Extreme Value Theory for dynamical systems, with applications in climate and neuroscience

Theophile Caby

Università degli Studi dell' Insubria

Université de Toulon

Phd in Mathematics

Phd Thesis
Extreme Value Theory for dynamical systems, with applications in climate and neuroscience

---

Date of defense: december 20^{th}, 2019.

Théophile Caby

Supervisors:
Giorgio Mantica, Sandro Vaienti
Jury:
G. Cristadoro, A. C. Freitas, J. Freitas, S. Galatolo, G. Mantica, S. Vaienti.

Center for Non-Linear and Complex Systems and Centre de Physique Théorique

Contents

1 Notations and basic definitions

- $\overline{X}$ denotes the complementary set of X
- log denotes the natural logarithm
- {C(x)} denotes the set of points x satisfying the condition $C(x)$
- $\mathbb{1}_X$ denotes the indicator function of the set X
- δ_x denotes the Dirac point measure at the point x
- $\mathbb{L}$ denotes the Lebesgue measure defined on the set that is mentioned
- $\lfloor x \rfloor$ denotes the floor function of x
- $X \otimes Y$ denotes the Cartesian product of the sets X and Y
- $X^{\otimes k}$ denotes the $k-$ary Cartesian power of the set X
- By $\rho - quantile$ of a distribution, for $0 < \rho < 1$ is meant the $100\rho^{th}$ quantile of the distribution
- $\overline{x_i}$ denotes the mean value of the quantities (x_i).

2 Introduction

The progressive discovery of the chaotic behavior of certain dynamical systems has been a major breakthrough in science in the last 150 years. First evidence of chaotic motion came from the description by Poincaré [1], at the end of the 19^{th} century, of the extreme complexity of the trajectories associated with the three body problem. Later, Lorenz discovered a 3-dimensional system defined by a simple set of non-linear differential equations that is very sensitive to initial conditions and whose trajectories tend to evolve on an object of complex geometrical structure [2]. This phenomenon is popularly known as the butterfly effect. Although this discovery seemed to impact only theoretical mechanics, the puzzling behavior of the Fermi Pasta Ulam model posed important questions of physical relevance [3, 13]. These properties, that now constitute the paradigm of chaotic dynamics, had a major impact on popular culture and on the way we think science. It became then clear that the traditional analytical tools used to describe the trajectories of such systems were insufficient and more appropriate probabilistic techniques were required.

The modern theory of dynamical system is then constructed on a set of three objects (M, T, μ). M is called the phase-space, that constitutes the set of all the possible configurations of the system. In many physical applications, such as the motion of some particles subject to some forces, M is taken as a subset of $\mathbb{R}^n$, but in all generality, M can be any abstract manifold, such as a symbolic space, which allows applications in number theory. The law of motion is given by the transformation T, that can be, for example, implicitly defined by a set of differential equations driving the dynamics (as for the Lorenz system), or an explicitly defined continuous flow, a discrete mapping, a shift in the symbolic space and can be either purely deterministic, or have a random component. On top of that, a probability measure μ is associated to the system, which is defined over a $\sigma-$algebra Σ_μ of the manifold M. The measure of a set $A \in \Sigma_\mu$ can acquire a physical interpretation, namely the frequency at which the system visits the set A. To possess such an interpretation, the measure μ must be left *invariant* by the dynamics, meaning that for all subset A of Σ_μ, we have the relation

$$\mu(T^{-1}A) = \mu(A). \tag{1}$$

Proving the existence of some non-trivial invariant measure can be difficult for many transformations, although theoretical results provide their existence in

certain particular cases, such as piecewise expanding maps of the interval [5], automorphisms of the torus [6], and even more complex systems such as the geometric Lorenz flow [7]. Sometimes, the invariant measure can be explicitly given in relation to the Lebesgue measure $\mathbb{L}$ and is absolutely continuous with respect to that matter. In practice, in order to study a dynamical system, we suppose the existence of a unique natural, non-trivial invariant measure, unveiled naturally by the dynamics of the system. As chaotic dynamics often settle on particular sets having complex, fractal geometry referred to as strange attractors, the natural invariant measure associated to the system can have a complex hierarchical structure. To characterize this fine geometrical structure, diverse dimensional quantities have been introduced, that will be described in the thesis.

It is a fundamental result that for *ergodic* systems (a class of systems that includes chaotic systems), the distribution of the values taken by an observable (*i.e.* a real-valued function defined on the phase-space M) along a generic trajectory of the system matches asymptotically the spatial distribution of the observable over the whole phase-space, provided we work with an invariant measure μ. This result, due to Birkhoff [4], allows the use of tools coming from statistics to study the dynamical properties of the system. This has led to the development of the rich and sophisticated ergodic theory of dynamical systems [8, 9]. In the last ten years, it was realized that a well-developed branch of statistics is particularly suited to study certain properties of such chaotic systems: Extreme Value Theory (EVT). At the origin [113], the theory aimed at predicting the occurrence of events of large intensity, such as epidemics, earthquakes or floods [114, 115]. Since the establishment of the connection between the existence of extreme values laws for some particular observables and the existence of hitting time statistics in little balls centered at a target point of phase space, by Freitas, Freitas and Todd [10], the development of the dynamical theory of extreme values has brought a whole new set of tools and powerful results. These tools have been succesfully applied in physics, especially in climate science [85, 105, 106, 72, 62], as new EVT-based methods were designed to compute, from observational time series, meaningful indicators of the dynamics of the underlying system.

Throughout the thesis, we will discuss, improve and provide a conceptual framework in which these methods can be understood. We will also provide new EVT-based methods to compute quantities of interest and introduce new useful indicators associated to the dynamics. An emphasis will be placed on physical

applications and numerical computations, as the use of such methods is developing rapidly. We will start by an introductory chapter to the dynamical theory of extreme events, in which we will describe the principal results of the theory that will be used throughout the thesis. After having introduced some objects that are characteristic of the invariant measure of the system, namely local dimensions and generalized dimensions, we devote the following chapters to the use of EVT to compute such dimensional quantities. One of these methods defines naturally a novel global indicator on the hyperbolic properties of the system. In these chapters, we will present several numerical applications of the methods, both in real world and idealized systems, and study the influence of different kinds of noise on these indicators. We will then investigate a matter of physical importance related to EVT: the statistics of visits in some particular small target subsets of the phase-space, in particular for partly random, noisy systems. The results presented in this section are mostly numerical and conjectural, but reveal some universal behavior of the statistics of visits. The eighth chapter makes the connection between several local quantities associated to the dynamics and computed using a finite amount of data (local dimensions, hitting times, return times) and the generalized dimensions of the system, that are computable by EVT methods. These relations, stated in the language of large deviation theory, have profound physical implications, and constitute a conceptual framework in which the distribution of such computed local quantities can be understood. We then take advantage of these connections to design further methods to compute the generalized dimensions of a system. Finally, in the last part of the thesis, which is more experimental, we extend the dynamical theory of extreme events to more complex observables, which will allow us to study phenomena evolving over long temporal scales. We will consider the example of firing cascades in a model of neural network. Through this example, we will introduce a novel approach to study such complex systems.

Let us now start by describing the dynamical theory of extreme events.

3 Extreme value theory for dynamical systems

We will give in this chapter a description of the principal objects and results of the dynamical theory of extreme events and introduce the estimators that will be used in the numerical applications throughout the thesis. This introductory chapter is largely inspired by the reference book in the domain [85].

3.1 General framework

Let us consider a dynamical system (M, T, μ) that is ergodic, meaning that it fulfills the following definition:

Definition 1 *We say that the invariant measure μ is ergodic for the transformation T if*

$$A \in \Sigma_\mu \text{ and } T^{-1}A = A \implies \mu(A) = 0 \text{ or } \mu(A) = 1. \tag{2}$$

As mentioned in the introduction, the Birkhoff ergodic theorem ensures that the distribution of the values of an observable sampled along a typical trajectory matches asymptotically the spatial distribution of the observable. Let us then take an observable $\phi : M \to \mathbb{R} \cup \{\infty\}$, a generic point $x_0 \in M$ and define the stationary process $X_i(x_0) := \phi(T^i(x_0))$. By accumulating the value taken by the observable over a long enough trajectory, statistical analysis can be performed; one can for example study the distribution of the extreme values of this process. The extreme events that are usually considered are characterized by large values of the observable ϕ and can be studied in two main ways.

The block maxima approach studies the distribution of the maximum value of the observable over blocks of time of large size n, $M_n(x_0) = \max\{X_0, X_1, ..., X_{n-1}\}$, while the peak over threshold approach studies the distribution of exceedances of the observable over a sequence of high thresholds. We will see that for systems with good mixing properties and suitable observables, the two approaches are deeply connected. Let us now describe the first one in detail and denote by

$$F(s) := \mu(x \in M, \phi(x) \leq s), \tag{3}$$

the cumulative distribution function (c.d.f.) of the process (X_i). We start by giving the following definition.

Definition 2 *We say that M_n satisfies an Extreme Value Law (EVL) if for all $s > 0$, there is a sequence of levels $u_n(s)$ and a non-degenerate function (i.e. in this case continuous) $H : \mathbb{R} \to [0, 1]$ that satisfy the two following properties:*

- $n\mu(x \in M, \phi(x) > u_n(s)) \xrightarrow[n\to\infty]{} s$
- $\mu(x \in M, M_n(x) > u_n(s)) \xrightarrow[n\to\infty]{} H(s)$.

Let us focus for the moment on the case of an independent identically distributed (i.i.d.) process (Z_i), with associated probability distribution $\mathbb{P}$. Of course the dynamical process (X_i) as defined above is not independent, as X_i and X_{i+1} are correlated by the deterministic action of T, but the i.i.d case will help us to understand the results of EVT. We will see later how, under certain conditions, one can deal with dependence effects. We have, for the independent case that:

$$\mathbb{P}(M_n \leq u_n(s)) = (1 - \mathbb{P}(Z_0 > u_n(s)))^n \underset{n\to\infty}{\sim} (1 - \frac{s}{n})^n \xrightarrow[n\to\infty]{} e^{-s}. \tag{4}$$

Therefore, if M_n satisfies an EVL, then

$$H(s) = 1 - e^{-s}. \tag{5}$$

Usually, with application in many physical situations, the sequence of levels $(u_n)_n$ is taken to be of the form $u_n = y/a_n + b_n$, where y is a real parameter and $a_n > 0$ for all n, so that $\mathbb{P}(M_n < u_n)$ can be written in the form $\mathbb{P}(a_n(M_n - b_n) < y)$. Suitably re-scaled, its asymptotic distribution can only be of three types: this is the fundamental theorem of Extreme Value Theory.

Theorem 1 *[113] If $X_0, ..., X_n$ are i.i.d. random variables and if there exists linear normalizing sequences (a_n), (b_n), $a_n > 0$, such that*

$$\mathbb{P}(a_n(M_n - b_n) < y) \xrightarrow[n\to\infty]{} G(y), \tag{6}$$

where the function G is non degenerate, then G is necessarily of the form

$$G(y) = GEV_\xi(y) = \begin{cases} e^{-(1+\xi y)^{-1/\xi}}, 1 + \xi y > 0 \text{ if } \xi \neq 0, \\ e^{-e^{-y}} \text{ if } \xi = 0. \end{cases} \tag{7}$$

The function GEV_ξ is called the Generalized Extreme Value distribution of shape parameter ξ. This latter parameter is determined by the tail of the complementary cumulative distribution (c.d.f.)

$$\bar{F} := 1 - F. \tag{8}$$

The value $\xi = 0$ corresponds to an exponential tail and the associated distribution is referred to as of *Gumbel* type. $\xi > 0$ corresponds to a heavy tail and it yields a *Fréchet* distribution. The last type corresponds to an upper bounded $\bar{F}$ and is called a *Weibull* distribution. By upper bounded, we mean that the value taken by the observable on the phase space is superiorly bounded.

It is a remarkable fact that in front of the diversity of possible distributions of the observable, the distribution of the maximum (if it exists) is always one of the above forms. In many applications of EVT, one is interested in determining the values of a_n and b_n for large n, since these quantities, as we will see in the rest of the thesis, can carry precious information on the system. When the asymptotic regime is achieved, meaning that taking larger n does not affect the value of the shape parameter ξ, we can write the equality in the above theorem in a form that is more adapted to numerical investigations:

$$\mathbb{P}(M_n \leq y) \underset{n\to\infty}{\sim} GEV_\xi(\frac{y - \beta(n)}{\sigma}). \tag{9}$$

The distribution of the maximum depends then asymptotically on only three parameters: the shape parameter ξ, the location parameter $\beta(n)$ and the scale parameter σ, that can be estimated by fitting the distribution of M_n for n large enough. Note that in this asymptotic case, the scale parameter σ also does not depend anymore on n. [85]

Complementary to the Block Maxima, the Peak Over Threshold approach for its part studies another characterization of extreme events: the exceedances of the *i.i.d.* process $Z_0, ..., Z_n$ over a high selected threshold u^*. Let us consider the distribution

$$F_{u^*}(y) = \mathbb{P}(Z_0 - u^* \leq y | Z_0 > u^*). \tag{10}$$

The following theorem makes the link between the two approaches:

Theorem 2 *If the conditions of theorem (1) apply and M_n follows a GEV of parameters ξ, σ, $\beta(n)$, then F_{u^*} converges uniformly to the function*

$$GPD_{\xi,\sigma} := 1 + \log(GEV_\xi(y/\sigma)), \tag{11}$$

as u^* tends to $u_F = \sup\{s : F(s) < 1\}$. The reverse is also true.

$GPD_{\xi,\sigma}$ is called the Generalized Pareto Distribution of shape parameter ξ and of scale parameter σ. Note that the location parameter $\beta(n)$ does not appear in this distribution, since the value of threshold is systematically subtracted from the random variable. Now we have two equivalent approaches to study rare events, each of them being more suited to some particular situations: the block maxima approach requires huge amounts of data, since by considering the maximum over large blocks of time, many extreme events are discarded. It has however proved to provide more stable results in many situations [85].

So far, we have stated the results for i.i.d. processes, but deterministic dynamical systems are driven by laws of motion, for which the state of the system at a given time depends on its state at previous time(s), and even for systems with fast decay of correlations, the process $X_0, X_1, ...$ defined earlier is not an independent sequence of random variables, but has a dependence structure. For example, the system may spend some time in a region of phase space where the observable takes high values, leading to a clustering of the large events. To account for this fact, and under some dependence conditions that will be presented, an extreme value law still exists for M_n, provided that an extremal index (EI) θ is introduced. We then have that

$$\lim_{n\to\infty} \mu(x \in M, M_n(x) \leq u_n(s)) = e^{\theta s}. \tag{12}$$

θ is a real number comprised between 0 and 1 which accounts for the eventual memory of the process and quantifies the tendency of the process to form clusters of high values. We will describe more in details this quantity in the following section.

3.2 Conditions of existence of extreme value laws and some formulae for the extremal index

The mere condition of ergodicity is not enough to have an extreme value law for a dynamical process. In fact, the conditions required must also be tied to the chosen observable as well. Indeed, even for systems with strong decay of correlations, if we chose an observable that is constant over the whole phase-space, one cannot expect to have a non-degenerate extreme value law associated to the process. Let us first refer to the classical (statistical) theory of EVT, where several results concerning

the existence of an extreme value law under more or less strong mixing conditions have been found.

3.2.1 Classical conditions

First results concerning the existence of Extreme Value laws for processes with good dependence structure are due to Leadbetter [31]. Later, Chernick Hsing and McCormick [32] proved the existence of an extreme value law under some less strict hypotheses; the condition $D(u_n)$ assuring asymptotic independence and the condition $D^m(u_n)$, controlling the maximal order of clustering, that we are going to define now.

Definition 3 *We say that the condition $D(u_n)$ is satisfied for the process $X_0, X_1, ...$ if for any integers $i_1 < ... < i_p$ and $j_1 < ... < j_k$ for which $j_1 - i_p > t$, and for all $n > n_0 \in \mathbb{N}$,*

$$|F_{i_1,...,i_p,j_1,...,j_k}(u_n) - F_{i_1...,i_p}(u_n)F_{j_1,...,j_k}(u_n)| \leq \alpha(n,t), \tag{13}$$

uniformly for every $p, q \in \mathbb{N}$. Here, $F_{i_1,...,i_p}$ is the joint distribution function of $X_{i_1}, ..., X_{i_p}$, that is

$$F_{i_1,...,i_p}(u_n) = \mu(X_{i_1} < u_n, ..., X_{i_p} < u_n).$$

The function α is such that $\alpha(n, t_n) \to 0$ for some sequence t_n that is asymptotically negligible with respect to n.

Definition 4 *We say that $D^m(u_n)$ holds if there exist sequences $(k_n)_n$ and $(t_n)_n$ such that:*

- $k_n \to \infty$
- $\lim\limits_{n\to\infty} k_n\alpha(n, t_n) = 0$, *where $\alpha(n, t_n)$ is as in definition 3.*
- $k_n t_n$ *is asymptotically negligible with respect to n.*
- $\lim\limits_{n\to\infty} n \sum_{j=m+1}^{\lfloor \frac{n}{k_n} \rfloor} \mu(X_0 > u_n \geq M_{m-1}(x_0) \cap X_{j-1} > u_n) = 0.$

Under these conditions, an extreme value law holds and the extremal index can be computed via what is called the O'Brien formula.

