^{-1}[B_{n,z,1}]}(x)p_1 + \mathbf{1}_{f^{-1}[B_{n,z,2}]}(x)p_2]h(x)dx}.$$

  When n goes to infinity only the term

$$p_1^2 \mu(f_1^{-1}(B_{n,z,1}) \cap T^{-1} f_1^{-1}(B_{n,z,1}))$$

  gives a non zero contribution. By using the same distortion arguments as in the proof of Proposition 2 we immediately get

$$q_0 = \frac{p_1}{|T'(z_{1,1})|} \left[1 + \sum_{i,j=1,2; i \neq j} \frac{p_j}{p_1} \frac{|f'(z_{1,1})| h(z_{i,j})}{|f'(z_{i,j})| h(z_{1,1})} \right]^{-1}$$

  and the extremal index will be $\theta = 1 - q_0$.

We now discuss the choice of the boundary levels u_n. First it is defined as

$$\mathbb{P}\left((x, \omega); -\log |f_\omega(x) - f(z)| > u_n\right) \to \tau/n.$$

By introducing the image measures

$$\mu_\omega^* := f_\omega^* \mu,$$

we finally have

$$\int \mu_\omega^*(B(f(z), e^{-u_n}) d\mathbb{G}(\omega) \to \tau/n.$$

It is interesting to explore whether we have a scaling of type

$$\mu_\omega^*(B(f(z), r) \approx r^{d^*},$$

and finally

$$\int \mu_\omega^*(B(f(z), r) d\mathbb{G}(\omega) \approx \int r^{d^*} d\mathbb{G}(\omega)$$

for some exponent d^*. Notice that contrarily to formula (24) we are now transporting the measure μ with some f_ω and computing this measure around the image of a point with a different f.

A simple trick allows us to restore the right framework and a quite general example will suggest some expected behavior. Take a countable family of prevalent observable indexed by $f_j, j = 1, \ldots, \infty$ each with a weight p_j such that $\sum_j p_j = 1$. This discrete measure is called $\mathbb{G}$. Fix one f and suppose that the range of each f_j contains $f(x)$; set $\mu_j^* = f_j^* \mu$. Then

$$\mu_j^*(B(f(x), r)) = \mu\left(f_j^{-1}(B(f_j(f_j^{-1}(f(x))), r))\right).$$

Call x_j one of the pre-images of $f(x)$ by f_j, $x_j \in f_j^{-1}(f(x))$. Then

$$\mu_j^*(B(f(x), r)) = \mu_j^*(B(f_j(x_j), r)).$$

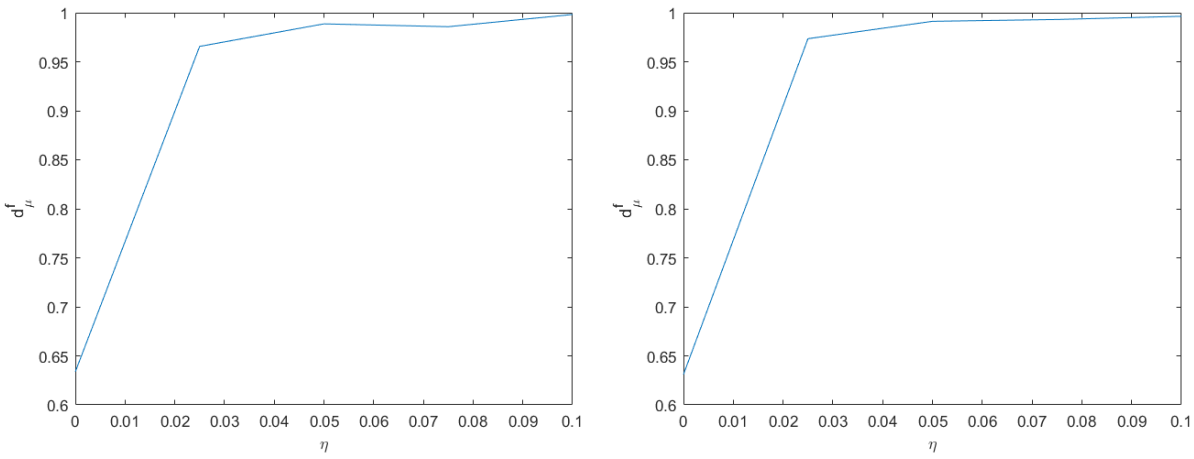


FIGURE 7. Influence of uniform noise of different intensities applied to the dynamics (left) and to the observable (right), for the motion on the product of Cantor sets and the non-prevalent observable $f(x, y) = x - y$, with $f_0 = f(0, 0) = 0$.

Each f_j is prevalent so by Theorem 1 we know that the quantity $\underline{d}_\mu^{f_j}(z)$ is equal to the minimum between the dimension of the range of f, which we take equal to m, and the lower point-wise dimension of μ at z, provided the latter is chosen μ-a.e. If we suppose that the point x_j is typical for f_j and also that μ is exact dimensional, we have that

$$\mu_j^*(B(f(x), r) \approx r^{d_\mu^{f_j}(x_j)}, \text{with } d_\mu^{f_j}(x_j) = \min(m, D_1(\mu)).$$

In conclusion

$$\int \mu_\omega^*(B(f(z), r) d\mathbb{G}(\omega) \approx r^{(\min(m, D_1(\mu)))}.$$

Therefore for scalar functions ($m = 1$) and attractors of high dimensionality, we expect to get a dimension equal to 1 when the observable is perturbed. Instead if the attractor, or repeller, have dimension less than m, the dimension of the image measure will jump to m.

We studied the effect of uniform additive noise of different intensities to the observable $f(x, y) = x - y$ for the dynamics on the product of two Cantor sets described earlier. At each iteration, we computed $f_k = f(T^k(x, y)) + \varepsilon_k$, where ε_k are i.i.d. random variables drawn with a uniform distribution in $[-\eta, \eta]$. Results are shown in figure 7. Similarly to the case where the dynamics is perturbed, we observe a convergence of d_μ^f to 1 as the intensity of noise η increases. We stress that this monotonic convergence to 1 depicted in the figure is a numerical artefact, since the image dimension becomes immediately 1 as soon as the noise is switched on. To compute d_μ^f, we simulated trajectories of 10^7 points and considered blocks of 10^3 points.

7. Open systems

In the paper [36] we considered the extreme value distribution for open systems, namely for systems with holes, where the orbits enter and disappear forever. That was motivated by the statistical description of phenomena where a perishable dynamics is approaching a fixed target state, but at the same time it deviates to another location where it is captured or vanishes. It is useful to extend that theory in presence of observables. A

close look at the proofs in the aforementioned paper, shows that such proofs can be easily translated to our present situations. One of the major results in [36] was to relate the extremal index to the escape rate (from the hole). That was achieved when the target set was chosen around periodic points. In presence of observables, periodicity is much more cumbersome, as we described in Proposition 2; it would be therefore interesting to have a version of such a proposition in the presence of holes. Before doing that we recall the main result in [36].