Theorem 3 *(Chernick, Hsing, McCormick [32]) Under the hypothesis $D(u_n)$ and $D^m(u_n)$, the stationary process $X_0, X_1, ...$ has an extreme value law with an extremal index given by*

$$\theta = \lim_{n\to\infty} \frac{\mu(X_0 \geq u_n \cap \max_{i=1,...,m} X_i < u_n)}{\mu(X_0 \geq u_n)}. \tag{14}$$

Freitas and Freitas [33] proposed some conditions that are easier to check in the case when the process originates from dynamical systems. These conditions have been used to obtain different results that we will mention later.

These results are valid for any process with good enough dependence structure, in the sense that the previous conditions are satisfied. In the context of dynamical systems however, one can take advantage of the underlying dynamics to prove the existence of an extreme value law in some more general settings. This can be done using the *spectral* theory developed by Keller and Liverani, which uses perturbation theory of transfer operators associated to mixing dynamical systems [37, 38].

3.2.2 The spectral approach

We will just describe briefly the part of the theory that applies to EVT, in the case of systems with invariant measures that are absolutely continuous with respect to Lebesgue (because so far it has been successfully applied only in these situations), but the theory is originally presented in the broader set up of an arbitrary invariant measure μ. To cover these situations, it is enough to replace the Lebesgue measure $\mathbb{L}$ in definition (5) by another reference measure μ_0 that is compatible with μ. The spectral approach applies to a class of mixing dynamical systems (M, T, μ), whose transfer operator has nice spectral properties (and in particular a spectral gap). Its main objective is to quantify the difference between the eigenvalues of the Perron-Frobenius operator of the system and an operator that is perturbed by the presence of a small hole in phase-space whose measure is shrinking to zero. This approach allows to estimate the escape rate of the system into this hole. The link with EVT is made by considering a family of sets (holes) of the form:

$$\Omega_n = \{x \in M, \phi(x) > u_n\}. \tag{15}$$

With this correspondence, studying the escape rate of the dynamics in the sets Ω_n for large n, in particular the probability that the dynamics has not entered Ω_n after n iterations matches the probability that the maximum value taken by the

observable along a long trajectory of size n is smaller than u_n, which constitutes the core of the block maxima analysis. This correspondence was used by Freitas, Freitas and Todd to show the equivalence between the existence of extreme value laws and hitting time statistics in small balls of shrinking measures [10]. Indeed, with the observable

$$\phi_z(x) = -\log \mathrm{dist}(z, x),$$

(that we will from now call the *standard* observable), the sets Ω_n are balls centered at z of radius e^{-u_n}. This subtile approach allows to benefit from the powerful results of EVT and to obtain a full asymptotic description of the hitting time statistics in these balls, which is described by a Gumbel law. As we will see, this approach can be generalized to arbitrary target sets. Let us now give an overview of the spectral approach.

Definition 5 *The Perron-Frobenius operator P of a dynamical system (M, T, μ) with absolutely continuous invariant measure $\mu = hd\mathbb{L}$ is defined by the following duality integral relation:*

$$\int_M P(f)gd\mathbb{L} = \int_M fg \circ Td\mathbb{L}, \tag{16}$$

where $\mathbb{L}$ is the Lebesgue measure, $f \in L^1$ and $g \in L^\infty$.

If P has a unique isolated eigenvalue equal to 1, the system has a unique mixing absolutely continuous invariant measure, whose density h is the eigenvector associated to the eigenvalue 1 [5]. The perturbed operator $\tilde{P}_n$ is defined by multiplying the argument of P by the indicator function of the complementary set of the hole:

$$\tilde{P}_n(f) = P(f\mathbb{1}_{\overline{\Omega}_n}). \tag{17}$$

The theory is applicable when the operators that we defined have some particular spectral properties that we now list.

Definition 6 *The operators P and $\tilde{P}_n$ are called Rare Event Perron-Frobenius Operators (REPFO) if there exist*

- *a functional space V endowed with a norm $||.||_V$ on M containing the constant functions,*
- *some constants $0 < \alpha < 1$ and $D > 0$,*
- *a second norm $|.|_W$ on V such that $|.|_W \leq ||.||_V$ and*

- *an integer* $N \in \mathbb{N}$

such that for all $n > N$, $\tilde{P}_n$ *satisfies the following properties;*

1. *The residual spectrum of each* $\tilde{P}_n$ *is contained in* $\{z \in \mathbb{C} : |z| \leq \alpha\}$

2. *For all* $f \in V$,
$$|\tilde{P}_n^n f|_w \leq D|f|_w$$
and
$$||\tilde{P}_n^n f||_V \leq D\alpha^n||f||_V + D|f|_w.$$

3. *There is a monotone upper continuous function* $\pi : [N, +\infty) \to [0, +\infty)$ *such that* $\lim_{n\to\infty} \pi(n) = 0$ *and for all* $f \in V$, *for all* $n \geq N$,
$$|\tilde{P}_n f - Pf|_w \leq \pi(n)||f||_V.$$

4. $\mathbb{1}_{\Omega_n}.f \in V$ *for all* $f \in V$.

5. $\mathbb{L}(\mathbb{1}_{\Omega_n} f)||\mathbb{1}_{\Omega_n} h||_V \leq const||f||_V \mu(h\mathbb{1}_{\Omega_n})$ *for all* $f \in V$.

6. *The following mixing condition for the operator holds:*
$$P = h \otimes \mathbb{L} + Q,$$
where Q *has spectral radius smaller than 1.*

If the operators have these good spectral properties, Keller proved, based on a spectral perturbation theory developed together with Liverani [38], that an extreme value law holds for the process $X_0, X_1, ...$ [37] (see theorem (4)). These properties can be difficult to check in practice, but are generally associated to a class of systems with super-polynomial decay of correlations and regular enough observables [99, 37]. With the standard observable, REPFO conditions hold for a class of expanding maps of the interval that have absolutely continuous invariant measures with bounded densities [38, 99]. The functional space V is often taken to be the space of bounded variations functions on M. We will throughout the thesis work with REPFO systems, or at least suppose that further analysis will prove that these properties hold, if rigorous proofs are lacking.

Theorem 4 *(Keller, Liverani [38]) For systems and observables with REPFO properties, an extreme value law holds for the process $X_0, X_1, ...$ Moreover, if the following limits exist:*

$$p_k := \lim_{n\to\infty} \frac{\mu(\Omega_n \cap (T^{-1}\Omega_n)^c \cap ... \cap (T^{-k}\Omega_n)^c \cap T^{-k-1}\Omega_n)}{\mu(\Omega_n)} \tag{18}$$

for $k \geq 1$ and

$$p_0 := \lim_{n\to\infty} \frac{\mu(\Omega_n \cap T^{-1}\Omega_n)}{\mu(\Omega_n)}, \tag{19}$$

then the extremal index is given by:

$$\theta = 1 - \sum_{k=0}^{\infty} p_k. \tag{20}$$

This theorem applies in more general settings than O'Brien result, and in particular it contains in some sense formula (14). To see that, let us define the quantity

$$\theta_m = \lim_{n\to\infty} \frac{\mu(X_0 \geq u_n \cap \max_{i=1,...,m} X_i < u_n)}{\mu(X_0 \geq u_n)}. \tag{21}$$

Proposition 1 *In the general set up of theorem (4), the extremal index is given by*

$$\theta = \lim_{m\to\infty} \theta_m. \tag{22}$$

To prove this proposition, we simply partition the event in the numerator of θ_m and recognize the p_k terms of equations (18) and (19):

$$
\begin{aligned}
\theta_m &= \lim_{n\to\infty} \frac{\mu(\phi(x) \geq u_n \cap \max_{i=1,\dots,m} \phi(T^i x) < u_n)}{\mu(\phi(x) \geq u_n)} \\
&= \lim_{n\to\infty} \frac{\mu(\phi(x) \geq u_n) - \mu(\phi(x) \geq u_n \cap \phi(Tx) \geq u_n) - \mu(\phi(x) \geq u_n \cap \phi(Tx) < u_n \cap \phi(T^2 x) \geq u_n)}{\mu(\phi(x) \geq u_n)} \\
&- \dots - \frac{\mu(\phi(x) \geq u_n \cap \max_{i=1,\dots,m} \phi(T^i x) < u_n \cap \phi(T^{m+1} x) \geq u_n)}{\mu(\phi(x) \geq u_n)} \\
&= 1 - \lim_{n\to\infty} \frac{\mu(\phi(x) \geq u_n \cap \phi(Tx) \geq u_n)}{\mu(\phi(x) \geq u_n)} - \frac{\mu(\phi(x) \geq u_n \cap \phi(Tx) < u_n \cap \phi(T^2 x) \geq u_n)}{\mu(\phi(x) \geq u_n)} \\
&- \dots - \frac{\mu(\phi(x) > u_n \cap \max_{i=1,\dots,m} \phi(T^i x) < u_n \cap \phi(T^{m+1} x) \geq u_n)}{\mu(\phi(x) \geq u_n)} \\
&= 1 - \sum_{k=0}^{m} p_k.
\end{aligned}
\tag{23}
$$

In the set up of theorem (4), the quantity θ_m converges to θ as m goes to infinity. Under the mixing condition $D^m(u_n)$ (4) limiting the characteristic size of the clusters, all the terms p_k for $k > m$ are zero and we have that $\theta_m = \theta$, which allows to recover the O'Brien formula (14). We will present in section (4.4) a situation where $D^m(u_n)$ is not satisfied, but in which the spectral theory can be applied, which makes it in some sense more general. In particular, we will give an example where none of the p_k term is null.

With these formulae, one is sometimes able to compute explicitly the extremal index associated with different observables. With the standard observable, the sets Ω_n considered in equation (15) are simply balls centered at the target point z and of radius e^{-u_n}. For unidimensional uniformly expanding systems preserving a measure absolutely continuous with respect to Lebesgue, it was proven (both using the conditions of Freitas, Freitas [65] and the spectral approach [37]) that an extreme value law associated to the maximum of ϕ_z holds with an extremal index that is equal to 1 everywhere, except for periodic target points z, for which

$$
\theta = 1 - p_{k-1} = 1 - \frac{1}{|DT^k(z)|}. \tag{24}
$$

k denotes here the period of z. In [86], an analogous result was proven for diffeomorphisms of the torus, where in this case, the extremal index is given by:

$$\theta = \frac{2}{\pi}(\arcsin \frac{|\lambda|^k}{\sqrt{|\lambda|^{2k+1}}} - \arcsin \frac{1}{\sqrt{|\lambda|^{2k+1}}}). \tag{25}$$

In the above formula λ denotes the leading eigenvalue of the matrix associated to the diffeomorphism. It is important to stress that this formula is obtained when the Euclidean distance is considered in the observable ϕ_z, while other distances yield different results. We remark that in the previous formulae, only the term p_{k-1} where k is the minimal period of the z is non-zero, and the extremal index contains in this case information on the periodicity of the target point.

These last examples show that the spectral theory can be used for balls centered at target points of M, but we will see in the rest of the thesis that equation (18) applies for more general target sets. In fact, when these target sets have a nice geometrical structure, EVT allows to extract meaningful dynamical information on the system. In section (5.3) we will consider tubular neighborhoods of the diagonal in the product space, while in section (9) we will consider a family of fractal sets of null measure. This approach was initiated in [30] and was put in a rigorous mathematical formulation in [11].

3.3 EVT and the statistics of the number of visits

We mentioned in section (3.1) that θ quantifies the tendency of the process to form clusters of high values. We give now a more precise meaning to this statement and see that it appears in the asymptotic statistics of the number of visits of the system to the set Ω_n. It is known from the classical statistical theory of extreme values that the number of exceedances of a process having an extreme value law, up to a re-scaled time, over a large threshold is given by a compound Poisson distribution [85]. Thanks to relation (15), exceedances over the high threshold u_n correspond in this context to visits of the system in the set Ω_n. Let us describe more in details the objects and the results.

Definition 7 *We say that the integer valued random variable P follows a Poissonian distribution $\mathscr{P}$ if*

$$\mathscr{P}(P = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{26}$$

for some parameter λ.

Definition 8 *We say that the integer valued random variable W is compound Poisson distributed if there exists*

- *Some integer valued i.i.d. random variables* $(X_j)_{j=1,2,...}$ *with the same probability distribution given by* $\pi_l = \mathbb{P}(X_j = l)$*, for* $l \in \mathbb{N}$
- *A Poisson distributed random variable* P

such that

$$W = \sum_{i=1}^{P} X_j. \tag{27}$$

Moreover, we say that a probability measure ν *is compound Poisson distributed with parameters* $t\pi_l$*,* $l = 1, 2, ...$ *if its generating function is of the form*

$$\mathscr{G}_\nu(z) = \exp \sum_{l=1}^{\infty} (z^l - 1) t\pi_l. \tag{28}$$

Let us define the quantity $N_n(t)$, that is the number of times the dynamics enters the set Ω_n up to a re-scaled time $t/\mu(\Omega_n)$:

$$N_n(t) = \sum_{l=1}^{\lfloor \frac{t}{\mu(\Omega_n)} \rfloor} \mathbb{1}_{\Omega_n}(T^l(x)). \tag{29}$$

Knowing the distribution of $N_n(t)$ can be interesting from a physical point of view. A first step to achieve this is to define the *cluster size distribution*:

$$\pi_l := \lim_{K\to\infty} \lim_{n\to\infty} \frac{\mu(\sum_{i=0}^{2K} \mathbb{1}_{\Omega_n} \circ T^i = l)}{\mu(\sum_{i=0}^{2K} \mathbb{1}_{\Omega_n} \circ T^i \geq 1)}. \tag{30}$$

Asymptotically, visits in the target set Ω_n occur by clusters whose sizes are distributed according to π_l. Each of these clusters is produced in an "unpredictable" way, in the sense that the number of these clusters follows a Poisson distribution. This is the principal idea of the following theorem.

Theorem 5 *([22]) For dynamical systems and observables having REPFO properties,*

$$\mu(N_n(t) = k) \underset{n\to\infty}{\to} \nu_t(k), \tag{31}$$

where ν_t *is a compound Poisson distribution of parameters* $t\pi_l$*.*

It also follows from the theory that the expected value of the cluster size is [22]:

$$\sum_{k=1}^{\infty} k\pi_k = \frac{1}{\theta}. \tag{32}$$

These results apply to very general target sets Ω_n, but in the standard case of balls centered at a point z, a theory has been developed in [68, 65], prior to the one in [22]. In the case where z is not a periodic point of the dynamics, we have that $\pi_l = \delta_1$ and ν_t is a pure Poisson distribution, in the sense that

$$\nu_t(k) = \frac{t^k e^{-t}}{k!}. \tag{33}$$

On the other hand, for periodic target points of minimal period m, the parameters of the compound Poisson distribution are geometric in the extremal index θ, in the sense that the cluster size distribution is given by

$$\pi_l = \theta(1-\theta)^l. \tag{34}$$

This yields a geometric Poisson distribution, also called a *Polyà-Aeppli* distribution:

$$\mu(N_n(t) = k) \xrightarrow[n\to\infty]{} e^{-\theta t} \sum_{j=1}^{k} (1-\theta)^{k-j} \theta^j \frac{(\theta t)^j}{j!} \binom{k-1}{j-1}. \tag{35}$$

In section (6.1), we will present numerical evidence that for this type of observable, a Polyà-Aeppli distribution parametrized by the extremal index seems to hold for certain stationary random systems as well. We will also see that in more general settings, distributions other than Polyà-Aeppli are found for the statistics of the number of visits. In particular, it is the case for the tubular neighborhood of the diagonal in the product space considered in section (5.3).

In the next section, we will present numerical estimates that can be used to evaluate the extremal index.

3.4 Numerical estimates of the extremal index

Let us consider a finite observational time series $X_0, ..., X_{n-1}$ which is obtained for example as in section (3.1), by computing the value taken by an observable along a trajectory. Although the extremal index is determined in terms of exceedances over some sequence of high thresholds, in practice we need to fix a single high threshold to define exceedances. Usually, a high quantile of the distribution $0 < \rho < 1$ is taken, assuring at the same time that enough data are at disposal for the statistics. In some situations, estimations of the extremal index at the order 0 (involving only an estimate of $1 - p_0$) suffice. This is the case of systems satisfying the property

$D^2(u_n)$. Under this assumption, the maximum likelihood estimate of Süveges [100] can be used. It is given by:

$$\hat{\theta}_{su}(\rho) = \frac{\sum_{i=1}^{N_c}(1-\rho)S_i + N + N_c - \sqrt{(\sum_{i=1}^{N_c}(1-\rho)S_i + N + N_c)^2 - 8N_c\sum_{i=1}^{N_c}(1-\rho)S_i}}{2\sum_{i=1}^{N_c}(1-\rho)S_i}, \tag{36}$$

where N is the number of exceedances over τ, N_c is the number of clusters with two or more exceedances, and S_i is the size of the cluster i. $\hat{\theta}_{su}$ is based on the point process character of the exceedances and is less biased than other estimators. For this reason, it is used in many applications in climate [105, 72]. It is presented in details in [100] and a *Matlab* code and numerical details related to this computation in the context of dynamical systems are available from [85]. Because it is an order zero estimate, this formula estimates fairly the extremal index for systems with a particular type of clustering: it takes into account only successive exceedances over the threshold. In more general situations, we need to estimate θ at a higher order. For example, we must be able to characterize the periodic points of formulae (24) and (25), for which only the term p_{k-1} is non zero (k being the period of the target point).

As the spectral approach is in some sense more general, we will stick to it as much as possible throughout the thesis, and take advantage of formula (20) to estimate numerically θ. To do so, we start by designing a suitable estimate $\hat{p_j}$ for p_j. We introduced this estimate in details in the article *T. Caby, D. Faranda, S. Vaienti, P. Yiou, On the computation of the extremal index for time series [112].* It is obtained by considering a trajectory $(x_0, Tx_0, ..., T^{n-1}x_0)$ of length N and computing for a high threshold $u(\rho)$ (that corresponds to the $\rho-$quantile of the time series) the quantity:

$$\hat{p}_j(\rho) = \frac{N\sum_{i=0}^{N-2-m}\mathbb{1}(A_i(\rho)\cap\bigcap_{l=1,...,j}\bar{A}_{l+i}(\rho)\cap A_{i+j+1}(\rho))}{(N-1-j)\sum_{i=0}^{N-1}\mathbb{1}(A_i(\rho))}, \tag{37}$$

where we denoted

$$A_i(\rho) = \{\phi(T^i x_0) \geq u(\rho)\}. \tag{38}$$

In the most general case, a good estimate of θ would require to compute an infinity of the above estimates and subtract them all from 1 (as in formula (20)), but this is of course impossible in practice. In fact, we will see that in many situations, either none or only one of p_k is non zero (like in the situations (24) and (25)) and even when it is not the case, we will see that (p_k) often decreases with k very fast to 0 and stopping the computation at a finite order m does provide a reliable estimate for θ. We thus define the estimate of order m:

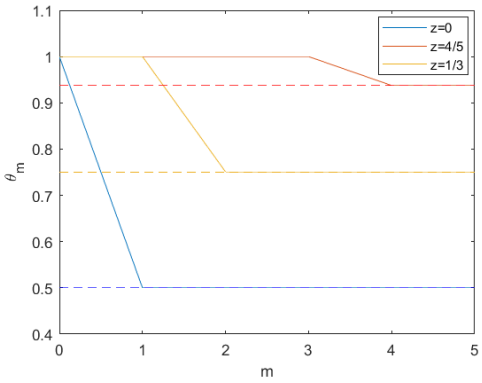
$$\hat{\theta}_m(\rho) := 1 - \hat{p}_0(\rho) - ... - \hat{p}_m(\rho). \tag{39}$$

In the situations considered in equations (24) and (25) and for a $k-$periodic target point z, we expect that all computed $\hat{p}_i$ but $\hat{p}_{k-1}$ are equal to 0, and an estimate of θ at order $k-1$ using $\hat{\theta}_k(\rho)$ should give a fair estimate of the extremal index. In table (4) we compute the extremal index associated to different periodic points of different maps and compare this estimate with the order 0 Süveges estimate.

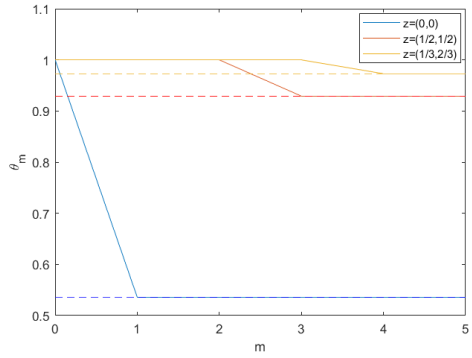
Map T	z	Period	Theor. value eqs. (24),(25)	$\hat{\theta}_{su}$	$\hat{\theta}_5$	Uncert.
$2x$ (mod 1)	$4/5$	4	0.9375	1	0.9375	0.0024
$2x$ (mod 1)	0	1	0.5	0.5007	0.5005	0.0028
$2x$ (mod 1)	$1/3$	2	0.75	1	0.7508	0.0017
$2x$ (mod 1)	$1/\pi$	not periodic	1	1	1	0
Cat map	$(1/3, 2/3)$	4	0.9730	1	0.9732	$8.3.10^{-4}$
Cat map	$(1/2, 1/2)$	3	0.9291	1	0.9292	$9.32.10^{-4}$
Cat map	$(0, 0)$	1	0.5354	0.5352	0.5350	0.0018
Cat map	$(1/\sqrt{2}, \pi - 3)$	not periodic	1	1	1	0

Table 1: Comparison of estimates of θ found by different methods for the $2x$ (mod 1) map and the Arnol'd cat map [93], defined on the torus $\mathbb{T}^2$ by $\Gamma(x, y) = (x+y, x+2y)$. For all of the computations, we averaged the results over 20 trajectories of 5.10^7 points and took as a threshold the 0.999-quantile of the observable distribution. The uncertainty is the standard deviation of the results.

In figure (1), we show the evolution of $\hat{\theta}_m$ with m for different maps and different target points. We observe that for a target point of period k, $\hat{\theta}_m$ is equal to 1 for $m < k$ and switches to the value θ for $m \geq k$, due to the fact that $\hat{p}_{k-1}$ is non zero. For $m \neq k$, we also have $\hat{p}_{m-1} = 0$, as expected. It becomes clear now that the estimate $\hat{\theta}_m$ is particularly suited to treat this kind of situation, as it unveils the presence of periodic points in the dynamics. A similar estimate, that makes



(a) $2x(\mod 1)$



(b) Cat map

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

directly use of the O'Brien formula (14), has been proposed in [79]:

$$\hat{\Theta}_m(\rho) := \frac{\sum_{i=0}^{N-1-m} \mathbb{1}(\phi(T^i x_0) > u(\rho) \cap \max_{j=1,\dots,m} \phi(T^{j+i} x_0) \leq u(\rho))/(N-m)}{\sum_{i=0}^{N-1} \mathbb{1}(\phi(T^i x_0) \geq u(\rho))/N}. \tag{40}$$

Nevertheless, making use of the spectral result (20) and estimating each p_k shows at least two advantages:

- Conditions $D(u_n)$ and $D^k(u_n)$ do not have to be checked to apply the method and it holds in more general settings (although the conditions of application of the spectral theory are sometimes difficult to check).
- We can determine the p_k that are non null and we can compute their values, which can be particularly interesting to reveal dynamical patterns, such as periodicity of the target points in the last examples.

For these reasons, throughout the thesis, we will use in many situations the estimate $\hat{\theta}_m$ for our numerical computations, at an order m that will depend on the situation. We give here a Matlab code that can be used for the computation of these estimates:

```
% Function extremal_index takes as entries a vector of time series X, a quantile rho and
% m is the order at which theta is estimated.

% The output is:

% theta: estimation of the extremal index.
% p: the vector countaining the estimates of the p_k.


function [theta p]= extremal_index(X,rho,m)

%The threshold is the p-quantile of the distribution
 u=quantile(X,rho);

 %Number of exceedences of Y over the threshold u
 N=length(find(X>u));
nn=zeros(m,1);

%Computation of p_0
 for i=1:length(X)-1
     if X(i)>u & X(i+1)>u
      nn(1)=nn(1)+1;
     end
 end
 p(1)=nn(1)/N;

 %Computation of qk
 for k=2:m
     for i=1:length(X)-k
         if X(i)>u
             if max(X(i+1:i+k-1))<u & X(i+k)>u
             nn(k)=nn(k)+1;
```

```
                end
            end
        end
        p(k)=nn(k)/N;
    end

%Computation of theta
theta=1;
for k=1:kmax
theta=theta-p(k);
end

end
```

We are now equipped with the principal tools of extreme value theory to be used in the study of chaotic dynamical systems. We will see that EVT is not only useful in a predictive way, to understand the distribution of events of large intensity, but also allows to extract useful indicators associated with the system. We already mentioned its relation to hitting times statistics (we will come back to hitting time statistics in section (104), in connection with the spectrum of generalized dimensions of the measure). Recently, EVT base methods have been used (especially in climate science) to compute diverse local quantities that provide valuable information on the local dynamics of the system. We will describe this approach in the next section.

4 Extreme Value Theory and local dynamical indicators

The methods that we will present give access the fine geometrical structure of the attractor by computing its local dimension. We now present this quantity along with other dimensional quantities associated with the invariant measure.

4.1 Fractal dimensions

Let (M, T, μ) be a dynamical system defined by the transformation T acting on the metric space M (usually of subset of $\mathbb{R}^n$ or of the torus $\mathbb{T}^n$) that leaves the probability measure μ invariant. Diverse dimensional quantities have been introduced to describe the geometrical structure of the measure, which can be extremely complex.

4.1.1 Local dimensions

To characterize the local scaling behavior of μ, Young [76] introduced the concept of local dimension.

Definition 9 *We call the local dimension d(z) of μ at the point $z \in M$ the quantity defined by the following limit (if it exists):*

$$d(z) := \lim_{r \to 0} \frac{\log \mu(B(z, r))}{\log r}, \tag{41}$$

where $B(z, r)$ denotes a ball centered in z of radius r.

$d(z)$ gives the scaling of the measure of a ball centered in z when its radius r goes to 0. For example, if we consider a transformation preserving the Lebesgue measure defined on the unit square, a point z in the interior of the square and denote $dist(.,.)$ the Euclidean distance, we have that $d(z) = \lim_{r \to 0} \frac{\log \pi r^2}{\log r} = 2$, which matches our usual concept of dimensions. For measures living on more complex fractal supports or with non uniform densities, $d(.)$ can also take non-integer values. For reasons that will be developed throughout the thesis, we also introduce the following quantity when the limit in equation (41) is not taken:

$$d(z, r) := \frac{\log(\mu(B(z, r))}{\log r}. \tag{42}$$

We call it the r-resolution local dimension at the point z. This quantity depends

on the invariant measure μ and both on the point z and the radius r considered.

Definition 10 *The measure μ is said to be exact dimensional if the limit defining the local dimension exists and is constant for almost all $z \in M$. This typical value, denoted D_1, is called the information dimension of the measure and is equal to the average value of the local dimensions over M:*

$$D_1 := \lim_{r \to 0} \frac{\int_M \log(\mu(B(z,r)))d\mu(z)}{\log r}. \tag{43}$$

Whenever the local dimensions exist, it is enough that μ is ergodic to be exact dimensional. The proof of this result is due to Lai-San Young [76]: due to the fact that μ is T-invariant, it is not difficult to show that $d(z)$ and $d(Tz)$ are equal for a generic point $z \in M$, so that all points of the trajectory starting from z have the same local dimension. Then an argument using the ergodicity of μ is used to conclude. We will from now on work with such exact dimensional systems, or at least suppose that they are.

4.1.2 Generalized dimensions

Even when the local dimension is constant $\mu - a.e.$, M may contain points with different scalings that do not contribute to the mass of μ. In that case, we say that the system is *multifractal*, as it is not described by a unique fractal dimension. It became apparent in the eighties that many physical systems seem to exhibit multifractal properties. Originally developed to study turbulence [14, 15, 19], some new tools have then been introduced along the years to deal with such multifractal systems. An effort has been made by many mathematicians to transpose these ideas to the study of strange attractors originating from chaotic dynamics [17],[18],[20],[24]. This effort has resulted in what is now called multifractal analysis, of which a comprehensive introduction is found in [25]. It is now applied in many different fields of science, such as signal processing [26] and image reconstruction [16]. We will now give a concise description of the matter. Let us start with some definitions.

Definition 11 *We call the α-level set Λ_α of the system the set of points of M whose local dimension is α.*

A characterization of the size of these sets in usual terms of measure is impossible, as for $\alpha \neq D_1$, we have seen that $\mu(\Lambda_\alpha) = 0$ for ergodic measures. Nevertheless, it is still possible to quantify the 'strength' of this set by indicating its Hausdorff dimension $f(\alpha)$. The function $f(\alpha)$, that is also called the singularity spectrum of the measure μ, is of fundamental importance since it characterizes the geometry of the attractor and its associated invariant measure. In practice, this quantity is extremely difficult to compute directly either analytically or via numerical investigations, as it involves distinguishing uncountably many sets of different scaling and exploring the fine structure of the measure. To approach in a more indirect way the details of this geometry, Grassberger and Procaccia introduced in [12] and [17] useful quantities.

Definition 12 *We call the Generalized dimension of order q the quantity (whenever the limit exists)*

$$D_q := \lim_{r \to 0} \frac{\log \int_M \mu(B(z,r))^{q-1} d\mu(z)}{(q-1)\log r}. \tag{44}$$

By introducing the weight exponents q in the integral, the quantity D_q is no more a simple average of the measures of local balls but it gives, in the case of multifractal measures, more weight to points of the attractor with different scaling than D_1 and it makes the integral deviate from this value. For such measures, the D_q spectrum is not constant and equal to D_1 but it varies non trivially with q. Note that D_1 is recovered by taking the limit $q \to 1$ and applying l'Hospital's rule, so that the quantities can be defined for all real q. It is a well known fact that D_0 is simply the Hausdorff dimension of the support of the measure; D_2 is called the *correlation dimension*, and D_∞ and $D_{-\infty}$ are respectively the minimum and the maximum of the local dimensions over M. It is now clear that a non trivial spectrum of generalized dimensions originates from the presence of points of different scalings hidden in the attractor.

The following will make this fact more precise: let us assume that the function $f(\alpha)$, defined as the Hausdorff dimension of the level set Λ_α, is continuous in its interval of definition $(\alpha_{min}, \alpha_{max})$. This assumption originates from considerations coming from the analysis of turbulence in physics and fluid dynamics [20],[19]. This correspondence also seems to indicate that for a small radius r, and physically relevant measure μ, the density of points $z \in M$ for which $\mu(B(z,r)) \in (r^\alpha, r^{\alpha+d\alpha})$ is of the form $\rho(\alpha) r^{-f(\alpha)} d\alpha$. Under these assumptions we can write:

$$\int_M \mu(B(z,r))^{q-1} d\mu = \int_\alpha r^{q\alpha} \rho(\alpha) r^{-f(\alpha)} d\alpha. \tag{45}$$

For a given q, when taking the limit of small r, the dominant term in the right hand side of the above equation is $\rho(\alpha^*) r^{\inf_\alpha\{q\alpha - f(\alpha)\}}$, where α^* is the value of α for which the infimum is reached. On the other hand, equation (44) implies that the left hand side of the above equation scales like $r^{(q-1)D_q}$ for small r. By combining the two scalings, we get that:

$$D_q(q-1) \approx \inf_\alpha \{q\alpha - f(\alpha)\}. \tag{46}$$

We therefore arrive at the fundamental result that the function $\tau(q) = D_q(q-1)$ is the Legendre transform of $f(\alpha)$, provided the latter is convex. As the functions f and τ are convex conjugates, we also have the relation:

$$f(\alpha) = \inf_q \{\alpha q - \tau(q)\}. \tag{47}$$

This relation is essential since the singularity spectrum, hardly accessible by numerical studies can be computed via a simple Legendre transform of the spectrum of generalized dimensions, for which different computational methods have been developed along the year [17, 27, 28, 39, 40].

Throughout the thesis, several new methods will be proposed to evaluate these spectra of generalized dimensions and their relation to different dynamical properties of the system will be revealed. Various of these methods make use of Extreme Value Theory. Indeed, we will see that for well chosen observables, like the one introduced in the previous chapter, the scale parameter σ of the GEV distribution is related to some dimensional quantities associated to the system. In the rest of the thesis, we will study in details the Extreme Value Laws associated to these observables, which will provide several numerical methods to compute such quantities.

4.2 Local dimensions and inverse persistence

In the last few years, the analysis of observational time series in term of local indicators of the dynamics estimated via EVT has gained importance. To introduce these methods, we will start by describing the physical importance of these quantities in the context of atmospheric dynamics. A pioneer work in the domain was published by Faranda, Messori and Yiou [105], where these tools were applied

to the study of the local properties of the North-Atlantic atmospheric attractor. The system under consideration is the daily evolution of the sea level atmospheric pressure on a grid of $53 * 20 = 1060$ points in the North Atlantic. For each atmospheric state z of the attractor (that consists of a vector in which each entry is the daily pressure at a point of the grid), two quantities were computed: the *instantaneous dimension* that corresponds to the finite resolution local dimension defined in equation (42) and the *inverse persistence*, which as we will see is an order 0 finite resolution estimate of the extremal index associated with the standard observable $\phi_z(x) = -\log(dist(z, x))$. We will see that the finite resolution local dimension can also be estimated by studying the extremes of the empirical distribution of ϕ_z, so that the two estimates can be evaluated simultaneously.
These quantities are useful indicators of the local dynamics of the system, since the local dimension can be interpreted as the number of different degrees of freedom, or directions along which the system can evolve near the state z, and is an indicator the local predictability of the system. The inverse persistence for its part was used to characterize the stability of the state z.

A wide distribution of each of these quantities was found, and it was shown to characterize particular patterns associated to well known climatic phenomena. In particular, for a wide range of values of r, the minima of the local dimensions $d(z, r)$ correspond to zonal flow circulation regimes, where the low pressure systems are confined to the polar regions, opposite to high pressure areas which insist on southern latitudes, in a North-South structure. On the other hand, the maxima of the local dimension correspond to blocked flows, where high and low pressure structures are distributed in the Est-West direction [105]. Strong correlations were found between inverse persistence and local dimensions, in the sense that states with high local dimension are more likely to be less persistent and states with low local dimensions are generally more stable [105].