Proposition 4. [36] *Let T be a uniformly expanding map of the interval I preserving a mixing measure. Let us fix a small absorbing region, a hole $H \subset I$; then there is an absolutely continuous conditionally invariant measure ν, supported on $X_0 = I \setminus H$ with density h_0. Write $\alpha = \nu(T^{-1}X_0)$. If the hole is small enough there is a probability measure μ_0 supported on the surviving set X_∞ such that the measure $\Lambda = h_0\mu_0$ is T-invariant; we assume that h_0 is bounded away from zero. Having fixed the positive number τ, we take the sequence u_n satisfying $n\Lambda(B(z, \exp(-u_n))) = \tau$, where $z \in X_\infty$. Then, we take the sequence of conditional probability measures $\mathbb{P}_n(A) = \frac{\nu(A\cap X_{n-1})}{\nu(X_{n-1})}$, for $A \subset I$ measurable, and define the random variable $M_n(x) := \max\{\phi(x), \dots, \phi(T^{n-1}x)\}$, where $\phi(x) = -\log|x - z|$. Moreover we suppose that all the iterates $T^n, n \geq 1$ are continuous at z and also that h_0 is continuous at z when the latter is a periodic point. Then we have:*

- *If z is not a periodic point:*

$$\mathbb{P}_n(M_n \leq u_n) \to e^{-\tau}.$$

- *If z is a periodic point of minimal period p, then*

$$\mathbb{P}_n(M_n \leq u_n) \to e^{-\tau\theta},$$

where the extremal index θ is given by:

$$\theta = 1 - \frac{1}{\alpha^p|(T^p)'|(z)}.$$

We remind that a probability measure ν which is absolutely continuous with respect to Lebesgue is called *a conditionally invariant probability measure* if it satisfies for any Borel set $A \subset I$ and for all $n > 0$ that

$$\nu(T^{-n}A \cap X_n) = \nu(A)\ \nu(X_n). \tag{50}$$

The *surviving set* is defined as $X_\infty = \bigcap_{n=1}^{\infty} X_n$, where $X_n = \bigcap_{i=0}^{n} T^{-i}X_0$ is the set of points that have not yet fallen into the hole at time n. Finally the *escape rate* η for our open system is usually defined as $\eta = -\log\alpha$.

Let us return to the proof of Proposition 2 trying to adapt it. The class of maps are the same as those in Proposition 4. The main change will concern the invariant measure which is now the singular measure Λ on the surviving (fractal) invariant set X_∞. Such an invariant measure is absolutely continuous with respect to the conformal measure called μ_0 in Proposition 4. This conformal measure plays the role of the Lebesgue measure in the proof of Proposition 2; in the latter we performed a change of variable which produced the terms $|(T^p)'|$, where p was related to the periodicity of the point where we computed the derivative. The conformal measure will give a multiplicative factor α^p. Moreover the density h in Proposition 2 will be now replaced with the density h_0 with respect to the conformal measure μ_0. In conclusion the term q_k in (30) will be now replaced by the

following one, which we call $q_k^{(o)}$ since it refers to open systems

$$q_k^{(o)} = \sum_{w\in A_k} \frac{1}{\alpha^{k+1}|T^{(k+1)}(w)'|} \frac{1}{1+\frac{|f'(w)|}{h_0(w)}\sum_{y\in B_k}\frac{h_0(y)}{|f'(y)|}}. \tag{51}$$

If we want to perform numerical computations, we should know the value of α. We already said that α is related to the size of the hole, in particular one can show that α is the largest eigenvalue of the perturbed transfer operator $\hat{P}g := P(\mathbf{1}_{H^c}g)$, compare with the perturbed operator $\tilde{P}_n$ of section 2. Therefore for small hole one could apply again the spectral technique of [45] and get α as an asymptotic perturbation of 1, the largest eigenvalue of P. It is not therefore surprising that such an expansion will be related to the location of the hole. In particular if the latter is around a point z which is not periodic, α will be equal to 1, instead it will be equal to $1-\frac{1}{|(T^p)'|(z)}$ if z is a periodic point of minimal period p. We point out again that those values hold in the limit of vanishing holes, so that one would get something slightly different for hole with finite size. An interesting case of a large hole is given in the next section.

7.1. **EVT on fractals I.** In this section and in its companion 8.2, we address the following question. Suppose we have a fractal invariant set which is a repeller and whose Lebesgue measure is zero. How could we get a good extreme value theory by using the Lebesgue measure as the underlying probability? In fact almost all the orbits leaving on sets of positive Lebesgue measure tend to escape from the repeller. On the other hand Lebesgue measure is the most accessible measure and repellers are widespread objects, for instance they constitute the basin boundaries between two, or more, basin of attraction, see [51] for applications to climate. The simplest non-trivial repeller is probably the ternary Cantor set, $\mathcal{C}$; in the above terminology, it is the surviving set of the map $T(x) = 3x$-mod1 having taken the hole as the open interval $(1/3, 2/3)$. We point out that other repellers could be generated as the surviving sets in open systems, so that the next considerations could be useful to understand larger class of fractal invariant sets. The first study dealing with the ternary Cantor set in connection with EVT was mostly numerical and it was given in [53]: the authors conjectured the existence of a limiting extreme value law with an EI equal to 1. A rigorous proof appeared recently in the paper [32]; in particular, the authors introduced the observable

$$\phi(x) = \begin{cases} n, & n \in C_n \\ \infty, & \text{otherwise} \end{cases}$$

where C_n is the disjoint union of the 2^n sets in the construction of the Cantor set $\mathcal{C}$.[11] Notice that the function ϕ will have his maximum (infinity) on the Cantor set, otherwise it says how fare we are from it: it is called the Cantor ladder function in [53, 32]. The probability was chosen as the Lebesgue measure Leb on the unit interval. Given $\tau > 0$ and by introducing the sequence of thresholds

$$w_n := \left\lfloor \tau \left(\frac{3}{2}\right)^n \right\rfloor,$$

it was proved in [32] that

$$\lim_{n\to\infty} \text{Leb}\,(M_{w_n} \le n) = e^{-\tau(1-\frac{2}{3})},$$

[11]The Cantor set is given by $\mathcal{C} = \cap_{n\ge 1} C_n$, where the C_n denotes the disjoint union of the 2^n (cylinder) sets obtained by removing the middle third part of each connected component of of C_{n-1}.

where M_n is the process as defined in (2). In this setting, the EI is therefore equal to 1/3. This result is interesting since the limit distribution is obtained with the Lebesgue measure, which allows us to look at the whole Cantor set as a rare event.

We now instead provide a local inspection to the Cantor set by giving the statistics of the hitting time around any point on the repeller. This statistics will be given by a measure which is absolutely continuous with respect to Lebesgue. All this will follow automatically from our Theorem 4 if it would hold for such a big hole like $H = (1/3, 2/3)$. Actually, in that theorem we required the hole to be small to be able to construct the conformal measure μ_0 and its density h_0 with a perturbative argument. In our case, we can do it directly since the map is easy enough. If we set P the Perron-Frobenius operator associated to T and we define the perturbed operator P_0 as $P_0(g) = P(g\mathbf{1}_{H^c})$, where g is a function of bounded variation, we check easily that, having set $\alpha = \frac{2}{3}$:

$$P_0 h_0 = \alpha h_0 \tag{52}$$

$$P_0^* \mu_0 = \alpha \mu_0, \tag{53}$$

where: P_0^* is the dual of P_0; $h_0 = 3/2$ on the unit interval and μ_0 is the balanced measure described below. The absolutely continuous conditionally invariant measure ν will be a measure with density 3/2 on the closed intervals $[0, 1/3], [2/3, 1]$, and 0 on the hole. Using it to construct the absolutely continuous probability $\mathbb{P}_n$ given in Theorem 4, which is numerically accessible, we could place target sets around any point $z \in \mathcal{C}$, as balls of radius e^{-u_n}, where the thresholds u_n can be chosen as $\frac{3}{2} n \mu_0(B(z, e^{-u_n})) = \tau$. Therefore we get convergence to Gumbel's law for our process M_n (2) with:
- the EI is equal to 1 if the point z is not periodic, thus partially supporting the conclusions of [53].
- if we choose z as a periodic point (they are dense in $\mathcal{C}$), of minimal period p, we get for the EI θ :