Now that the physical relevance of this type of analysis has been discussed, we will describe the method that has been used in [105] and extensively described in [85] to compute these quantities. Let us place ourselves in the general set up of a chaotic dynamical system (M, T, μ) and consider the observable ϕ_z defined for all point $x \in M$ by $\phi_z(x) = -\log(dist(z, x))$, where z is a point of the attractor. As mentioned in section (3), the observable is defined in such a way that the complementary distribution function of $\phi_z(.)$

$$\bar{F}_z(u) = \mu(\{x \in M, \phi_z(x) > u\}) \tag{48}$$

coincides with $\mu(B(z, e^{-u}))$, the measure of a ball centered in z and of radius e^{-u}. This quantity is equal, by definition (42) to the quantity $e^{-ud(z,e^{-u})}$. EVT can then be applied to study the tail of $\bar{F}$ and allows us to extract the local dimension of the point z at resolution $r = e^{-u}$. By letting u tend to $+\infty$, this quantity converges to $d(z)$.

To extract this quantity, one can use the POT approach described in section (3): consider a long trajectory and at each iteration compute the observable ϕ_z. Once a long enough time series is obtained, a high quantile of the observable distribution is taken as a threshold u_z^* and the exceedances over the threshold are considered, so that an empirical distribution of these extreme events is obtained. As $\bar{F}_z(u) = \mu(B(z, e^{-u}))$ decreases (for regular enough measures) exponentially fast with u, the observable distribution is flat tailed and we expect the distribution of the exceedances over a high enough threshold to be a Generalized Pareto Distribution of shape parameter 0 (which is in fact an exponential distribution) and of scale parameter σ. This reads:

$$\mu(\{x, \phi_z(x) - u_z^* > s\}) \sim e^{-s/\sigma}. \tag{49}$$

By fitting the numerically obtained empirical distribution of exceedances, one can estimate the parameter σ whose inverse provides an estimate of the $(e^{-u_z^*})$-resolution local dimension around the point z. This distribution being exponential, the scale parameter σ is also equal to the inverse of its mean and can be estimated as such. This method of evaluation is generally more stable than performing a fit of the distribution, since it does not suffer from the systematic errors associated to fitting procedures and it is often used in physical situations (see for example [105]). This method is easy to implement, does not require enormous amount of data and allows fast computations. A detailed review of the computational aspects of the method is available in [85], from section 9.3.

This technique allows to compute at the same time the extremal index appearing in the extreme value law, which can provide relevant information on the dynamics of the system: whenever z is not a fixed point of the dynamics, we expect that in the limit of small radius, any point x inside the ball centered at z will leave it immediately after one iteration. This translates into the fact that the quantity

p_0 of equation (18) is zero and $\theta_0 = 1 - p_0 = 1$, as discussed in section (3.4). However, when estimating this quantity with a finite threshold $u_z^* = -\log r$ (that corresponds to the $\rho-$quantile of the empirical distribution of ϕ_z), the Süveges estimate $\hat{\theta}(\rho)_{su}$ carries a dynamical meaning: in fact, a low value (close to 0) is characteristic of a high mean clustering of the exceedances of ϕ_z, indicating that the system spends, in average, a long time in the neighborhood of the point z (the ball of center z and of radius r). On the contrary, if its value is close to 1, the motion tends to leave the ball almost immediately after it entered it, showing an instability of the state z. This quantity, which contains information on the stability of the state z, is named in [105] the *inverse persistence* and is denoted $\theta(z, r)$.

This description of the extremal index is interesting from a physical point of view, but is quite restrictive. Not only it is an estimate at finite resolution, but since it is an estimate of order 0, it accounts only for consecutive exceedances over the threshold, and in particular it does not account for eventual periodic returns to the neighborhood of z. The estimate $\hat{\theta}_m$ that we introduced does account for this effect. To evaluate the differences between the two estimates in real world-time series, we test them on the daily sea-level pressure data extensively described in [105], in which the Süveges estimate was used for the computation of θ. As expected, estimates using $\hat{\theta}_5$ are systematically lower than those with the order 0 method (in average 0.065 less), due to the contribution of some $\hat{p}_k$ for $k > 0$ (see the empirical distribution in figure (2)). To be compared fairly, both estimates are computed using as a threshold the 99^{th} quantile of the observable distribution. We believe that this method suggests a physical interpretation of the extremal index different from mere inverse persistence near a given state, since it reflects a more general form of clustering, taking into account periodic returns of the system to the neighborhood of the point z.

Let us emphasize that, theoretically, both the local dimension $d(z)$ and the extremal index θ_z are constant almost everywhere for the kind of chaotic physical systems we are considering (in fact the described method can be applied only if the system verifies these prescriptions). The fact that we find a broad distribution of both local dimensions and inverse persistence is due to the fact that we are considering estimates of these quantities when the convergence is not yet achieved. We will discuss this matter in the following section and see that the broad distribution of finite resolution local dimension can be interpreted as a sign that the physical invariant measure associated to the atmospheric attractor is in fact

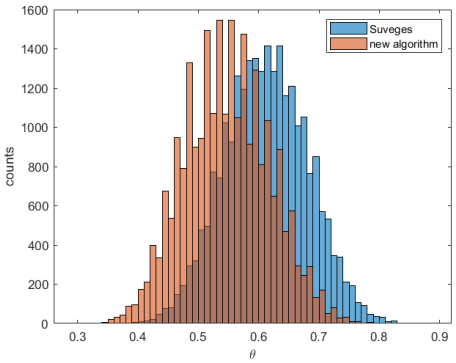


Figure 2: Comparison on the distribution of θ with the two estimates. The blue distribution is obtained using the Süveges estimate and the red one using $\hat{\theta}_5$ (taking for both as a threshold the 0.99-quantile of the observable distribution).

multifractal.

4.3 Local dimensions and generalized dimensions

We mentioned is section (4.1) that a non constant spectrum of generalized dimensions is a signature of multiple local scaling behaviors of the measure. This can be seen in the very definition of these quantities, given by equation (44) that, combined with relation (42), yields the following expression for D_q:

$$D_q = \lim_{r \to 0} \frac{\log \int_M r^{(q-1)d(z,r)} d\mu(z)}{(q-1)\log r}. \tag{50}$$

Written in this way, the influence of points with different local dimensions becomes even more apparent, and one can take advantage of this relation to propose a technique to estimate D_q, as we did in *Caby et al. Generalized dimensions, large deviations and the distribution of rare events [51]* . Indeed, we described in section (4.2) how the quantity $d(z, r)$ can be estimated via EVT. The integral at the right hand side of equation (50) can be estimated via a Birkhoff sum over many points z_i of the attractor chosen at random with distribution μ (for example, chosen along a generic trajectory of the system). Each of the different $d(z_i, r)$ can be computed via the EVT procedure described in section (4.2), taking as a threshold $u_z^* = -\log r$. Let us define the quantity

$$S(r, q) := \frac{1}{N} \sum_{i=1}^{N} r^{(q-1)d(z_i,r)}. \tag{51}$$

the computation of the quantity

$$\Delta(r,q) := \frac{\log S(r,q)}{(q-1)\log r},$$

with r sufficiently small, provides an estimate of D_q. Clearly, a constant distribution of $d(z_i,r) = \hat{D}_1$ gives estimates of D_q constant over q and equal to $\hat{D}_1$. On the other hand, when a wide distribution of $d(z_i,r)$ is found, $\Delta(r,q)$ varies non trivially with q and we observe that for $r < 1$:

- $\Delta(r,q) \xrightarrow[q\to\infty]{} \min_i\{d(z_i,r)\}$ as $q \to \infty$.
- $\Delta(r,q) \xrightarrow[q\to\infty]{} \max_i\{d(z_i,r)\}$ as $q \to -\infty$.
- $\Delta(r,q) \xrightarrow[q\to\infty]{} \overline{d(z_i,r)}$.

$\overline{d(z_i,r)}$ denotes the mean value of the distribution $(d(z_i,r))_i$. These properties also apply to the spectrum of generalized dimensions that they estimate. To test this technique, we will use the atmospheric data presented in the previous section, which constitute an archetype of observational time series originating from the chaotic dynamic of an underlying high dimensional system. We recall that the system under consideration is the daily evolution of the sea level atmospheric pressure on a grid of $53 * 20 = 1060$ points in the North Atlantic. Using the method described in the previous section, the authors in [105] provided estimates of finite resolution local dimensions using as a threshold the 98^{th} quantile of the observable distribution. A broad distribution on these values was found, ranging from 4 to 25, with an average value of about 12 (the system having potentially 1060 degrees of freedom). These results, which are stable against different choices of quantiles suggest the existence of a low dimensional attractor associated to atmospheric flows [105].

In principle, the method that we just described requires the computation of a distribution of local dimensions at a uniform resolution r_{cut}. However, as large variations of local dimensions are observed, the data are not suited for an analysis at fixed resolution, since this may over-sample or under-sample the distributions of extreme events of the observable $\phi_z(x) = -\log(d(z,x))$, depending on the region of the point z we are considering. For this reason, we follow the approach described in [105] and [106] and use as threshold values $T_{\rho,z}$ for the observable ϕ_z the ρ-quantile of the distribution of ϕ_z, where ρ is fixed. This ensures that the extreme value

statistics is computed with the same sample size at all points. The effective radius considered for the computation of the generalized dimensions is then taken to be the average of $e^{-T_{\rho,z}}$ over z. Applying formula (50) to the computation of D_q, one obtains the non-linear behavior pictured in figure (3). Let us briefly describe the results found when adopting different quantiles reported in table (2). When the quantile ρ is relatively low, a large sample of recurrences is used (corresponding to a larger average cutoff radius). This implies a smaller spread of the distributions of $d(z,r)$. To the contrary, when ρ is larger, the sample statistics contains fewer recurrences and the spread in $d(z,r)$ increases. Note that, although $\min(d(z,r))$ and $\max(d(z,r))$ seem to experience large variations with different quantiles, these have to be compared with the dimension of the phase space, which corresponds here to the number of grid points of the sea-level pressure fields used, 1060. The relative variation is therefore very small, less than 1%. Depending on the size of the data sets, one can then look for the best estimates of the $d(z,r)$ distribution for several values of q and look for a range of stable estimates in q space.

ρ	$D_{-\infty} = \max(d(z,r))$	$D_1 = \overline{d(z,r)}$	D_2	$D_\infty = \min(d(z,r))$
0.95	20.7	11.2	9.5	6.0
0.97	24.2	12.2	10.2	6.4
0.98	25.7	13.0	10.5	6.4
0.99	29.1	14.3	11.1	6.5
0.995	39.6	15.5	11.3	6.2

Table 2: Values of D_q found with different quantiles. For all of them, estimates $D_{-\infty}$ and D_∞ match with the extrema of the local dimensions, and D_1 with the average of the local dimensions. Estimates of D_2 are larger than the value of 8.9 found with a different technique in [50].

This method cannot be expected to give precise estimates of generalized dimensions, for various reasons. First, we are working with estimates of quantities that may be far from convergence and may lead to poor results. The method is also very sensitive to the extrema of the distribution of local dimensions, that can take pathological values at rarefied regions of the attractor. An example of this phenomena is found in the supplementary material of [105], where the authors found, for points near the extremity of the wings of the Lorenz attractor, finite resolution local dimensions much larger than 3, the dimension of the ambient space. In fact, due to the particular location of these points, the ball surrounding them intersects the attractor in only a small region, leading to inconsistency of

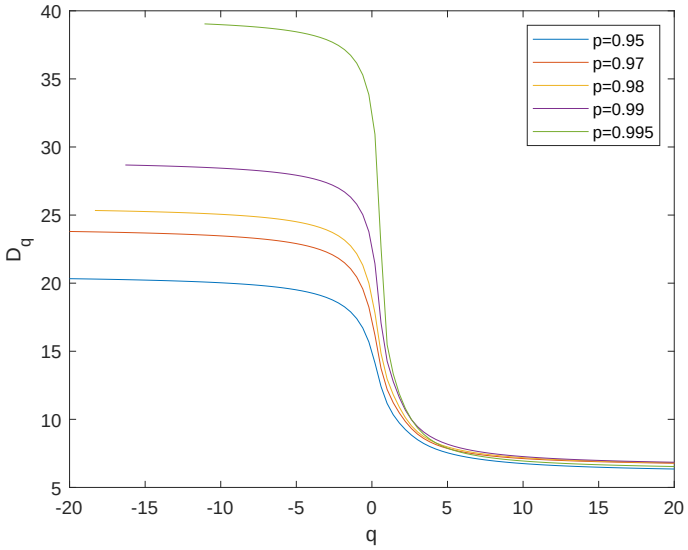


Figure 3: D_q spectrum obtained from climate data, using equation (51) and the techniques described in the text. Curves are displayed for different values of the quantile ρ.

the statistic of recurrences. This explains the apparent wide variations of the curves for different quantiles for negative q.

Despite these limitations, this approach can be seen as an illustration of the merits of the analysis of chaotic systems in terms of local dimensions and of the fact that was mentioned earlier: when analyzing finite time series, we are in fact extrapolating properties of the invariant measure living on the attractor using finite trajectories. The word *penultimate attractor* has been proposed to describe the object defined by a set of finite trajectories supposedly hosted by the attractor. Within this perspective, the finite resolution local dimensions that we compute are objects associated with the penultimate attractor, and are deeply connected to the generalized dimensions of the system, which are characteristic of the *ultimate* fine structure of the attractor. The idea that multifractality can be seen as a "gate" between penultimate local indicators of the system and the ultimate measure will be explored in a more systematic way in section (7.1), where large deviation results concerning several local indicators of the dynamics in connection with the generalized dimensions of the system will be described.

4.4 Influence of noise on θ_z and on the local dimensions

The influence on the value of θ_z of a uniform noise perturbing the dynamics has been studied both analytically and numerically and have shown to immediately turn the value of θ_z to 1, due to the fact that perturbations destroy the eventual periodicity of the orbits. Extensive results and discussions of this matter can be found in [85, 63]. The aim of this section is to investigate whether different kinds of discrete perturbations of a deterministic dynamics, or some discrete observational noise affect the value of the extremal index in a different way, as these situations may be paradigmatic of what happens in real world systems like the atmospheric example presented earlier. We will see that the presence of discrete noise can affect in an important way the computation of the extremal index. The following results can be found in *T. Caby, D. Faranda, S. Vaienti, P. Yiou, On the computation of the extremal index for time series [112].*

4.4.1 I.i.d. random tranformations

A common way to construct a stochastically perturbed dynamical system is the following: a deterministic dynamical system (M, T, μ) is considered, but instead of successively mapping a point x by T, the motion is defined by successive iterations of randomly perturbed maps $T_{\xi_i} = T + \varepsilon\xi_i$, where ε models the *intensity of noise* and $\xi_i \in \Xi$ are i.i.d. random variables in a probability space $(\Xi, \mathbb{P})$. A classic example is obtained by taking $\mathbb{P}$ equal to the uniform distribution in $\Xi = B(0, 1)$. One must be careful however that such a perturbed dynamics is well defined and that the noise does not push the motion away from the basin of attraction of the system, making the trajectories diverge. When these requirements are satisfied, it is sometimes possible to prove the existence of a stationary measure μ_s for the system, defined as follows.

Definition 13 *We say that μ_s is a stationary measure for the system if for all μ_s-integrable function g,*

$$\int_\Xi \int_M g(T_\xi(x)) d\mu_s(x) d\mathbb{P}(\xi) = \int_M g(x) d\mu_s(x). \tag{52}$$

Whenever the system has polynomial decay of correlations and when there exists an absolutely continuous stationary measure μ_s, it was proven in [59] that an extreme value law is associated to the standard observable ϕ_z and equation (18) still holds, provided the measure considered is replaced by the product measure $\nu = \mathbb{P} \times \mu_s$. In the case when $\mathbb{P}$ is the uniform distribution in $B(0, 1)$, the authors

of [59] also showed that the extremal index is 1 whenever $\varepsilon > 0$. We will now present an example of discrete perturbations (i.e. the distribution of the noise is a sum of Dirac masses), where the extremal index can be computed explicitly and is different from 1.
Let us consider:

- Two maps, $T_0 = 2x \pmod 1$ and $T_1 = 2x + \varepsilon \pmod 1$, $0 < \varepsilon < 1$.
- A *fixed* target set around 0: $B_n(0) := B(0, e^{-u_n})$: the ball of radius e^{-u_n} around 0. Notice that 0 is a fixed point of T_0, but not of T_1.

For this map, $T = T_0 = 2x \pmod 1$, ξ_i can take the value 0 or 1 and is a Bernoulli random variable with distribution that we take equal to $\mathbb{P} = \rho\delta_0 + (1-\rho)\delta_1$, with $\rho = \frac{1}{2}$. Since both T_0 and T_1 preserve the Lebesgue measure, this measure is stationary for the system, leading to the following formula for p_0:

$$\begin{aligned} p_0 &= \lim_{n\to\infty} \frac{\int d\mathbb{P}(\xi)\mathbb{L}(B_n(0) \cap T_\xi^{-1} B_n(0))}{\mathbb{L}(B_n(0))} \\ &= \lim_{n\to\infty} \left[\rho \frac{\mathbb{L}(B_n(0) \cap T_0^{-1} B_n(0))}{\mathbb{L}(B_n(0))} + (1-\rho)\frac{\mathbb{L}(B_n(0) \cap T_1^{-1} B_n(0))}{\mathbb{L}(B_n(0))}\right] \end{aligned} \tag{53}$$

In [65], it was proven that for expanding maps of the interval, if a target point z does not have period $k+1$, then the term p_k is 0. Therefore, as 0 is not a fixed point of the dynamics induced by T_1, the last term in the bracket of the above equation vanishes in the limit. On the other hand, using arguments from [37], if 0 is a fixed point of the dynamics, the value of the term of order 0 appearing in the spectral formula (20) is non zero and it is given in terms of the inverse of the derivative of the map at the target point. Zero being a fixed point of the map T_0, the limit of the first term in the bracket can be computed explicitly with the following result:

$$p_0 = \rho/|T'(0)| = 1/4.$$

The preceding computation of p_0 is valid for all values of $\varepsilon \in (0,1)$. We may now ask the question whether other values of p_k can be different from zero. By a similar reasoning, this problem can be reduced to the existence of a concatenation $T_{\xi_{k+1}} \circ ... \circ T_{\xi_1}$ mapping the point 0 to itself after $k+1$ iterations and not before. Indeed, in such case we have that

$$p_k \geq \lim_{n\to\infty} \frac{\mathbb{P}(\xi_1,...,\xi_{k+1})\mathbb{L}(B_n(0)\cap(T_{\xi_1}^{-1}B_n(0))^c\cap...\cap(T_{\xi_k}^{-k}B_n(0))^c\cap T_{\xi_{k+1}}^{-k-1}B_n(0))}{\mathbb{L}(B_n(0))}$$
$$= \frac{\mathbb{P}(\xi_1,...,\xi_{k+1})}{2^k}. \tag{54}$$

On the other hand, if the dynamics cannot come back to 0, then $p_k = 0$ for $k \neq 0$ (using again the argument from [65]).
To investigate the existence of such a concatenation, let us distinguish between two cases.

- It is easily seen that if ε is irrational and $\xi_1 \neq 0$, any concatenation of any length will map 0 into an irrational point, so that the dynamics cannot return to 0 after having left from it. For this reason, all the p_k but p_0 are null.
- It remains to investigate the case when $0 < \varepsilon = p/q < 1$ is rational, with $p, q \in \mathbb{N}$ mutually prime.

We start by showing by induction that for $n \geq 1$ and for all $i \in [2^n, 2^{n+1} - 1]$, the dynamics starting from p/q can attain the points $\frac{ip}{q} \pmod 1$ (and only them) in exactly n steps.
This proposition is true at rank $n = 1$, since $T_0(p/q) = 2p/q$, $T_1(p/q) = 3p/q \pmod 1$. Suppose now that it holds at rank n. Then we have that

- $T_0(2^n p/q \mod 1) = 2^{n+1}p/q \pmod 1$,
- $T_1(2^n p/q \mod 1) = (2^{n+1} + 1)p/q \pmod 1$,
- ...
- $T_0((2^{n+1} - 1)p/q \mod 1) = (2^{n+2} - 2)p/q \pmod 1$,
- $T_1((2^{n+1} - 1)p/q \mod 1) = (2^{n+2} - 1)p/q \pmod 1$,

which shows that the proposition is true at rank $n+1$. This result implies that any point of the form $jp/q \pmod 1$, with $j \in \mathbb{N}, j \geq 2$ can be attained by the dynamics starting from the point p/q.

Let us consider $T_{\xi_m} \circ ... \circ T_{\xi_2} \circ T_1$, the concatenation of minimal length $m \geq 2$ such that $T_{\xi_m} \circ ... \circ T_{\xi_2} \circ T_1(0) = qp/q \pmod 1 = 0$. Note that the first map involved in this concatenation is T_1, so that the first iterate of the dynamics is p/q. By the preceding argument, there exists such a concatenation and since we consider the shortest concatenation(s) having this property, the dynamics leaving 0 does not come back to 0 before m iterations. This implies, as discussed earlier, a non zero value of p_{m-1}.

In fact, there exists infinitely many non zero p_k: consider the concatenation of minimal length $m' \geq 2$ $T_{\xi'_{m'}} \circ ... \circ T_{\xi'_2} \circ T_1$ such that $T_{\xi'_{m'}} \circ ... \circ T_{\xi'_2} \circ T_1(0) = (q+1)p/q \pmod 1 = p/q \pmod 1$. By the argument in the induction, we have that either $m' = m$ or $m' = m+1$. Then it is easily seen that

$$T_{\xi_m} \circ ... \circ T_{\xi_2} \circ T_{\xi'_{m'}} \circ ... \circ T_{\xi'_2} \circ T_1(0) = T_{\xi_m} \circ ... \circ T\xi_2(p/q) = 0.$$

This proves the existence of a concatenation returning the point 0 to itself after exactly either $2m$ or $2m+1$ iterations (and not before). Again we have proven that either p_{2m} or p_{2m-1} is non zero. Applying the same reasoning, we can prove that infinitely many p_k terms have positive values.

An interesting situation happens when $\varepsilon = 1/2$. In this case, for all k, 0 can come back to itself after exactly $k+1$ iterations, and this happens only for the concatenation $T_0 \circ T_1 \circ \cdots \circ T_1 \circ T_1$, where the map T_1 is applied successively k times. This sequence having probability $(1-\rho)^k \rho = 1/2^{k+1}$, it is easily seen (by the definition (18) of p_k and analogously to the proof of the computation of p_0 in equation (53)) that none of the p_k is zero; in fact p_k is equal to:

$$\begin{aligned} p_k &= \frac{1}{2^{k+1}} \lim_{n\to\infty} \frac{\mathbb{L}(B_n(0) \cap (T_1^{-1}B_n(0))^c \cap ... \cap (T_1^{-k}B_n(0))^c \cap T_0^{-k-1}B_n(0))}{\mathbb{L}(B_n(0))} \\ &= \frac{1}{4^{k+1}}. \end{aligned} \tag{55}$$

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

$$\theta = 1 - \sum_{k=0}^{\infty} (\frac{1}{4})^{k+1} = 2/3. \tag{56}$$

We computed in table (3) some estimates $\hat{p}_k$, defined by formula (37) and a very good agreement is found.

	Theoretical values	Estimate $\hat{p}_k$	Uncertainty
p_0	0.25	0.250	0.0020
p_1	0.0625	0.062	0.0016
p_2	0.015625	0.016	8.10^{-4}
p_3	0.00390625	0.0039	4.10^{-4}
p_4	$9.765625.10^{-4}$	$9.3.10^{-4}$	$1.7.10^{-4}$

Table 3: Estimations of the p_k using $\hat{p}_k$ for the precedent example with $\varepsilon = 1/2$. These results compare favorably with theoretical values to a precision of the order of 10^{-4}. Results are averaged over 20 trajectories of 5.10^6 points and we took the 0.995-quantile of the observable distribution as a threshold. The uncertainty is the standard deviation of the results.

To sum up: we have obtained the interesting result that under some particular discrete stochastic perturbation, the extremal index θ_z can be different from 1, and in some cases, all p_k are non zero. We now comment these results.
In this example, the condition $D^m(u_n)$ of theorem (3) is clearly not satisfied, because the value of p_k is different from 0 for all $k \in \mathbb{N}$. In this sense, the spectral technique can be applied in more general situations than the classical theory of O'Brien.
The preceding example can also be interpreted as a perturbation of the map $T_0 = 2x \pmod 1$ around its fixed point 0. Instead, we change the point of view and consider it as a perturbation of the map $T_1 = 2x + \varepsilon \pmod 1$ around 0. If we suppose that the map T_0 is chosen with probability ρ, we see that when $\rho = 0$ (absence of perturbation), the $EI = 1$, since 0 is not a fixed point for T_1, but this value jumps to $EI \leq 1 - 0.5\ p_0$ as soon as $\rho > 0$, for any value of $\varepsilon > 0$. This behavior is analogue to what happens in the case of uniform additive noise: it is enough to switch on the noise, no matter how small, to make the EI equal to one. Here, the EI changes in term of the probability of appearance of the perturbed map, no matter of its topological distance from the unperturbed one, which is the value of ε in this case.

However, notice that it is easy to construct examples where the EI depends also on the *distance* between the maps. For instance, take now $T_1 = (2 + j)x \pmod 1$,

$j \in \mathbb{N}$, which could be interpreted as a strong perturbation of T_0. By repeating the argument above we immediately get that the $EI \leq 1 - \rho\ \frac{1}{2+j}$.
In this example, computations of p_k and of the EI can be made explicitly. We will use this result in section (6.1) to study numerically the statistics of the number of visits for stochastically perturbed systems.

Let us now investigate the influence of other families of noise of physical relevance.

4.4.2 Moving target

We will now study another kind of disturbance: consider a deterministic system and a target set that is slightly moving at each iteration, either continuously or in a discrete way. This can be regarded as a simple model of observational noise.

The discrete case. Let $z \in X$, consider σ to be the right-sided shift on m symbols and set $\Omega = \{1, \cdots, m\}^{\mathbb{Z}}$. To each symbol, $j = 1, \ldots, m$, associate a point $z_j \in B(z, \varepsilon)$. Take as a probability $\mathbb{G}(\omega), \omega = (\omega_0, \omega_1, \ldots) \in \Omega$ the Bernoulli measure of weights $\rho_i = \frac{1}{m}, i = 1, \ldots, m$. This measure is invariant under the shift σ. The value of ϵ is fixed and models the strength of the uncertainty. At each temporal step, the map T moves x and the target point moves as well around z under the action of σ. We now construct the direct product $S(x, \omega) := (Tx, \sigma\omega)$. The probability measure $\mathbb{P} := \mu \times \mathbb{G}$ is invariant under S, implying the existence of an extreme value law [112] with an extremal index θ that can be expressed by adapting to the present case the equation (18). In this case, p_0 reads:

$$p_0 = \lim_{n \to \infty} \frac{\int d\mathbb{G}\ \mu\left(B_n(z_{\omega'}) \cap T^{-1} B_n(z_{\sigma^{-1}\omega'})\right)}{\int d\mathbb{G}\ \mu(B_n(z_{\omega'}))}. \tag{57}$$

We now give a simple example of such a situation, where $p_0 > 0$, which implies that the extremal index is strictly less than 1. Take as T the map $T(x) = 2x \pmod 1$, and an alphabet of 4 letters: $\{0, 1, 2, 3\}$, with equal weights $1/4$. The invariant measure is here the Lebesgue measure $\mathbb{L}$. We set the associations:

$$0 \to z_0,\ 1 \to z_1,\ 2 \to z_2,\ 3 \to z_3, \tag{58}$$

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

$$T(z_1) = T(z_2) = z_0,\ T(z_0) = z_3,\ T(z_3) \neq z_i,\ i = 0, 1, 2, 3. \tag{59}$$

The numerator of p_0 in equation 57 is:

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

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

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

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

$$\int d\mathbb{G}\mathbb{L}(B_n(z_{\omega'})) = \sum_{l=1}^{4}\int_{[\omega_0=l]} d\mathbb{G}\mathbb{L}(B_n(z_l)) = \sum_{l=1}^{4}\rho_l\ \mathbb{L}(B_n(z_l)) = \mathbb{L}\left(B_n(z_0)\right), \tag{60}$$

by Lebesgue translational invariance.

Using the same arguments as in equation (53), we see that for all cylinders in the numerator different from the three explicitly given above, the corresponding integrals vanish in the limit of large n. For the first of the three cases above (the others being similar), we get

$$\begin{aligned}\int_{[\omega_0=1,\omega_{-1}=0]} \mathbb{L}\left(B_n(z_1)\cap T^{-1}B_n(z_0)\right) d\mathbb{G} &= \frac{1}{2}\int_{[\omega_0=1,\omega_{-1}=0]} \mathbb{L}\left(B_n(z_0)\right) d\mathbb{G} \\ &= \frac{1}{2}(\frac{1}{4})^2\mathbb{L}\left(B_n(z_0)\right).\end{aligned} \tag{61}$$

Therefore the numerator equals $\frac{3}{2}(\frac{1}{4})^2\mathbb{L}\left(B_n(z_0)\right)$.

The quantity $\mathbb{L}\left(B_n(z_0)\right)$ cancels out with the one in the denominator and we get

$$p_0 = \frac{3}{2}(\frac{1}{4})^2 = 0.09375. \tag{62}$$

To obtain the value of p_1, we need to compute the following quantity:

$$p_1 = \lim_{n\to\infty} \frac{\int d\mathbb{G}\ \mathbb{L}\left(B(z_{\omega'},n)\cap T^{-1}B(z_{\sigma^{-1}\omega'},n)^c \cap T^{-2}B(z_{\sigma^{-2}\omega'},n)\right)}{\int d\mathbb{G}\ \mathbb{L}(B(z_{\omega'},n))}. \tag{63}$$

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

$$\begin{aligned}\int_{[3,1,1]} d\mathbb{G}\ \mathbb{L}\left(B_n(z(\omega')) \cap T^{-1}B_n(z_{\sigma^{-1}\omega'})^c \cap T^{-2}B_n(z_{\sigma^{-2}\omega'})\right) \\ = (\frac{1}{2})^2 \int_{[3,1,1]} d\mathbb{G}\ \mathbb{L}(B_n(z_3)) = (\frac{1}{4})^3(\frac{1}{2})^2\mathbb{L}(B_n(z_3)).\end{aligned} \tag{64}$$

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

$$p_1 = 6(\frac{1}{4})^3(\frac{1}{2})^2. \tag{65}$$

We tested this result numerically with different sets of target points satisfying the aforementioned assumptions. We find good agreements of $\hat{p}_0$ with the theoretical value of $p_0 = 0.09375$ and of $\hat{p}_1$ with the theoretical value of $p_1 = 0.0234375$ with a precision of order 10^{-3} (see table 4).
In the case where none of the points z_i is the k^{th} iterate of another (in particular none of the points z_i is k-periodic), it can be shown that p_{k-1} is equal to 0 for $k > 2$. This is indeed what we find in our numerical simulations, which were performed up to order five.

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

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

This example is in some sense atypical since the four points $z_0, \ldots, z_3$ could be far away from each other. For instance, z_1 and z_2 are the two predecessors of z_0 and therefore they are on the opposite sides of $\frac{1}{2}$. By constraining the points to be in a small neighborhood of a given *privileged* center, the previous effects should be

absent and the EI should be one, or very close to it. However this example shows that in the presence of a moving target, it is not the periodicity which makes the EI eventually less than one. This gives another concrete example where the series in the spectral formula (20) has at least two terms different from zero.

The uniform case We claimed above that when the center of the target ball can take any value in a small neighborhood of a given point z_0, the EI collapses to 1, even for periodic points z_0. To model this more physically realistic situation, let us fix $\varepsilon > 0$ and consider the set $Z_\varepsilon = B(z_0, \varepsilon)$. We define a map f acting on Z_ε with an associated invariant probability measure ν that drives the dynamics of the target point $z \in Z_\varepsilon$. We suppose that ν is not atomic, meaning that for all $z \in M$, $\nu(\{z\}) = 0$. The observable considered is now $\phi(x, z) = -\log|x - z|$ on the product space $\{X \times Z_\varepsilon, \mu \times \nu\}$. By similar arguments to the ones we described in the case of discrete perturbation of the target point, we can show the existence of an extreme value law for the process $\phi \circ (T^k, f^k)$, with an EI given by equation (18), where

$$p_k = \lim_{n\to\infty} \frac{\int_{Z_\varepsilon} d\nu(z)\mu\left(B_n(z) \cap T^{-1}B_n(f(z))^c \cdots \cap T^{-k-1}B_n(f^{k+1}(z))\right)}{\int_{Z_\varepsilon} d\nu(z)\mu(B_n(z))}, \qquad (66)$$

where $B_n(y)$ denotes a ball around y of radius e^{-u_n}. We are now ready to state the following proposition:

Proposition 2 *Suppose that for all $k \in \mathbb{N}$, $\nu(\{z \in Z_\varepsilon | T^{k+1}(z) = f^{k+1}(z)\}) = 0$ and that μ is absolutely continuous with respect to Lebesgue with a bounded density h such that $h \geq \iota > 0$. Then the extremal index is* 1.

Proof: We compute the term q_k given by equation (66) and denote

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

where $r_n = e^{-u_n}$. We can write the numerator in equation (66) as a sum of integrals over $Z_1^{n,k}$ and its complementary set $Z_2^{n,k} = Z_\varepsilon \backslash Z_1^{n,k}$.
Since for $z \in Z_1^{n,k}$, z and z^* are at a distance larger than r_n, for all $z^* \in T^{-k-1}B_n(f^{k+1}(z))$, we have that $B(z, n) \cap T^{-k-1}B_n(f^{k+1}(z)) = \emptyset$, and the integral over $Z_1^{n,k}$ is zero for all n.