$$\theta = 1 - \frac{1}{2^p}.$$

The global [32] and our local approaches to the EVT distribution just described, considered the Cantor set as the non-wandering set of a dynamical system defined on the unit interval. One could consider the dynamics defined directly on the Cantor sets, which means to study the system $(\mathcal{C}, T_{|\mathcal{C}})$. We took this point of view in [26], where the transfer operator was defined directly on $\mathcal{C}$ with the potential $|T'|^{-d_H}$, where d_H was the Hausdorff dimension of the Cantor set ($d_H = \frac{\log 2}{\log 3}$ for the ternary Cantor set). It turns out that the invariant (Gibbs) measure for that potential is exactly the measure μ_0 introduced above [12]. But there was another reason for having chosen such a potential; in fact the conformality of this measure implies that for any measurable set A where T is one-to-one, we have $\mu_0(TA) = 2\mu_0(A)$. Therefore that measure gives 2^{-n} masses to the 2^n intervals C_n of length 3^{-n} at the n-th generation in the construction of $\mathcal{C}$. One could also show that this measure is the weak-limit of the sequence of point masses measures constructed with the pre-images of *each* point in the interval both weighted by 1/2. This is a sort of ergodic theorem for repellers, which makes μ_0 accessible for numerical purposes: it is often called a *balanced* measure. *Using μ_0 as the probability for the EVT distribution directly on the Cantor set, we find Gumbel laws with the same behavior for the EI described above.* We will use again this balanced measure in section 8.2.

[12]Notice that μ_0 is also invariant in the framework of Proposition 4 since it differs from the measure Λ by the constant $f_0 = 3/2$. In this respect Λ is not a probability measure.

8. Hitting time statistics in the neighborhood of sets

8.1. **Smooth sets.** Our approach allows us to compute the hitting time statistics (and the statistics of the number of visits) in shrinking neighborhoods of a C^1 surface $\Gamma \subset \mathbb{R}^k$. At this regard, it is enough to consider an observable $f \in C^1(\mathbb{R}^k, \mathbb{R})$ such that $f(x) = \mathrm{dist}(x, \Gamma)$. In this case, we have for all n the identity

$$\{x \in X, |f(x) - 0| < e^{-u_n}\} = \{x \in X, \mathrm{dist}(x, \Gamma) < e^{-u_n}\}.$$

The hitting time statistics in the target sets $\Gamma_n = \{x \in X, |f(x) - 0| < e^{-u_n}\}$ can be deduced from our theory and is given by the distribution of M_n, which converges to the Gumbel law. We then automatically obtain the hitting time statistics in the sets $\{x \in X, \mathrm{dist}(x, \Gamma) < e^{-u_n}\}$. The parameters of this limit law are often computable explicitly (see section 4 for the computation of d_μ^f).

We now give two examples based on the baker map and on the product of two Cantor sets.

- Let us take Γ a straight line of equation $ax + by + c = 0$. The distance from a point $z = (x, y) \in X$ to Γ is given by

$$\mathrm{dist}(z, \Gamma) = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}}.$$

  Let us take the C^1 observable

$$f(x, y) = \frac{ax + by + c}{\sqrt{a^2 + b^2}},$$

  so that with this choice of observable, we have for all n the identity

$$\{z \in X, \mathrm{dist}(z, \Gamma) < e^{-u_n}\} = \{z \in X, |f(z) - 0| < e^{-u_n}\}. \tag{54}$$

  The hitting times statistics in the set in the right hand side is given for large n by the Gumbel law with scale parameter $1/d_\mu^f$. For the baker's map d_μ^f is 1 if $b \neq 0$ and less than 1 if $b = 0$ (see section 4). The extremal index was computed numerically by two of us in [26] in the neighborhood of the diagonal and we found a value strictly less than 1.
  For the product of the two Cantor sets, we found in section 4 that for straight lines parallel to the coordinate axis and in the neighborhood of the diagonal, d_μ^f was strictly less than one. In [26] we proved analytically that the EI computed around the diagonal was equal to 1/2.
- Let us take now Γ as the circle in $\mathbb{R}^2$ of center (a, b) and of radius R, of equation

$$(x - a)^2 + (y - b)^2 - R^2 = 0.$$

  The distance from a point $z = (x, y)$ to Γ is given by

$$\mathrm{dist}(z, \Gamma) = |\sqrt{(x - a)^2 + (y - b)^2} - R|.$$

  Let us take the C^1 observable

$$f(x, y) = \sqrt{(x - a)^2 + (y - b)^2} - R,$$

  so that with this choice, we have again the equivalence (54).
  As before, the hitting times statistics in the neighborhood of the circle is given for large n by the Gumbel law with scale parameter $1/d_\mu^f$. For the baker's map d_μ^f is D_1 when $R = 0$ as we already showed, and 1 whenever $R > 0$, as it easy to see by adapting the argument given for the double Cantor set in the neighborhood of

the diagonal.
For the product of two Cantor sets we have again $d_\mu^f < 1$ for $R = 0$ and also $d_\mu^f < 1$ for $R > 0$, proving that the observable is not prevalent. In both cases the EI follows the usual dichotomy for $R = 0$. We do not dispose of rigorous results in the other case $R > 0$.

One can generalize this approach to higher dimensional $REPFO$ systems and generic C^1 hypersurfaces.

8.2. **EVT of fractals II.** We could now wonder what happens if we consider the distance with respect to a fractal set, for instance the ternary Cantor set introduced in section 4. An easy way to do it and which uses the ideas of this section, is to consider the ternary Cantor set placed along the y-axis in the Cartesian product studied in section 4.2. If we consider the function $f(x, y) = x$, we are led to study the EVT distribution for the observable (1) with $\phi(x, y) = -\log(|x|)$. As the underlying probability we take the balanced measure $\mu^{(2)}$ introduced in section 4.2 and described in section 7.1. We are now interested in computing the EI. Since the Cantor set on the y-axis is invariant for the product map $T \times T$ on the unit square (the map T was defined in section 4.2), we can adapt the proof "along the diagonal" given by us in [26] section II B or in [22, 38], and find easily that only the term q_0 in the expansion of the extremal index will not vanish. Then we use the conformality of the factor measure μ along the x-axis and we get $q_0 = 1/2$, giving also an extremal index equal to 0.5. It is interesting to compare this result with that in [32] described in section 7.1 and with a global approach to the Cantor set as a rare event: the EV found there was 1/3.

Up to now, EVT has been applied to compute hitting time statistics in the neighborhood of some sets of points [33, 5], some Cantor sets [32, 52], or the diagonal in product spaces [26, 16]. We now provide generalizations to arbitrary C^1 surfaces. We point out that a few results have been obtained in that direction in [18], where for the Arnol'd cat map and a C^1 curve $\Gamma \subset [0, 1]^2$, the asymptotic behavior of the shortest distance of the system to Γ up to a time n was derived for Lebesgue almost every starting point.