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

$$\begin{aligned} p_k &\leq \lim_{n\to\infty} \frac{\int_{Z_2^{n,k}} d\nu(z)\mu(B_n(z))}{\int_{Z_\varepsilon} d\nu(z)\mu_n(B_n(z))} \leq \lim_{n\to\infty} \nu(Z_2^{n,k}) \frac{\sup_{x\in M}\mu(B_n(x))}{\inf_{x\in M}\mu(B_n(x))} \\ &= \lim_{n\to\infty} \nu(Z_2^{n,k}) \frac{\sup_{x\in M} h(x)}{\inf_{x\in M} h(x)}. \end{aligned} \tag{67}$$

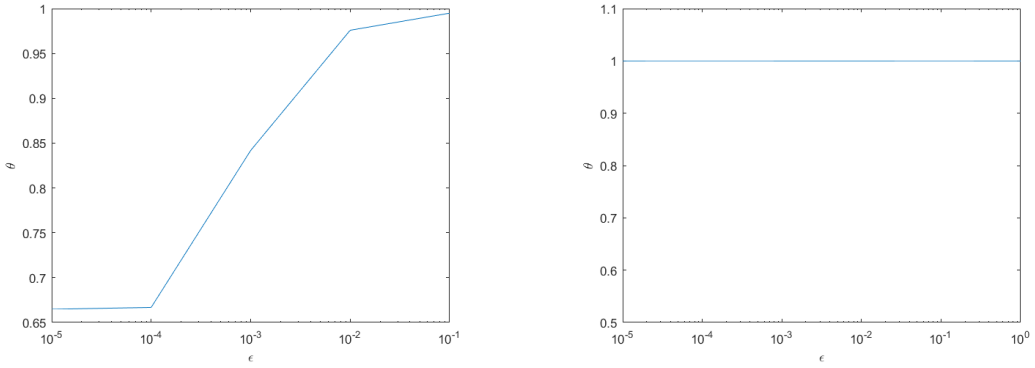
With our assumptions on h, the fraction in the last term is finite. Let us consider the limit set $Z_2^k = \cap_n Z_2^{n,k} = \{z \in Z_\varepsilon | T^{k+1}(z) = f^{k+1}(z)\}$. We have that $\lim_{n\to\infty} \nu(Z_2^{n,k}) = \nu(Z_2^k) = 0$, by hypothesis and so $p_k = 0$ for all k and therefore $\theta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

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

Observational noise. Instead of considering a perturbation of the target set driven by the map f, we now consider the center of the target ball as a random variable uniformly distributed in the neighborhood of size ε of the point z_0. This situation can model data with uncertainty or disturbances in their detection and is equivalent to the case *observational noise* which was considered in [90]. In this approach, the dynamics of a point x is given by $T^k x + \xi_k$, where ξ_k is a random variable uniformly distributed in a ball centered at 0 of radius ε. The process is therefore given by $-\log(dist(T^k x + \xi_k, z_0))$. When this is translated in the moving target case, it becomes $-\log(dist(T^k x, z_0 + \xi_k))$ and the two approaches define the same process, which give the same stationary distribution. In [90] the authors proved rigorously in some cases and numerically in others, that for a large class of chaotic maps, an extreme value law holds with an extremal index equal to 1. By the equivalence of the two processes, an extreme value law also holds in the moving target scenario and the EI is equal to 1.

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

is equal to 1 for all ε (see figure (4-B)).



(a) $z_0 = 1/2$ (fixed point)

(b) Generic z_0

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

5 Extreme value theory and global dynamical indicators

We have seen that EVT is useful to extract geometrical and dynamical information on a system, via the introduction of a suitable observable. Let us now describe a further application of EVT. It will allow us to compute the spectrum of generalized dimension associated to a chaotic dynamical system (M, T, μ), as well as a spectrum of indicators of the hyperbolicity of the system. The method can compute the spectrum D_q for integer $q \geq 2$ and is a direct generalization of the method introduced for the computation of the correlation dimension D_2 in [50]. It is based on the statistics of close encounters of q long independent orbits of the system. We introduced and described in details the method in *T. Caby, D. Faranda, G. Mantica, S. Vaienti, P. Yiou, Generalized dimensions, large deviations and the distribution of rare events [51].*

5.1 A new observable

Let us consider, for integer $q \geq 2$, the direct product $(M, T, \mu)^{\otimes q}$ with the direct product map $T_q = T \otimes \cdots \otimes T$ acting on the product space M^q and the product measure $\mu_q = \mu \otimes \cdots \otimes \mu$. We define the observable ϕ_q on M^q as:

$$\phi_q(x_1, x_2, \ldots, x_q) = -\log(\max_{i=2,\ldots,q} d(x_1, x_i)), \tag{68}$$

where each $x_i \in M$. We also write $\overline{x}_q = (x_1, x_2, \ldots, x_q)$ and $T_q(\overline{x}_q) = (Tx_1, \ldots, Tx_q)$.

The observable ϕ_q is defined in such a way that it attains high values, say larger that u_n, when all the points in the q subsystems are in the same region of phase space, that is the ball centered at a point in the first system of radius e^{-u_n}. Since x_1 can be arbitrary, in the product space M^q, the region so defined becomes a tubular neighborhood of the diagonal:

$$\Delta_n^q = \{\overline{x}_q \in M^q \text{ s.t. } d(x_1, x_2) < e^{-u_n}, \ldots, d(x_1, x_q) < e^{-u_n}\}. \tag{69}$$

Considering the complementary cumulative distribution function of the observable ϕ_q, $\bar{F}_q(u) = \mu_q(\{\overline{x}_q \in M^q \text{ s.t. } \phi_q(\overline{x}_q) > u\})$, we see that:

$$\bar{F}_q(u) = \int_{M^q} \mathbb{1}_{B(x_1,e^{-u})}(x_2)\cdots\mathbb{1}_{B(x_1,e^{-u})}(x_q)d\mu_q(\overline{x}_q) = \int_M \mu(B(x_1,e^{-u}))^{q-1}d\mu(x_1). \tag{70}$$

In the above, we factorized the $q-1$ integrals at the right hand side of the equation, by independence of the q systems. This last expression, by equation (12), has the following scaling behavior:

$$\bar{F}_q(u) \underset{u\to\infty}{\sim} e^{-uD_q(q-1)}. \tag{71}$$

This equation reveals that D_q can be evaluated by looking at the tail of $\bar{F}_q$. This can be done using EVT as we will do later; we first pursue an approach consisting in evaluating $\bar{F}_q(u)$ for different equally spaced values of u and finding the scaling with respect to large u to extract the desired quantity. Indeed, relation (71) yields:

$$\log \bar{F}_q(u) \underset{u\to\infty}{\sim} -D_q(q-1)u.$$

Therefore, a linear fitting of the empirical complementary cumulative distribution function $\bar{F}_q$ in semi-logarithmic scale allows to extract the generalized dimension of order q. This approach allows us to have a visual control on the convergence of this method.

To evaluate the value of $F_q(u)$ for different u, we run simultaneously q generic trajectories of length N, with distinct starting points given by the entries of the vector $\overline{x}_{q,0}$. Then $F_q(u)$ can be estimated by the Birkhoff sum

$$1/N \sum_{i=0}^{n-1} \mathbb{1}_{\phi_q>u}(T^q\bar{x}_{q,0}). \tag{72}$$

We test this technique with the area preserving automorphism of the torus known as the Arnol'd cat map [93], defined in the $2-$torus $\mathbb{T}^2$ by

$$\Gamma(x,y) = (2x+y, x+y) \mod 1. \tag{73}$$

The torus itself being an invariant set and the Lebesgue measure $\mathbb{L}$ being the physical invariant measure, it is easily seen that the spectrum of generalized dimension is equal to 2 for all q and we have the following theoretical result:

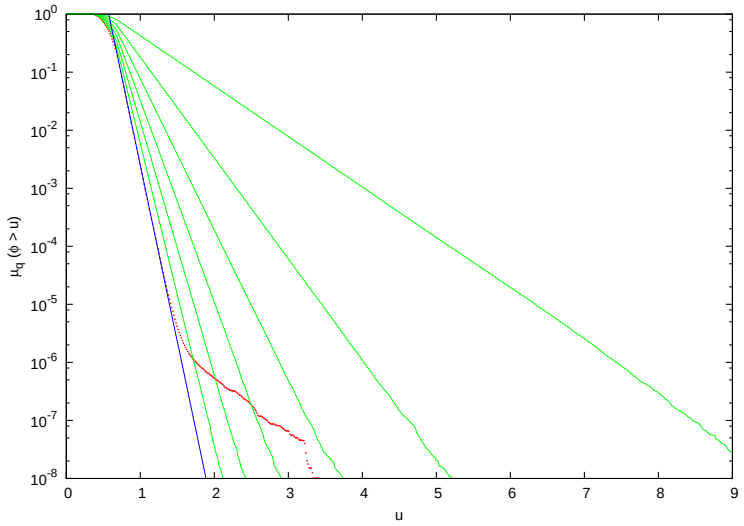


Figure 5: Probability of large events, $\bar{F}_q(u) = \mu_q(\phi_q > u)$, versus u, in the case of the Arnol'd cat. It has been estimated via a Birkhoff average over 32 trajectories of length 10^{10}. Data for $q = 2$ (highest curve) to $q = 7$ (green) and $q = 8$ (red, lowest) have been reported. The theoretical result for $q = 8$ is $\mathbb{L}(\phi_q > u) = \pi^7 e^{-14u}$ (blue).

$$\bar{F}_q(u) = \mathbb{L}(\phi_q > u) = \int_{\mathbb{T}^2} \mathbb{L}(B(x, e^{-u}))^{q-1} dx = \int_{\mathbb{T}^2} (\pi e^{-2u})^{q-1} dx = (\pi e^{-2u})^{q-1}. \tag{74}$$

This technique is tested numerically and good agreement is found with theoretical values (see figure (5)) for values of q between 2 and 8. As q gets larger, rare events associated to the observable ϕ_q become scarcer and scarcer, since asking for all the q system to be in the same small region of phase space is a strong requirement. For this reason, we are forced to perform the fit on a smaller, possibly not asymptotic range of u, leading to a discrepancy of the estimates.

To test the validity of this procedure for systems with a multifractal physical measure, we consider the Hénon map [107], that was introduced as a simplified model of a Poincaré section of the Lorenz model (that we will study later in the thesis). The discrete dynamics is defined in $\mathbb{R}^2$ as follows:

$$T(x, y) = (1 - ax^2 + y, bx). \tag{75}$$

For the standard values $a = 1.4$, $b = 0.3$ and a starting point in the basin of attraction, the dynamics create an attractor with multifractal properties suited to our investigation (see figure (6) for a numerical representation). Evidence of

multifractality of this measure along with extensive numerical determinations are presented in the literature, for example in [18, 29, 95]. We stress that we lack a formal proof that the Hénon system has REPFO properties and that the spectral method can effectively be applied. However, as in the case of the cat map, the proposed method provides reliable results for small values of q (things getting worse for large q as discussed before) and are very close to the values found by extensive computations based on box-counting techniques in [95]. In particular, we find the values $D_2 = 1.2$, $D_3 = 1.16$ and $D_4 = 1.1$ (see figure (7)).

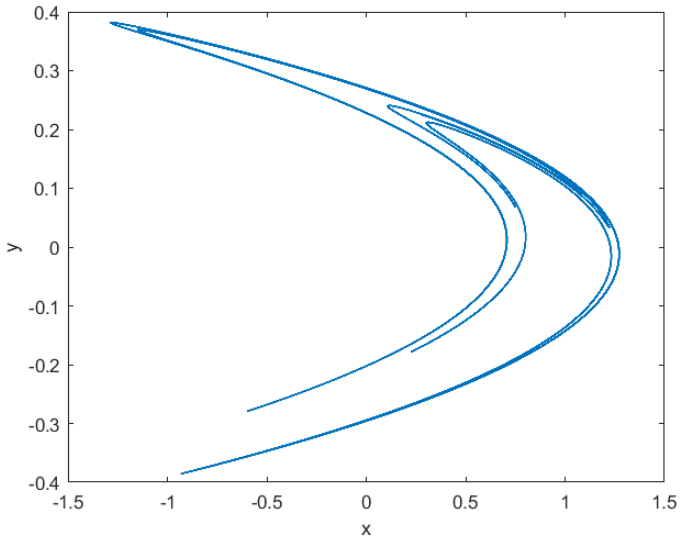


Figure 6: Numerical reconstruction of the Hénon attractor, using 10000 points.

5.2 Generalized dimensions and the Block Maxima approach

We have seen that the tail of the distribution $\bar{F}_q(u)$ contains information on the value D_q. Studying this tail is the core of EVT and it can be done for example following the block maxima approach, described in section (3). Let us consider the maximum value attained by the function ϕ_q over a *block of times* of length n:

$$M_n^q(\overline{x}_q) = \max\{\phi_q(\overline{x}_q), \ldots, \phi(T_q^{n-1}(\overline{x}_q)\}, \tag{76}$$

and its distribution function $G_{q,n}(u)$:

$$G_{q,n}(u) = \mu_q(\{\overline{x}_q \in M^q \text{ s.t. } M_n(\overline{x}_q) \leq u\}). \tag{77}$$

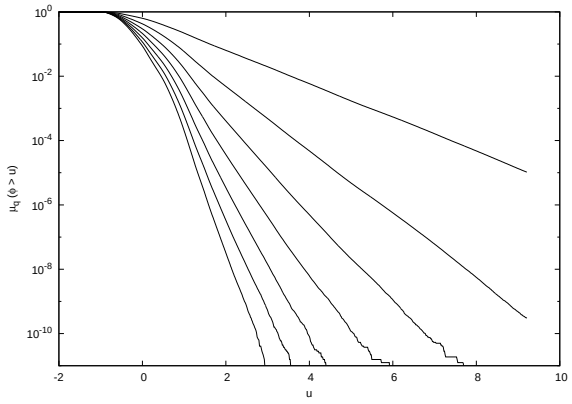


Figure 7: Probability of large events, $\bar{F}_q(u) = \mu_q(\phi_q > u)$, versus u, in the case of the Hénon attractor physical measure. It has been estimated via a Birkhoff average over 32 trajectories of length 10^{10}. Data for $q = 2$ (highest curve) to $q = 8$ (lowest) have been reported.

To follow the common procedure, it is required that for all $s > 0$, there exists a sequence of levels $u_n(s)$ such that:

$$\bar{F}_q(u_n(s)) \underset{n\to\infty}{\sim} \frac{s}{n}, \tag{78}$$

and such that the distribution of the maximum is non degenerate. Equation (71) implies an exponential tail for the complementary observable distribution $\bar{F}_q$. Such a flat tail implies, as discussed in section (3), the convergence of $F_{q,n}$, suitably rescaled, to a Gumbel law:

$$G_q(s) = e^{-\theta_q s}. \tag{79}$$

We call the quantity θ_q the *dynamical extremal index DEI* of order q and it will be studied in the next section. This convergence can be investigated numerically and can provide an estimate of the generalized dimensions. In fact, combining relations (71) and (78), we have that

$$\bar{F}_q(u_n) \underset{n\to\infty}{\sim} e^{-u_n D_q(q-1)} \underset{n\to\infty}{\sim} \frac{s}{n}, \tag{80}$$

so that

$$u_n \underset{n\to\infty}{\sim} \frac{-\log s}{D_q(q-1)} + \frac{\log n}{D_q(q-1)} = \frac{-\log s}{a_n} + b_n. \tag{81}$$

As mentioned in section (3), a_n converges to the scale parameter σ in equation (9) from which, by the previous equation, D_q can be extracted. The parameters ξ, σ and β_n of the GEV are computed via the maximum likelihood estimates provided

by the Matlab function *gevfit*, for which detailed information is found at the web page [98]). In figure (8) we show an example of the best fit obtained in the case of the observable ϕ_3 for the $3x$ (mod 1) map.

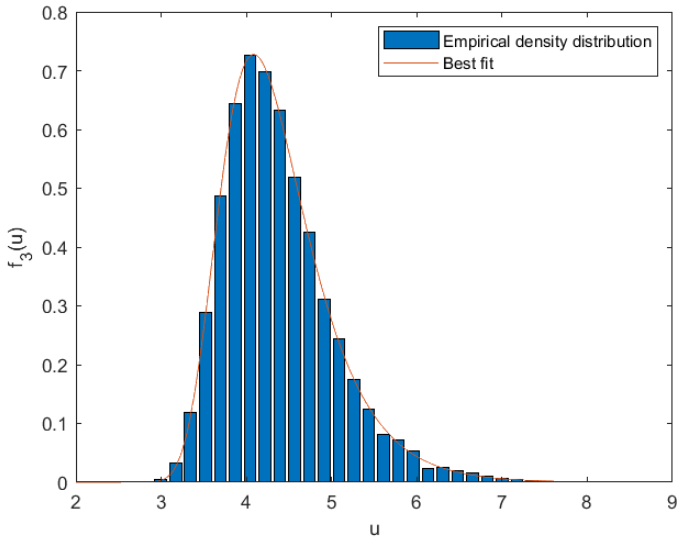


Figure 8: Best fit of the empirical density distribution of the maximum associated to the observable ϕ_3 and the map $3x$ (mod 1). We used trajectories of 10^7 points and a blocksize of 1000. The estimate for the shape parameter is as expected very close to 0 (≈ -0.002) and the one for the scale parameter 0.502 yields an estimate for D_3 of 0.996 that compares favorably with the theoretical value of 1.

For the numerical computations that will be performed next, we use the following procedure:

1-Generation in parallel of q distinct long trajectories of N points and, at each iteration, computation and storage of the observable $\phi_q(x_1, ..., x_q)$

2-Division of time into blocks of length n and extraction of the maximum value of μ_q from each block.

3-Fit of the obtained empirical distribution of the maxima, and maximum likelihood estimation of the GEV parameters.

4-Computation of $\hat{D}_q = \frac{1}{\hat{\sigma}(q-1)}$, where $\hat{\sigma}$ is the estimate of the scale parameter obtained by the fit.

5-Averaging the result over l trials.

We now give a Matlab code of this procedure for the map $3x \mod 1$:

```
clear all

qmax=10; %maximum value of q considered
M=10000000; %length of the simulated trajectories
blocklength=10000; %blocklength
logextr=zeros(floor(M/blocklength),qmax); %vector
    countaining the maxima
ntraj=20;%number of trajectories over which the estimates
    are averaged

for l=1:ntraj
ind=0;
    %Initialization of the q trajectories
for q=1:qmax
    x(q) = rand;
end

maxx=-100*ones(qmax,1);

for i=1:M
    mm=0;

    for q=1:qmax
   x(q)=mod(3*x(q),1); %Iteration of the dynamic

    if  abs(x(1)-x(q))>mm
        mm=abs(x(1)-x(q));
    end

    %The maximal value of the observable over the current
        block is stored
    if q>1
    if -log(mm)>maxx(q)
        maxx(q)=-log(mm);
    end
```

```
33      end
34
35      end
36
37 %The maximal value of the observable over the block is
       stored in logextr
38
39    if mod(i,blocklength)==0
40          ind=ind+1;
41       logextr(ind,:)=maxx;
42    maxx=-100*ones(qmax,1);
43    end
44
45 end
46
47
48  for q=2:qmax
49 % fit of the empirical distributions of maximum
50 [tpar,tpari]=gevfit(logextr(:,q));
51 %Estimation of Dq for the trajectory l
52 d(q,l)=1/((q-1)*tpar(2));
53
54  end
55 end
56
57 %Averaging of the result and computation of the uncertainty
58 for q=2:qmax
59 Dq(q)=mean(d(q,:));
60 uncertainty(q)=std(d(q,:));
61 end
```

We apply this procedure to the case of an Iterated Function System measure [49] on the Sierpinski gasket, defined by the stochastic process on the unit square $M = [0,1]^2$ realized by iteration of the maps f_i, $i = 1,2,3$ chosen at random with

probability p_i:

$$\begin{cases} f_1(x,y) = (x/2, (y+1)/2),\ p_1 = 1/4, \\ f_2(x,y) = ((x+1)/2, (y+1)/2),\ p_2 = 1/4, \\ f_3(x,y) = (x/2, y/2),\ p_3 = 1/2. \end{cases} \tag{82}$$

We use the parameters $N = 2{\cdot}10^8$, $n = 5000$, $l = 20$. In the left panel of figure (10), the numerically obtained generalized dimensions are compared with the analytical values [39] equal to

$$D_q = \frac{\log_2(p_1^q + p_2^q + p_3^q)}{1-q}. \tag{83}$$

Good agreement is found for small values of q, which later worsens. Indeed, as discussed earlier, the scarcity of close encounters increases as q gets larger; moreover, since we are considering time blocks of fixed length, events that may not be considered as extremes are selected in each block and the distribution of the maxima does not reach its asymptotic form, leading to poor estimates for large q.

We now test the method in the context of continuous time systems: we compute the D_q spectrum for the well known Lorenz 1963 model [2], a continuous–time dynamical system reconstructed by the Euler method with step size 0.013. The dynamic is defined by the following set of differential equations:

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(r - z) - y \\ \dot{z} = xy - \beta z, \end{cases} \tag{84}$$

As for the Hénon system, we have no formal evidence that the spectral theory can be applied to the Lorenz system.

where $\beta = 8/3$, $r = 28$ and $\sigma = 10$. To apply the method to this system, we consider blocks of 10^4 units of time and the results are obtained with trajectories of 10^8 points. Dimensions estimates, reported in figure (10), are obtained by averages over 20 trajectories, uncertainties being the standard deviations of these results. The estimate for D_2 (≈ 2.025) is slightly inferior to the value of 2.05 found in other studies [17], and a lack of well established estimates of D_q for larger q in the literature makes comparisons difficult (see however [96] for some computations and discussions).

Despite its limited range of application to positive, integer and not too large q, the main advantage of the proposed method, in opposition to box counting tech-

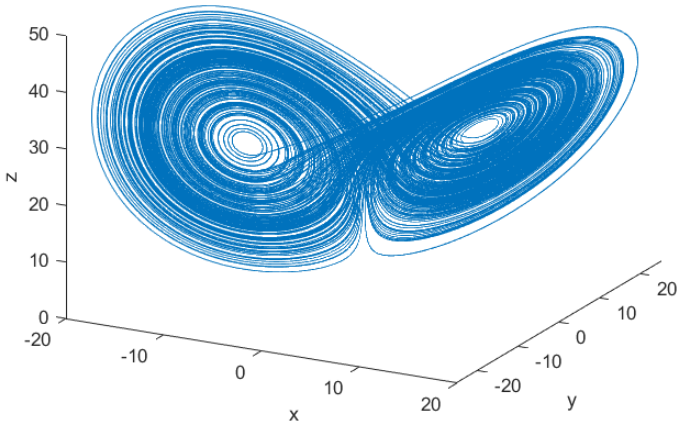


Figure 9: Numerical reconstruction of the Lorenz attractor, using the Euler method with a step $h = 0.013$ and 5.10^4 points.

niques, is that it seems suited to the study of high dimensional dynamical systems and could possibly be applied to time series originating from many dimensional natural systems, provided enough data are at disposal. In the case where only one observational time series is at disposal, and with the assumption that the underlying system has strong enough decay of correlations, q distinct trajectories can be obtained by taking spaced enough starting points within the signal.

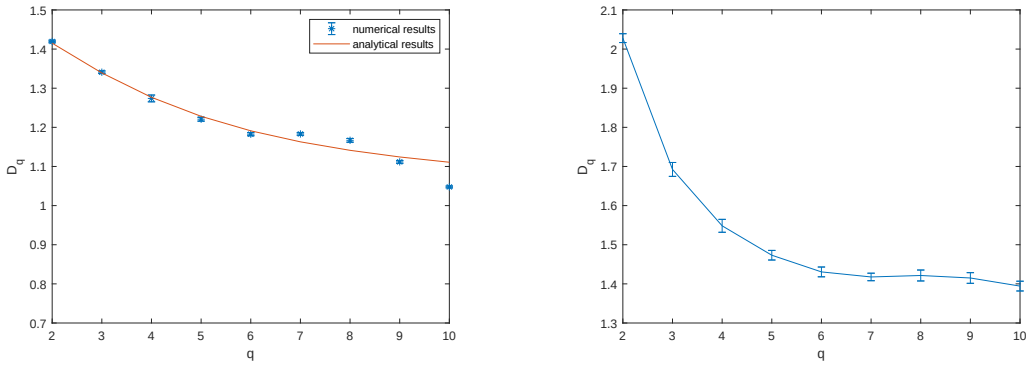


Figure 10: Left panel: numerical estimates of D_q for the Sierpinski gasket (blue symbols) and theoretical value (red curve). Right panel: numerical estimates of D_q for the Lorenz 1963 model. The uncertainty is the standard deviation of the results obtained over 20 trajectories. See text for parameters and discussion.

5.3 The dynamical extremal index and the hyperbolicity index of order q

We mentioned earlier that the extremal index associated with the distribution of maxima gives precious information on the hyperbolicity of the system. In [50], the authors relate the extremal index associated to the observable ϕ_2, that they called the *Dynamical Extremal Index* (*DEI*), to the Lyapunov exponent of the system for one-dimensional systems, and to the entropy for systems of higher dimension.
Let us now study in detail the parameter θ_q appearing in the exponent of the Gumbel law (79). In this context, formula (18) reads

$$\theta_q = 1 - \sum_{k=0}^{\infty} p_{q,k}, \tag{85}$$

where

$$p_{q,k} = \lim_{n\to\infty} \frac{\mu(\phi_q \geq u_n \cap \max_{i=1,\dots,k} \phi_q(T^i x) < u_n \cap T^{k+1}x \geq u_n)}{\mu(\phi_q \geq u_n)}. \tag{86}$$

We have that $p_{q,k}$ can be written as:

$$p_{q,k} = \lim_{n\to\infty} \frac{\mu(\Delta_n^q \cap (T_q^{-1}\Delta_n^q)^c \cap \dots \cap (T_q^{-k}(\Delta_n^q))^c \cap T_q^{-k-1}(\Delta_n^q))}{\mu(\Delta_n^q))}. \tag{87}$$

It is expected that, for systems with exponential decay of correlations, the values of $p_{q,k}$ for $k > 0$ are all equal to 0. Indeed, the distance between the trajectory in the first subsystem and the trajectories in the other $q-1$ subsystems grow exponentially over time, so that one expects that if at least one of the $q-1$ subsystem is at a given time outside of the ball centered at a point of the first system, it will not come back to it after a few iterations of the map. In fact, the mixing properties of the system are such that the dynamics will come back to Δ_n^q (by the Poincaré recurrence theorem), but after a time that grows asymptotically with n and in an irregular way. This heuristic argument suggests that the $p_{q,k}$ for $k \geq 1$ are all equal to 0. This result has indeed first been proposed for mixing piece-wise expanding maps of the interval with invariant measures absolutely continuous with respect to Lebesgue in the case $q = 2$ in [37], then extended to larger q and proven rigorously in [99]. In the latter, the proof is made in the context of coupled map lattices under some assumptions, namely that the map T is C^2 by parts, that the absolute value of the derivative is always greater than 1 (uniformly hyperbolic system) and under some (not so restrictive) spectral properties needed to use the spectral approach (assumptions P1 to P5 and P8 in [99]). In this set up, the authors were also able

to give a formula for θ_q that depends on the derivative T' of the map, the density h of the measure and the coupling factor of their map-lattices (see their proposition (5.5)). By adapting slightly the formula, taking a coupling factor equal to 0 to yield an independent evolution of the subsystems, we get the formula:

$$\theta_q = 1 - p_{q,0} = 1 - \frac{\int \frac{h(x)^q}{|DT(x)|^{q-1}} dx}{\int h(x)^q dx}. \tag{88}$$

We now test numerically this formula for diverse maps of the interval satisfying the required assumptions. A direct application of formula (88) to the map $3x$ (mod 1) gives the result $\theta_q = 1 - 3^{1-q}$, and indeed very good agreement is found between numerical simulations and predictions (see figure (11)). In this situation, only p_0 participates to the value of the extremal indices, so that the order 0 Süveges estimator can be used. For each q, we took as a threshold the 0.997-quantile of the distribution $F(u)$ computed on a pre-runned trajectory of 10^6 points.

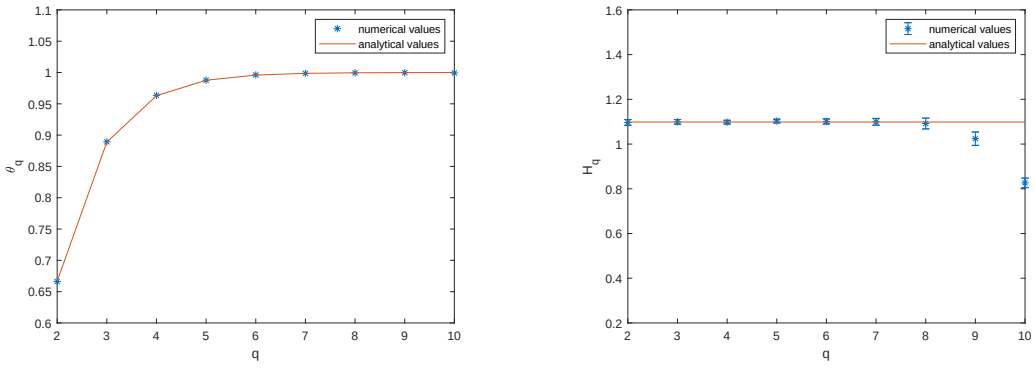
The fact that the derivative of the map is present in the formula reflects the fact that θ_q carries information on the hyperbolicity properties of the system. This statement will be made more concrete after the following definition.

Definition 14 *We call the hyperbolicity index of order $q \geq 2$ the quantity:*

$$H_q = \frac{\log(1 - \theta_q)}{q - 1}. \tag{89}$$

For maps of the interval with constant derivative, like $3x$ (mod 1), we observe that this quantity is constant equal to $\log |DT|$, which is exactly the Lyapunov exponent of the system [53]. This is checked numerically in figure (11), where good agreement is found for small q. For larger q, estimates get less reliable for two main reasons: as discussed earlier for the determination of the generalized dimensions, collecting extreme events becomes more and more difficult as q gets large; it is also likely that since the quantity $\log(1 - \theta_q)$ diverges as q tends to infinity, this yields less precise results (we observe that the error bars get wider for large q).

When the derivative is not constant, a non trivial spectrum H_q is found that accounts for regions of different hyperbolic behavior in the attractor. We stress that the quantity H_q depends both on the derivative of the map and on the density. It was stated in [50] that for maps whose derivative does not vary too much, the quantity H_2 is a fair approximation of the Lyapunov exponent of the system. On



(a) θ_q (b) H_q

Figure 11: Indicators θ_q and H_q of the map $x \mapsto 3x \pmod 1$, obtained by averaging the results over 20 trajectories of $2 \cdot 10^7$ points. As only $p_{q,0}$ contributes to the computation, we used the order 0 Süveges estimator for θ_q. The uncertainty is the standard deviation of the results. See text for further discussion.

the other hand, a look at equations (88) and (89) tells us that for high values of q, the value of H_q is principally determined by regions of phase space with either low absolute value of the derivative, or high values of the density, or both. This suggests that the H_q spectrum contains information on the different hyperbolic behaviors of the system. This spectrum of indicators, which appears naturally in the method can be extended, at least for maps of the interval for which equation (88) holds, to real values of q and to cover all real values except $q = 1$, for which H_q can be extended by continuity applying l'Hospital rule, just like for the D_q spectrum. We find that

$$H_1 = \lim_{q \to 1} H_q = \lim_{q \to 0} -\frac{d}{dq} \log \int \frac{h(x)}{|DT(x)|^q} dx, \tag{90}$$

which, by applying Leibniz integral rule, equals

$$H_1 = \lim_{q \to 0} -\int \frac{d}{dq} \frac{h(x)}{|DT(x)|^q} dx = \int \log |DT(x)| h(x) dx. \tag{91}$$

We find that H_1 is the Lyapunov exponent of the system. As mentioned in [50], the value H_2 is close to the Lyapunov exponent when the derivative does not change too much. Indeed, by continuity of the H_q spectrum, the values of H_1 and H_2 are close, and the difference is highly influenced by the variations of $|DT|$. For the extended H_q spectrum, it is easily seen that very negative q characterize regions of low density and/or high absolute value of the derivative. Once this extension is made, the spectrum of H_q contains information on the dynamics over the whole attractor.

To illustrate these facts, let us consider a map of the interval with non constant and absolutely continuous density and non constant derivative: the Gauss map, defined as $x \to \frac{1}{x} \pmod 1$. The density is known and is equal to $h(x) = \frac{1}{\log 2(1+x)}$ [5].

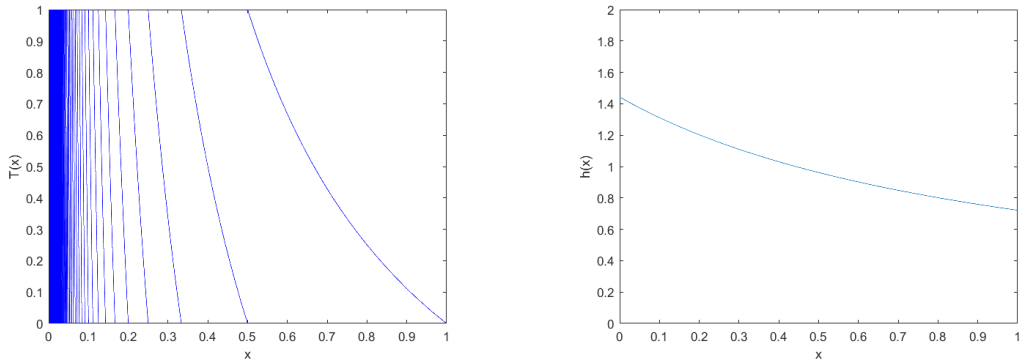
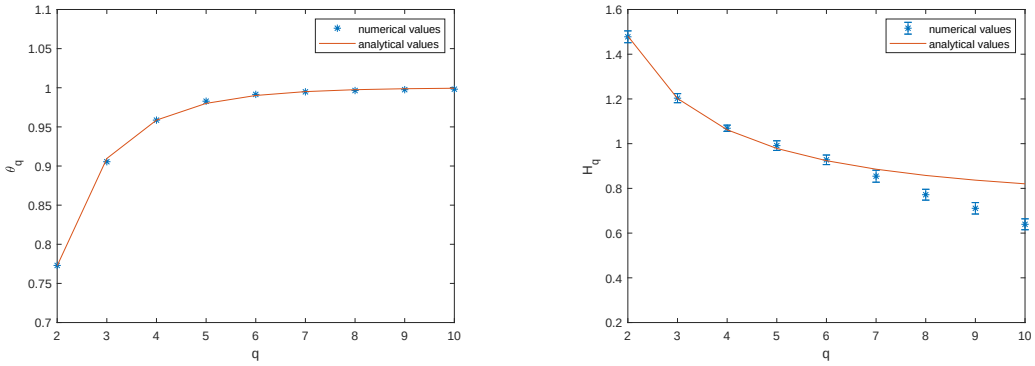


Figure 12: Graphical representation of the Gauss map (left) and its associated invariant density (right).