9. Large deviations

We pointed out in the Introduction and experienced in the preceding sections, that one the most useful, and used, consequences of the EVT applied to dynamical systems, is the possibility to compute numerically the point-wise (also named local), dimensions of the invariant sets. It turns out that in several time series given by natural phenomena or experimental signals, these local dimensions deviate significantly from each other, while in the ergodic setting they should coincide almost everywhere. Instead of seeing in this behavior only a numerical effect, we attributed it to the presence of large deviations in the convergence to the local dimension. The latter manifest themselves on small, but not negligible, scales, a regime which we called *penultimate* [16]. The presence of large deviations for the point-wise dimensions has been rigorously proved for conformal repellers[13] in the paper [19]. Suppose we have an exact dimensional measure μ, call D_1

[13]These are the invariant sets of uniformly expanding $C^{1+\alpha}$ maps, defined on smooth manifolds and whose derivative is a scalar times an isometry. The repeller arises as the attractor of pre-images of the map, see [8] for an exhaustive description. Dynamically generated Cantor sets on the line, Iterated Function Systems with the open set condition, disconnected hyperbolic Julia sets, are all examples [9] of conformal repellers. It is worth mentioning that such repellers can be coded by a subshift of a finite type

the μ-almost sure limit, and suppose that the following limit exists

$$D_q = \lim_{r\to 0} \frac{\log \int \mu(B(x,r))^{q-1} df_*\mu(x)}{(q-1)\log r}. \tag{55}$$

for all $q \in \mathbb{R}$ and moreover the function $\tau(q) = D_q(q-1)$ is C^1 over $\mathbb{R}$ and strictly convex.[14] This is what happens for conformal repellers, where the limit (55) exists for real q [57]. Then we are in the setting of the large deviation result by Gardner-Ellis, see for instance [20], which allows us to state for all interval I:

$$\lim_{r\to 0} \frac{1}{\log r} \log \mu \left(\left\{ z \in X \text{ s.t. } \frac{\log \mu(B(z,r))}{\log r} \in I \right\} \right) = \inf_{s\in I} Q(s). \tag{56}$$

The *rate function* $Q(s)$ is determined by the D_q:

$$Q(s) = \sup_{q\in\mathbb{R}} \{-qs + qD_{q+1}\}. \tag{57}$$

Remark 4. *We notice that when the limit (55) exists in some interval of values of q, then we have to restrict the interval I to a suitable neighborhood $\tilde{I}$ of the information dimension D_1, and for $s \in \tilde{I}$ we can control only deviations larger than D_1, namely we have*

$$\lim_{r\to 0} \frac{1}{\log r} \log \mu \left(\left\{ z \in X \ s.t. \ \frac{\log \mu(B(z,r))}{\log r} > D_1 + s \right\} \right) = \inf_{s\in\tilde{I}} Q(D_1 + s). \tag{58}$$

see [39], *Lemma XIII.2, for the details.*

It is interesting to ask whether large deviations are present when an observable is applied to the measure. Let us start by defining the generalized dimension of order q of $f_*\mu$ (if it exists) as:

$$D_q^f = \lim_{r\to 0} \frac{\log \int f_*\mu(B(x,r))^{q-1} df_*\mu(x)}{(q-1)\log r}. \tag{59}$$

Suppose now that the image measure $f_*\mu$ is exact dimensional; if the function D_q^f exists and is differentiable in some interval of values of q and moreover it is there strictly convex, we have a large deviation principle like (56), eventually slightly modified as in (58). Actually, Remark 4 becomes particularly pertinent in view of the next result by Hunt and Kaloshin. They in fact showed that for a prevalent set of C^1 observables $f : \mathbb{R}^n \to \mathbb{R}^m$, and for $1 \le q \le 2$, D_q^f is given by

$$D_q^f = \min(D_q, m). \tag{60}$$

This result implies that when m is smaller than the D_q's, the image measure is not anymore multifractal, in the sense that all the D_q^f are equal to m, for $1 \le q \le 2$. The function $(q-1)D_q^f$ is not strictly convex and therefore no large deviation principle holds for d_μ^f.

On the other hand, if m is larger than the D_q, we have from equation (60) that $D_q^f = D_q$, at least for $1 \le q \le 2$. Therefore, the image measure inherits some part of the generalized dimensions spectrum from the original measure, which could influence the fluctuations of d_μ^f around D_1.

Apart the threshold imposed by m, the observable f will not exhibit itself explicitly in the detection of d_μ^f given by equation (60) in the range $q \in [1, 2]$. One could ask if the

and they support invariant measures which are Gibbs equilibrium states. This makes them particularly suited for the application of the thermodynamic formalism.

[14]For $q = 1$, the value for D_q is obtained by l'Hopital rule.

influence of f will manifest for values of q outside the interval $[1,2]$. Hunt and Kaloshin gave examples of dynamical systems where $D_q \neq D_q^f$ for $q \notin [1,2]$.
We will instead show that in presence of non prevalent observable the image measure will not in general satisfy Eq. (60). In conclusion: *the signature of the observable f could become apparent by affecting the typical value of d_μ^f for large and small q, or when f is not prevalent*[15]. This issue could be important when we analyze time series generated by physical observables, especially if the underlying dynamical systems is high dimensional. We will study a few of those cases in a future publication.

9.1. **Examples.** In the following examples we will mostly consider the baker's map for which we can establish rigorous results. The baker map does not give a conformal repeller, but we could reduce to it by conditioning on the invariant manifolds.

9.1.1. *Vertical linear observable.* We consider the baker map studied in the previous sections. The attractor of this map has a multifractal structure [56, 10].

Let us take the observable $f(x,y) = x$ and consider

$$\int f_*\mu(B(z,r))^{q-1} df_*\mu(z),$$

where $z \in \mathbb{R}$.
First, by definition of image measure we bring the integration over the SRB measure supported on the baker's attractor:

$$\int f_*\mu(B(f(\overline{v}),r))^{q-1} d\mu(\overline{v}),\ \overline{v} = (x,y).$$

We notice that $f_*\mu(B(f(\overline{v}),r))$ is exactly the SRB measure of a vertical strip centered at x and with width r. We now use disintegration and write

$$\Sigma(x) := \{(x,y); |x-y| < r, y \in [0,1]\} = f^{-1}(B(x,r)),$$

and

$$\int f_*\mu(B(x,r))^{q-1} d\mu(\overline{v}) = \int_{\mathcal{F}_s} \int_{W_{s,\nu}} f_*\mu(B(x,r))^{q-1} d\mu_{s,\nu} d\zeta(\nu),$$

where $W_{s,\nu}$ denotes an horizontal stable manifold indexed with ν and ζ is the counting measure over the stable foliation $\mathcal{F}_s$. Since stable manifolds are horizontal segments of length 1 emanating from all but countably many points y on the y-axis, we will, from now on, identify ν with y and the first integral on the right hand side of the expression above will be evaluated between 0 and 1. It has been proved in section 4.2 that

$$f_*\mu(B(x,r)) = \mu_{s,y}(\Sigma(x) \cap W_{s,y}),$$

which is the conditional measure of a ball of radius r around the point with abscissa x. This measure does not depend on y, and also the conditional measures are the same on all stable fibers and, as we said in section 4.1, they are the invariant (*balanced*), measure of a one-dimensional conformal repeller with two linear branches of slopes λ_1^{-1} and λ_2^{-1} and weights α and $1-\alpha$. In conclusion

$$D_q^f = \lim_{r\to 0} \frac{\log \int f_*\mu(B(x,r))^{q-1} df_*\mu(x)}{(q-1)\log r} = \lim_{r\to 0} \frac{\log \int_0^1 \mu_{s,y}(\Sigma(x) \cap W_{s,y})^{q-1} d\mu_{s,y}}{(q-1)\log r}$$

[15]We remind however that the observable manifests itself in the computation and in the detection of the extremal index, as we showed in formula (30).

Therefore the generalized dimensions spectrum D_q^f of the image measure will be that of the associated $1-D$ IFS which are the solution of the transcendental equation [56, 10]:

$$\alpha^q \lambda_a^{(1-q)D_q^f} + (1-\alpha)^q \lambda_b^{(1-q)D_q^f} = 1,$$

which differs from the generalized dimensions D_q of the baker attractor (in fact we have $D_q = 1 + D_q^f$) [56]. Therefore, this observable does not belong to the prevalent set of the Hunt-Kaloshin theorem, but we already proved that the observable $f(x,y) = x$ is not prevalent.

9.1.2. *Horizontal and oblique linear observable.* We take now first the observable $f(x,y) = y$ (which is prevalent for d_μ^f) and we disintegrate along the unstable manifolds $W_{u,\iota}$ (see section 4.1), where the index ι characterizes the uncountable family of unstable leaves which foliate baker's attractor. In this case we move up a horizontal strip of width r: $\Sigma(y) := \{(x,y); |x-y| < r, x \in [0,1]\} = f^{-1}(B(f(y),r))$. This strip has a measure which is independent of its height and of the unstable leaf $W_{u,\iota}$; it is therefore given by $2r$ (the vertical thickness) times 1 which is the full balanced measure along the x-axis. Remember also that each unstable manifold carries a normalized Lebesgue measure Leb. Therefore we have

$$D_q^f = \lim_{r\to 0} \frac{\log \int f_*\mu(B(x,r))^{q-1} df_*\mu(x)}{(q-1)\log r} = \lim_{r\to 0} \frac{\log \int \mathrm{Leb}(\Sigma(y) \cap W_{u,\iota})^{q-1} d\mathrm{Leb}(y)}{(q-1)\log r} =$$

$$\lim_{r\to 0} \frac{\log \int (2r)^{q-1} d\mathrm{Leb}(y)}{(q-1)\log r} = 1,$$

which shows that all the generalized dimensions for the image measure are equal to 1, and *this proof works for any* q. The same proof immediately generalizes to linear scalar observables of the form $f(x,y) = ax + by + c$, with $b \neq 0$ and it will give that the $D_q^f = 1, \forall q$, just establishing that there are no deviations from the typical value $D_1 = 1$.

9.1.3. *Numerical verification.* We computed numerically the D_q^f of the baker map for the linear observables introduced at the end of the previous section, using the EVT based method developed in [16]. For different values of a and $b \neq 0$, we found a spectrum of generalized dimensions very close to 1 up to $q = 5$ (the discrepancy of the method for high q yields imprecise results for $q > 5$). In conclusion, we believe that for a prevalent set of smooth scalar observable (or more generally for those for which the dimensionality m is smaller than the generalized dimensions of the system), D_q^f will be 1 (or m) for all q. We successfully tested this matter numerically for different C^1 scalar observables. Similar results are found for the Hénon system and for a multifractal Sierpinski gasket that we constructed with the iterated functions system technique presented in [16]. We took the probabilities $p_1 = p_2 = 1/4$ and $p_3 = 1/2$. The generalized dimensions for this system are explicit (see [16]) and comprised between 1 and 2. When we take an observable in $\mathbb{R}^2$, we find a perfect agreement between D_q^f and D_q (see figure 8) for q ranging from 2 to 5. This is a sign that the result of Hunt-Kaloshin may hold for a large class of systems in a much broader range for q than the interval $[1,2]$. We proceeded our computations using the EVT based method, as for the baker map. We took trajectories of length 10^8 and blocks of size 10^4.

10. Applications

Extreme Value Theory for dynamical systems has been a promising framework for devising metrics to study the climate system. Several heuristic studies [27, 16, 24, 14] have focused on the applicability to climate variables. Although those studies did not

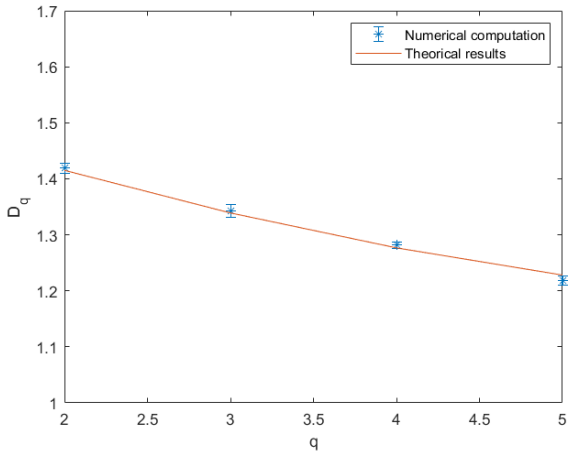


FIGURE 8. D_q^f computed for a Sierpinski gasket and the observable $f(x, y) = (0.2x + 2y, x^2)$, compared with theoretical values of D_q.

have an immediate mathematical justification, they provided insights on the multifractal, non-stationary nature of the climate attractor. Here we give an a posteriori justification of those results, showing that the EVT can be applied to a wide range of observables. It is worth mentioning that in the physical applications, instead of looking at an observable defined on the phase space, we will follow it in time,which would be equivalent by assuming ergodicity of the transformations.

10.1. **From scalar to vector-valued observables.** A common approach to compute the dimensions of the attractor is to use embedding techniques. It consists in taking a C^2 scalar observation $\alpha(x)$, that is accessible through measurement, and constructing a delay observable $f_k(x) = (\alpha(x), \alpha(T(x)), \ldots, \alpha(T^k(x)))$ (a lag parameter is sometimes added in the numerical studies) [63]. Takens' Theorem states that when the map T is a C^2 diffeomorphism defined on a smooth manifold of dimension D and α is a C^2 function, the observable f_k is *generically* an embedding into $\mathbb{R}^{2D+1}$. The notion of (topological) genericity echoes with the notion of prevalence used in the Hunt-Kaloshin Theorem. In fact Takens' Theorem has been strengthened in [63] just by using prevalent observables defined directly on a compact invariant subset of some $\mathbb{R}^l$ and where D is now the box counting dimension of that compact set. As f_k is a diffeormorphism, it preserves the fine structure of the attractor, it allows to reconstruct it and its dimension can be computed numerically. The results of Hunt-Kaloshin, however, state that it is enough that $k > D_1$ to have access to the local dimensions of the attractor, provided the delay coordinate observable f_k is prevalent (which is surely the case for a dense set of C^1 observables α). Therefore the Hunt-Kaloshin result provides a more efficient way to access the dimension of the attractor, but it is surely not enough, in general, to reconstruct it.

Remark 5. *It is very important to point out that in order to get the dimension we* **do not need to reconstruct the attractor***, since the dimension of the image measure is provided directly by the extreme value technique as one of the parameter in the numerical detection of Gumbel's law.*

It is well known that embedding techniques often work efficiently for a number of delay coordinates k that is much smaller than the theoretical value of $2D$ prescribed by the Takens theorem, at least when it comes to the computations of the attractor dimensions [63]. For the Lorenz system for example, it is enough that $k = 3$ [64]. This result is

particularly well understood with the Hunt-Kaloshin results, for which it suffices to have $k > D_1$ to have that $d_\mu^f = D_1$ almost everywhere. To illustrate this fact let us begin with a more general consideration. Let f be a C^1 scalar function defined on a neighborhood $\mathcal{U}$ of our attractor and let the map T be smooth enough in order to apply Theorem 1. We then define the vector-valued function, with values in $\mathbb{R}^k$ and components $(f_1, \ldots, f_k)$:

$$f_j(x) = f(T^{j-1}x), \ x \in \mathcal{U}, \ j = 1, \ldots k. \tag{61}$$

We are now in position to apply the Hunt-Kaloshin Theorem and to look at the least embedding dimension. We tested it on the Lorenz map introduced in section 4.3, reconstructed with the Euler method with step $h = 0.01$. With this iterative procedure, the system could be seen as a discrete mapping, to which we apply our theory.

In figure 9, we computed the local dimensions associated with a vector-valued function like Eq. (61) with f defined by the projection on the x axis, $f(x, y, z) = x$. For the computations, we generated trajectories of $4 \cdot 10^7$ points and took a block size of 10^4. The results are averaged over 20 different trajectories and target points. We find indeed that $d_\mu^f = k$ for $k < D_1 \approx 2.04$ and becomes constant equal to D_1 for $k > D_1$. This suggests that the EVT based methods to compute local dimensions are suited to determine the dimensions of the attractor from a scalar observation f: it is enough to construct the delay-coordinate observable f_k and compute its associated dimension $d_\mu^{f_k}$ for different values of k, until they do not vary anymore as we increase k or until a non-integer value is obtained. We have then attained the dimension of the attractor D_1.

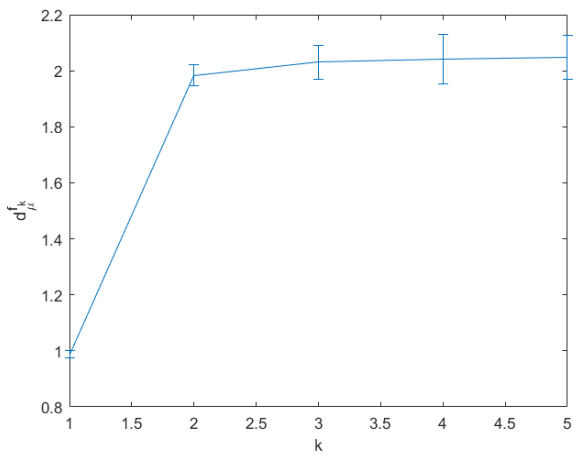


FIGURE 9. Value of $d_\mu^{f_k}$ found for different values of k, for the Lorenz system and the scalar observable $f(x, y, z) = x$. The parameters used are described in the text. The error bars are the standard deviations of the results.

10.2. **Vector-valued observables.** At the end of the previous section we showed how to construct a vector-valued observable by composing a given scalar function with the dynamics. This is the spirit of the embedding approach. It turns out that this procedure has its limits, for instance if the map T is not regular enough to construct prevalent delayed-coordinate observables. Sometimes a variety of observables are available and could be used to compute the local dimensions of the original system, provided their cardinality, say L is large enough. Physicists can measure various quantities associated

to the system (temperatures, pressures, velocities, positions...), which could be arranged as the outcomes of a function with values in $\mathbb{R}^L$. It is enough that L is larger than the information dimension of the system to be able to compute the latter. For example, in order to study the dynamics of the atmospheric circulation over the North Atlantic, several authors [27, 16, 24, 14] have considered an observable that is a vector containing the values of the sea-level atmospheric pressure on a grid of $\approx 10^3$ locations over the North Atlantic. These data were analyzed by [27] and [16]. The source of variability of the sea-level pressure (SLP) atmospheric pressure data has been related to the properties of the atmospheric circulation, namely the switching between different weather regimes [24], the occurrence of extreme weather events [14] and the non-stationarity of the underlying attractor due to climate change [27]. The local dimensions computed from these observations were centered around the value 13, which is much smaller than the dimension of the space where evolves the observable ($\approx 10^3$) [27]. The Hunt-Kaloshin theorem can provide a justification to these results: it is enough that the dimension of the ambient space k where the observable evolves (here $\approx 10^3$) is larger than D_1 (the information dimension of the underlying system) to get that d_μ^f is equal to D_1 almost everywhere. In other words, when we compute some low and non-integer values for d_μ^f, it could be the sign that the information dimension of the underlying system is also not integer and much smaller than that of the ambient space.

To study what happens when the dimension of the observable is smaller than the information dimension of the attractor, we use again the SLP data from the reanalysis of the National Centers for Environmental Prediction (NCEP) [43], but we now investigate the effect of averaging the information over all the grid points. To this purpose, we define the function

$$\phi_Z(X) = -\log|\langle SLP(Z)\rangle - \langle SLP(X)\rangle|,$$

where $\langle SLP(X)\rangle$ is the spatial arithmetic average value of the x_i:

$$\langle SLP(X)\rangle = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

and $Z = (z_1, \ldots, z_m)$ corresponds to a particular configuration of the pressure field.

To compute θ and d_μ^f associated with this observable, we perform a computation of the empirical distribution of the variable M_{50} defined in Eq. (2), for different points $Z \in X$. For each of them, we find that the best fit of the empirical distribution is a Gumbel law of scale parameter close to 1 and estimates of the extremal index are close to 1, like in the baker map situation. The fittings are performed using the Matlab function *gevfit*. For the computation of the extremal index, we used the estimate $\hat{\theta}_0$ introduced in [15], with a threshold value equal to the 0.99-quantile of the observable distribution. Both the values of the extremal index and of $d_{f*\mu}$ that we found have small variability around 1, due to finite effects. In figure 10 the distributions of the values found for θ and $d_{f*\mu}$ over the different points Z of the attractor are represented. These estimates are in perfect agreement with the results presented in this work.

We now study the statistics of the number of visits of the observable in the neighborhood of a particular value for the presented climate data, as we explained in section 5. We can observe in figure 10 on the left, that finite effects lead to an estimate of θ that is slightly smaller than 1, due, among other reasons, to the persistence of the orbits in the neighborhood of the point z. This clustering is very likely to disappear if the amount of data allows to take a higher threshold, and we would observe a pure Poisson distribution

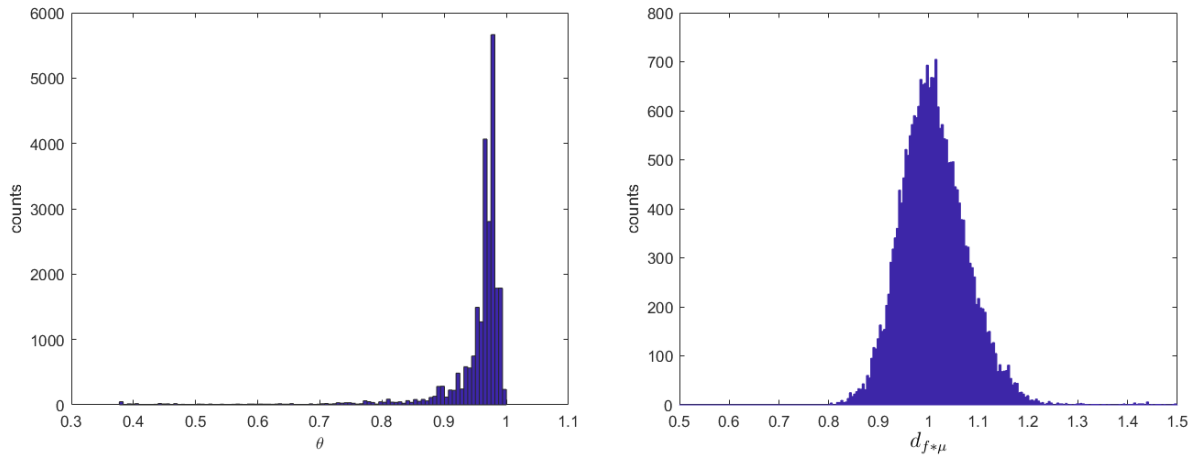


FIGURE 10. Distributions for the atmospheric circulation (SLP) data presented in the text of the EI (left) and of the $d_{f*\mu}(f(z))$ (right) found for different points z of the attractor.

at the limit of high threshold. We get a distribution that is very close to Pòlya-Aeppli of parameters t and the extremal index computed at a finite resolution. This is consistent with the discussion on the figure 5: in many situations, although we know we do not have a Pòlya-Aeppli distribution, it seems that it still models the limit law quite well. In our computations, we studied visits of the observable $f(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$ in the ball $(f_0 - r, f_0 + r)$, where f_0 is the value taken by the observable on July, 1st, 1948 and $r = e^{-u}$, u being the 0.98− quantile of the distribution of ϕ.

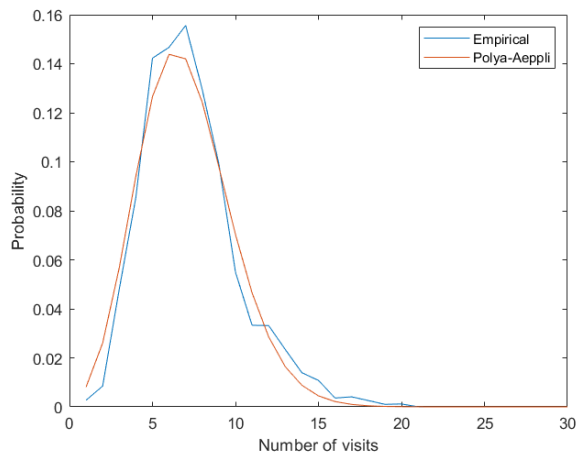


FIGURE 11. Comparison between the empirical distributions of the number of visits of the observable mean value in a ball centered at $f(z)$ and a Polya-Aeppli distribution for the climate data presented in the text.

The existence of extreme value laws for recurrences of physical observables also justifies the results obtained in [28, 29, 2, 35, 59], where the rate of statistical convergence to the extreme value laws was used to estimate the characteristic recurrence time of temperature values (termed recurrence spectra in [2]). They show that despite the slow convergence of

the dynamical systems metrics towards unknown asymptotic values, their distribution is reminiscent of an underlying high-dimensional attractor. On this object, the recurrences around high dimensional fixed or periodic points determine interesting dynamical behaviors such as switching between metastable states [27], critical phenomena [35] or different basin of attraction [12] that can be detected by deviations of the dynamical indicators from their expected asymptotic behavior.

11. Conclusions

This paper contains a few rigorous results illustrated by several examples. The latter are worked out relatively easily, but it was important for us to show that the statistical indicators established by the theory can be explicitly computed and compared with the numerical simulations. The dynamical systems we considered have strong mixing properties, in particular they exhibit exponential decay of correlations on suitable spaces of observables. This allowed us to use a very efficient perturbative theory and compute the extremal index in a broad variety of situations [15]. We believe that our results could be generalised to larger class of systems, even non-uniformly hyperbolic, or exhibiting intermittency [30, 31], using for instance techniques with more probabilistic flavor [33, 34, 1, 17, 50]. In this perspective, we also considered more complex systems and physical time series that we analyzed numerically and that can be tested and interpreted in the framework of our theory. As we explicitly shown in the last section, we believe that our results are useful for physical and natural systems, in the sense that they provide a formal framework for the applications presented in [27, 23, 12, 59, 35]. They partially answer the concern raised in [13] about the slow convergence of dynamical system metrics for climate data and make useful asymptotic theorems for finite data sets.

12. Acknowledgements

S.V. thanks the Laboratoire International Associé LIA LYSM, the INdAM (Italy), the UMI-CNRS 3483, Laboratoire Fibonacci (Pisa) where this work has been completed under a CNRS delegation, the *Centro de Giorgi* in Pisa and UniCredit Bank RD group for financial support through the Dynamics and Information Theory Institute at the Scuola Normale Superiore. SV thanks N. Haydn for enlightening discussions on this paper. The authors thank A. Coutinho and W. Bahsoun for useful exchanges on a few part of this article.

References

[1] H. Aytac, J.M. Freitas, S. Vaienti, Laws of rare events for deterministic and random dynamical systems, Trans. Amer Math. Soc., 367 (2015), 8229-8278.

[2] M.C., Alvarez-Castro, D. Faranda, T. Noël, P. Yiou. Recurrence Spectra of European Temperature in Historical Climate Simulations. Atmosphere, 10-166, 2019.

[3] J. Alves, V. Araujo, Random perturbations of nonuniformly expanding maps, Astérisque, 286, 25-62, 2003.

[4] J. Atnip, N. Haydn, S. Vaienti, Extreme Value Theory with Spectral Techniques: application to a simple attractor, submitted, https://arxiv.org/pdf/2002.10863.pdf.

[5] D. Azevedo, A-C. Moreira Freitas, J-M. Freitas, F.B. Rodrigues, Extreme Value Laws for Dynamical Systems with Countable Extremal Sets, J. Stat. Phys., 167 (2017), no. 5, 12441261.

[6] W. Bahsoun, H. Hu, S. Vaienti, Pseudo-Orbits, Stationary Measures and Metastability, Dynamical Systems: an International Journal, Volume 29, Issue 3, July 2014, pages 322-336.

[7] W. Bahsoun, S. Vaienti, Escape rates formulae and metastability for randomly perturbed maps, Nonlinearity, 26 (2013) 1415–1438.

[8] L. Barreira, Dimension and Recurrence in Hyperbolic Dynamics, in: Progress in Mathematics, vol. 272, Birkäuser Verlag, Basel, 2008.

[9] M.F. Barnsley, S.G. Demko, Iterated function systems and the global construction of fractals, Proc. R. Soc. London A 399 (1985) 243–275.
[10] D. Bessis, G. Paladin, G. Turchetti, S. Vaienti: Generalized dimensions, entropies and Lyapunov exponents from the pressure function for strange sets, J. Stat. Phys., 51, 109-134, (1988).
[11] M. Boshernitzan, Quantitative recurrence results. Invent. Math., 113(1), 617-631, (1993).
[12] M. Brunetti, J. Kasparian, Ch. Vérard, Co-existing climate attractors in a coupled aquaplanet, Climate Dynamics 53.9-10 (2019): 6293-6308.
[13] S. Buschow, P. Friederichs: Local dimension and recurrent circulation patterns in long-term climate simulations, Chaos 28(8), 083124 (2018).
[14] R. Caballero, D. Faranda, G. Messori, A dynamical systems approach to studying midlatitude weather extremes, Geophys. Res. Lett. 44(7) (2017) 3346–3354.
[15] T. Caby, D. Faranda, S. Vaienti, P. Yiou, On the computation of the extremal index for time series, to appear in Journal of Statistical Physics,(2019)
[16] T. Caby, D. Faranda, G. Mantica, S. Vaienti, P. Yiou, Generalized dimensions, large deviations and the distribution of rare events, Phys. D: Nonlinear Phenomena, Volume 400, (2019), 132143.
[17] M. Carvalho, A. C. Moreira Freitas, J. Milhazes Freitas, M. Nicol, Extremal dichotomy for uniformly hyperbolic systems (2015), Dynamical Systems 30(4).
[18] M. Carney, M. Holland, M. Nicol, Extremes and extremal indices for level set observables on hyperbolic systems, arXiv:1909.04748
[19] A. Coutinho, J. Rousseau, B. Saussol, Large deviation for return times, Nonlinearity 31 (11) (2018) 5162–5179.
[20] [38] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Springer Verlag, Berlin, (2010).
[21] M. Demers: A Gentle Introduction to Anisotropic Banach Spaces, Chaos, Solitons and Fractals, 16 (2018), 29-42.
[22] D. Faranda, H. Ghoudi, P. Guiraud, S. Vaienti, Extreme Value Theory for synchronization of Coupled Map Lattices, Nonlinearity 31 (2018) 3326–3358.
[23] 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).
[24] Faranda, D., Sato, Y., Messori, G., Moloney, N. R., Yiou, P. (2019). Minimal dynamical systems model of the Northern Hemisphere jet stream via embedding of climate data. Earth System Dynamics, 10(3), 555-567.
[25] D. Faranda, V. Lucarini, G. Turchetti, S. Vaienti, Numerical convergence of the block-maxima approach to the Generalized Extreme Value distribution, Journal of Statistical Physics, 145 1156-1180, (2011)
[26] D. Faranda, S. Vaienti, Correlation dimension and phase space contraction via extreme value theory, Chaos 28 (2018) 041103.
[27] D. Faranda, G. Messori, P. Yiou, Dynamical proxies of North Atlantic predictability and extremes, Sci. rep. 7 (2017) 41278.
[28] D. Faranda and S. Vaienti. A recurrence-based technique for detecting genuine extremes in instrumental temperature records. Geophysical Research Letters, 40:1-5, (2013).
[29] D. Faranda, M.C. Alvarez-Castro, P. Yiou: Robust Return Times of hot and cold days via recurrences and extreme value theory. Climate Dynamics, 1-13, (2016).
[30] A.C.M. Freitas, J.M. Freitas, M. Todd, S. Vaienti, Rare events for the Manneville-Pomeau map, StochasticProcesses and their applications, 126, Issue 11, pp 3463-3479, (2016).
[31] A.C.M. Freitas, J.M. Freitas, S. Vaienti, Extreme value laws for sequences of intermittent maps, Proceedings of the AMS, 146, (2018) 2103-2116.
[32] A.C.M. Freitas, J. Milhazes Freitas, F. B. Rodrigues, J.V Soares, Rare events for Cantor target sets, to appear on Communication in Mathematical Physics, https://arxiv.org/abs/1903.07200
[33] 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.
[34] A.C. M. Freitas, J. M. Freitas, and M. Todd, The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics, Comm. Math. Phys. 321 (2013), no. 2, 483527.
[35] A. Gualandi, J.-P. Avouac, S. Michel, Davide Faranda, The predictable chaos of slow earthquakes, Science Advances 6.27 (2020): eaaz5548.
[36] P. Giulietti, P. Koltai, S. Vaienti, Targets and holes, https://arxiv.org/pdf/1909.09358.pdf

[37] N. Haydn, S. Vaienti,The compound Poisson distribution and return times in dynamical systems. Probab. Theory Related Fields, 144 (3-4), 517–542, (2009).
[38] N. Haydn, S. Vaienti, Limiting entry and return times distributions for arbitrary null sets, Communication in Mathematical Physics, DOI 10.1007/s00220-020-03795-0, (2020), arXiv:1904.08733
[39] H. Hennion, L. Hervé, Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness, Lect. Notes Math., 1766, Springer, (2001).
[40] P.C. Hemmer, The exact invariant density for a cusp-shaped return map, J. Phys. A: Math. Gen. 17 (1984) L247–9.
[41] B. R. Hunt, and V. Kaloshin, How projections affect the dimension spectrum of fractal measures, (1997).
[42] B. R. Hunt, T. Sauer, J. Yorke, Prevalence: a translation-invariant almost every for infinite dimensional spaces, Bull. Am. Math. Soc. 27 217-38, (1992).
[43] E. Kalnay, et al. (1996). The NCEP/NCAR 40-year reanalysis project. Bulletin of the American meteorological Society, 77(3), 437-472.
[44] G. Keller, Rare events, exponential hitting times and extremal indices via spectral perturbation, Dyn. Syst. 27 (2012) 11–27.
[45] G. Keller and C. Liverani, Rare events, escape rates and quasistationarity: some exact formulae, J. Stat. Phys. 135 (2009), 519-534.
[46] M.B. Kennel, R. Brown, and H. DI Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Physical review A 45.6:3403 (1992).
[47] MR Leadbetter, G. Lindgren, and H. Rootzen. Extremes and related properties of random sequences and processes. Springer, New York, 1983.
[48] C. Liverani, V. Maume-Deschamps, Lasota–Yorke maps with holes: conditionally invariant probability measures and invariant probability measures on the survivor set, Annales de l'Institut Henri Poincaré Probability and Statistics, 39 (3), 385-412 (2003).
[49] E. N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci., vol. 20, no 2, (1963), p. 130-141.
[50] Lucarini V, Faranda D, Freitas A M, Freitas J M, Holland M, Kuna T, Nicol M, Todd M and Vaienti S, Extremes and Recurrence in Dynamical Systems (New York: Wiley), 2016.
[51] V. Lucarini, T. Bodei, Global Stability Properties of the Climate: Melancholia States, Invariant Measures, and Phase Transitions, arXiv:1903.08348
[52] G. Mantica, L.Perotti, Extreme value laws for fractal intensity functions in dynamical systems: Minkowski analysis, Journal of Physics A: Mathematical and Theoretical, Volume 49, Number 37 (2016).
[53] G. Mantica, L. Perotti, Extreme value laws for fractal intensity functions in dynamical systems: Minkowski analysis. J. Phys. A, 49(37):374001, 21, (2016).
[54] Mathworks, Generalized extreme value parameter estimates. https://fr.mathworks.com/help/stats/gevfit.html
[55] W. Ott, J. A. Yorke, Prevalence, Bulletin (New Series) of the American Mathematical Society Volume 42, Number 3, Pages 263–290 S 0273-0979(05)01060-8, (2005).
[56] E. Ott Chaos in dynamical systems, Cambridge university press, (1993).
[57] Y. Pesin, H. Weiss, A multifractal analysis of equilibrium measures for conformal expanding maps and moran-like geometric constructions, J. Stat. Phys. 86 (1997) 233–275.
[58] J. Pickands III, Moment convergence of sample extremes, The Annals of Mathematical Statistics, 39(3):881–889, 1968. 71–782.
[59] D. Rodrigues, M. C. Alvarez-Castro, G. Messori, P. Yiou, Y. Robin, D. Faranda, Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c reanalysis, Journal of Climate 31.15 (2018): 6097-6111
[60] J. Rousseau, B. Saussol, Poincaré recurrence for observations, Trans. Amer. Math. Soc. 362 (2010), no. 11, 5845-5859.
[61] F. Takens, Detecting strange attractors in turbulence, D. A. Rand and L.-S. Young (ed.). Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898. Springer-Verlag. pp. 366–381 (1981).
[62] J. Rousseau. Récurrence de Poincaré pour les observations, page 30. Mathématiques [math]. Université de Bretagne occidentale - Brest, 2010.
[63] T. Sauer, J. A. Yorke, M.Casdagli. Embedology. J. Statist. Phys., 65(3-4):579–616, 1991. 1, 2.
[64] T. Sauer, Attractor reconstruction, Scholarpedia.

[65] L. G. Souza França et al., Fractal and Multifractal Properties of Electrographic Recordings of Human Brain Activity: Toward Its Use as a Signal Feature for Machine Learning in Clinical Applications, Front. Physiol., (2018).

[66] F. Takens, Detecting strange attractors in turbulence, In D. A. Rand and L.-S. Young (ed.). Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898. Springer-Verlag. pp. 366–381 (1981)

[67] M. Viana, Stochastic dynamics of deterministic systems Brazillian Math. Colloquium (1997), IMPA.

[68] Wikipedia, Prevalent and shy sets.
`https://en.wikipedia.org/wiki/Prevalent_and_shy_sets`

[69] L. S. Young, Dimension, entropy and Lyapunov exponents, Ergodic Theory and Dynamical Systems, 2 (1), 109-124, (1982).