Analytical values for θ_q can be computed from formula (88) and give:

$$\theta_q = 1 - \frac{[\sum_{k=0,k\neq q-1}^{2(q-1)} (-1)^k \binom{2(q-1)}{k} \frac{2^{k-q+1}-1}{k-q+1}] + (-1)^{q-1} \binom{2(q-1)}{q-1} \log 2}{\frac{2^{1-q}-1}{1-q}}. \tag{92}$$

It can be seen in figure (13) that good agreement is found between theory and simulations for the θ_q spectrum. As discussed earlier, estimates of H_q for large q are less reliable.



(a) θ_q (b) H_q

Figure 13: θ_q and H_q of the Gauss map absolutely continuous invariant measure. Parameters of the numerical estimation are as in figure 11.

An explicit computation of θ_q can be performed for other maps of the interval for which we know density and derivative. This is the case of the *piece-wise linear*

Markov maps defined as [5]:

$$T(x) = \begin{cases} T_1(x) = 3x \text{ if } x \in I_1 = [0, 1/3), \\ T_2(x) = 5/3 - 2x \text{ if } x \in I_2 = [1/3, 2/3), \\ T_3(x) = -2 + 3x \text{ if } x \in I_3 = [2/3, 1). \end{cases} \tag{93}$$

Spectral techniques introduced in [5] allow to compute the piece-wise constant density:

$$h(x) = \begin{cases} h_1 = 3/5 \text{ if } \quad x \in I_1, \\ h_2 = 6/5 \text{ if } \quad x \in I_2, \\ h_3 = 6/5 \text{ if } \quad x \in I_3. \end{cases} \tag{94}$$

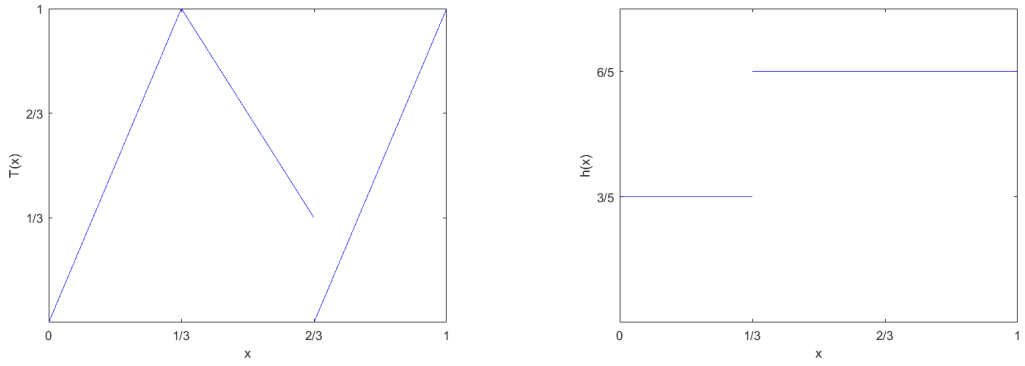


Figure 14: Graphical representation of the Markovian map (left) defined by equation (93) and its associated invariant density (right).

θ_q can be easily computed by formula (88) and it reads:

$$\theta_q = 1 - \frac{\frac{h_1^q}{(T_1')^{q-1}} + \frac{h_2^q}{(T_2')^{q-1}} + \frac{h_3^q}{(T_3')^{q-1}}}{h_1^q + h_2^q + h_3^q}. \tag{95}$$

For this map, the spectrum of H_q can be extended explicitly to negative q, and it is easily seen that for large q, H_q converges to the logarithm of the minimum of the absolute value of the derivative of the map over phase space; the limit when q goes to $+\infty$ is the maximum of the logarithm of the absolute value of the derivative (see figure (15) for a graphical representation of this spectrum).

We have so far stated examples for which the spectral theory can be applied. It is in practice very difficult to check that the system has REPFO properties in dimension more than 1. We will therefore suppose that such conditions hold and we will now discuss the interpretation of our indicators H_q for higher dimensional systems. It was stated in [50] that the quantity H_2 relates to the metric entropy of the system,

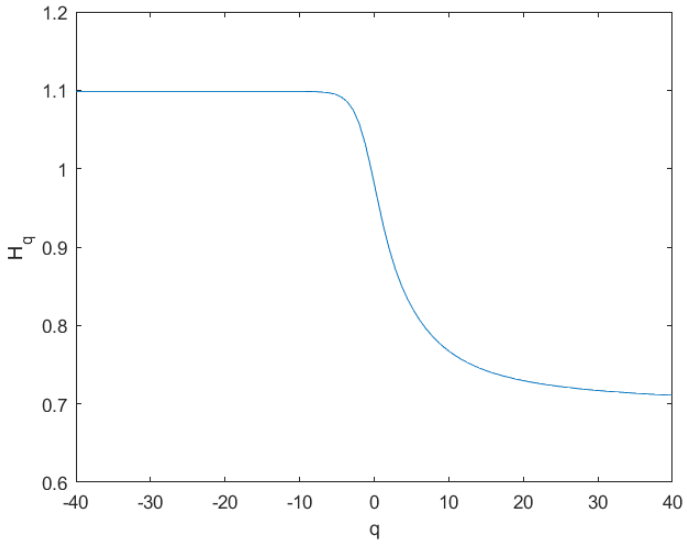


Figure 15: Graphical representation of the extended H_q spectrum for the Markovian map described in the text. The limit as q goes to $+\infty$ is $\log 2 \approx 0.693$ and $\log 3 \approx 1.099$ as q goes to $-\infty$, that correspond to, respectively, the logarithms of the minimum and the maximum of the absolute value of the derivative.

that by Pesin formula [87] is equal to the sum of the positive Lyapunov exponents. As for uni-dimensional systems with constant derivative, we expect that for maps with constant Jacobian norm, the quantity H_q does not vary with q. We test this hypothesis with the Lozi map, a map of constant Jacobian norm introduced in the seventies by Lozi [103] as a kind of lineralized version of the Hénon map, for which the explicit computation of periodic orbits is possible. The map is defined in $\mathbb{R}^2$ by:

$$\begin{cases} x_{n+1} = -a|x_n| + y_n + 1 \\ y_{n+1} = bx_n \end{cases} \tag{96}$$

For a range of parameters including the values $a = 1.7$ and $b = 0.5$ (used in our simulations), the dynamics tend to evolve on a strange attractor (see figure (16) for a numerical reconstruction). Since the hyperbolic properties of this map do not depend on the region of phase space, we expect for this map a constant spectrum of H_q, equal to the largest Lyapunov exponent of the system, which is approximately 0.47 [92]. This is indeed in agreement with the numerical simulations presented in figure (17), which shows small systematic fluctuations around this value for some q, probably due to the role of the contracting subspace. Note that again, a discrepancy of the results is observed for values of q greater than 8.

For maps with non constant Jacobian, we expect, as for uni-dimensional systems

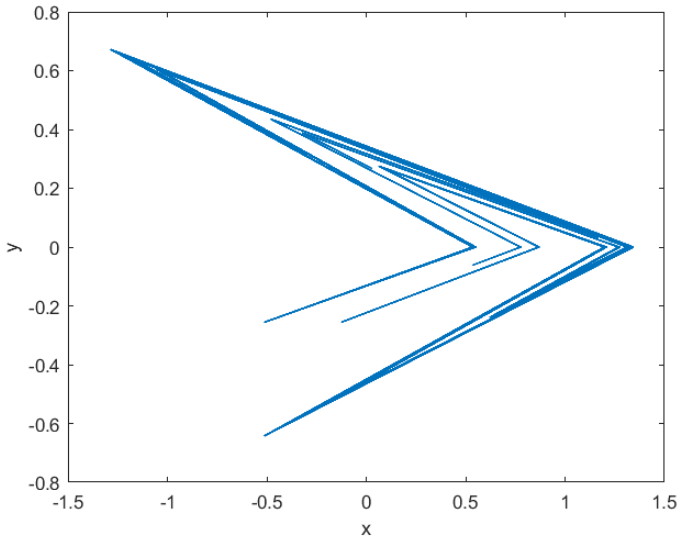
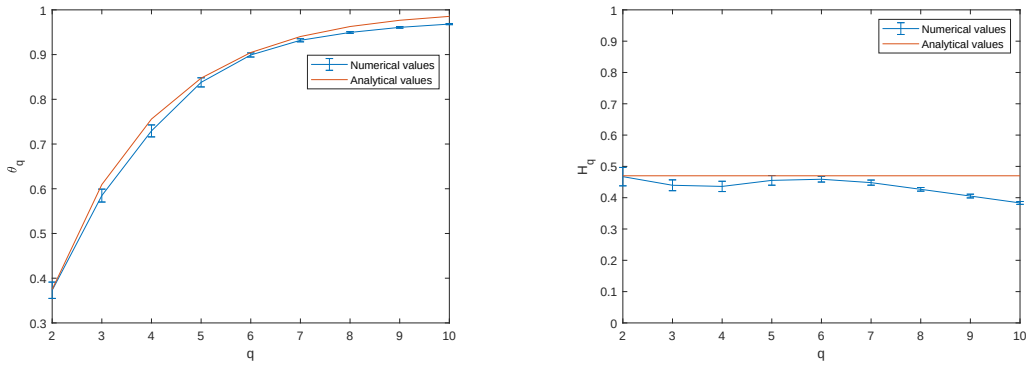


Figure 16: Numerical reconstruction of the Lozi attractor, using 10^5 points.



(a) θ_q (b) H_q

Figure 17: θ_q and H_q of the Lozi map. We simulated trajectories of 5.10^7 points and used the Suveges estimate with a threshold corresponding to the $0.997-$quantile of the observable distribution. Results are averaged over 20 trajectories, the uncertainty being the standard deviation of the results. The red curves represent $H_q = 0.47$ for the right panel and $\theta_q = 1 - A^{q-1}$, with $A = 0.47$ for the left panel.

with non constant derivative, a non trivial variation of H_q with q.

Now that we have discussed in details the parameters appearing in the extreme value law associated to the observable ϕ_q, let us discuss the implications of this result. The quantity $M_n^q(x_1, ..., x_q)$ can be interpreted as the "closest encounter" between q trajectories of length n starting from the points $x_1, ...x_q$. Its asymptotic probability distribution is, for the class of systems considered here, a Gumbel law, whose parameters are determined entirely by:

- The length n of the trajectories,

- D_q: the generalized dimension of order q,

- θ_q: the dynamical extremal index of order q.

In a recent work [74], the asymptotic behavior of the shortest distance between two orbits was studied. This constitutes a different problem to the one of the closest encounter between two trajectories, because shifts are allowed in this situation. For the two problems, the asymptotic behavior is governed by the correlation dimension D_2, although when we do not allow shifts, the full asymptotic distribution can be obtained. An extension to the case of more than 2 orbits could be made, the asymptotic behavior of the shortest distance between q orbits would then be governed by the generalized dimension of order q. These problems have, as pointed out in [74], interesting connections with the problem of the size of the longest substring common to different chains of characters.

5.4 Influence of noise on the global dynamical indicators

It is now natural to investigate the influence of a random perturbation on our estimates, as in many natural phenomena observational noise is present, either due to the nature of the dynamics itself, or to some observational noise associated to imprecision in the measurement process.

5.4.1 Influence of uniform noise on the D_q spectrum

Let us first investigate the influence of some uniform noise on the estimates. By uniform noise, we mean that, as in section (4.4.1), at each iteration, to the original map is added a small perturbation $\varepsilon\xi$, where ξ follows a probability law $\mathbb{P}$ that is uniform in the ball $B(0,1)$. While doing so, one must be careful that the noise does not drive the dynamics away from the basin of attraction of the attractor. In such cases the dynamics can diverge, making the method inapplicable. This is typically what happens when we perturb the motion on the Lorenz or the Hénon attractor with some uniform noise. When the system lives in a smooth compact manifold the dynamics cannot diverge, and it is expected that adding a small perturbation of uniform type will make the dynamics 'spread' and destroy the fine structure of the

attractor. The local dimensions (and as a consequence the generalized dimensions) are therefore expected to approach the dimension of the ambient space. This is indeed what we observe in figure (18), where an IFS dynamics on the common Cantor set (similar to the one on the Sierpinski gasket defined in equation (97)) is perturbed by a uniform noise of intensity ε. The dynamics is defined by the stochastic process on the unit interval $M = [0, 1]$ realized by iteration of the maps f_i, $i = 1, 2$ chosen at random with probability $1/2$:

$$\begin{cases} f_1(x) = x/3, \\ f_2(x) = (x+2)/3. \end{cases} \tag{97}$$

The generalized dimensions, that are all equal to the constant value of the local dimension $\frac{\log 2}{\log 3} \approx 0.63$ of the unperturbed system (see for example to [53]), all shift to the Hausdorff dimension of the ambient space which is 1 when the intensity of noise increases. As it can be observed in figure (18), this convergence is much slower for higher values of q. We interpret this fact as follows: some regions of phase space are left by chance relatively unperturbed by the noise and the local structure is preserved around these points. For large q, the generalized dimensions characterize regions of low local dimensions, so that the spectrum is here determined mostly by points having preserved their local fractal structure. When the intensity of the noise increases, however, less local structures are left unaffected by noise and we observe a convergence of all D_q's to the dimension of the ambient space.

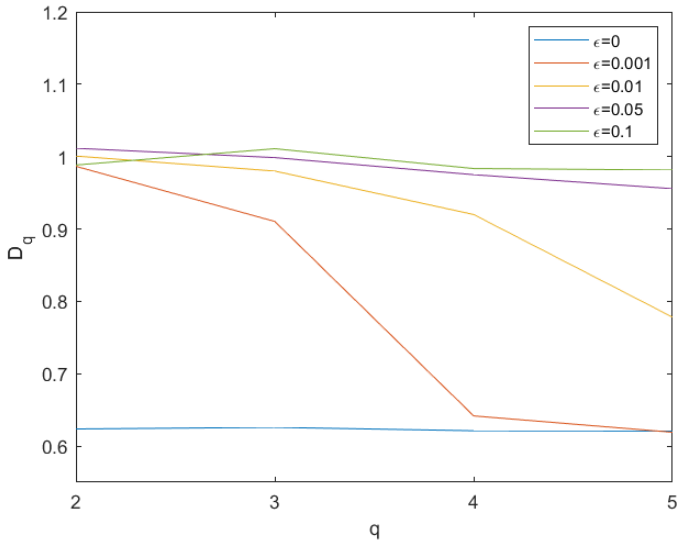


Figure 18: Influence of uniform noise on the computed D_q spectrum associated to the dynamics on the Cantor set for different intensities of noise. We simulated trajectories of 2.10^8 points, and used blocks of length 50000.

5.4.2 Influence of uniform noise on the DEI spectrum

The influence of uniform additive noise on θ_2 has been investigated in [50]. The authors found, for different maps and with the kind of perturbations considered in the last section, that the numerical estimates of θ_2 tend to 1 as the noise increases. Indeed, the diagonal in the product space:

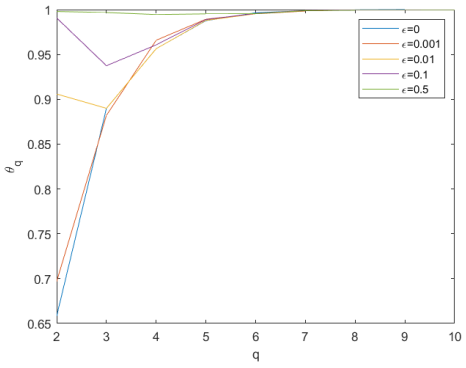
$$\Delta^2 = \bigcap_{n\in\mathbb{N}} \Delta_n^2, \tag{98}$$

that is invariant by action the deterministic map T is no more invariant in the perturbed case. As $p_{2,0}$ in equation (88) measures the proportion of points staying in Δ_n^2 after one iteration of the map in the limit of large n, in the presence of noise this quantity tends to 0, because the perturbation separates the iterates. This reasoning can be generalized to larger q and we expect all the quantities θ_q to turn to 1 in the presence of noise. This is indeed what is observed in figure (19), in which we test these considerations for two maps of the circle: $3x \pmod 1$ and the Gauss map. For both, trajectories of 10^7 points were simulated and the $0.999-$quantile of the distribution of the observable ϕ_q was selected as a threshold. We choose the Süveges estimate for our computations, otherwise we are confronted to the following issue: theoretically, at least for the unperturbed case, all the $p_{q,k}$ but $p_{q,0}$ are 0, but (especially for large q) some of the computed $\hat{p}_k$ are small but different from 0. This is a numerical effect, probability due to the difficulty of the computation of the DEI for large q. Adding these small errors would increase significantly the global error and we want to avoid this fact. This effect surely affects the computation of this indicators for real world stationary time series with noisy underlying dynamics. As in section (4.4.1), we will see that with suitable discrete noise, the dynamical extremal indices can be different from 1 and can sometimes be computed explicitly.

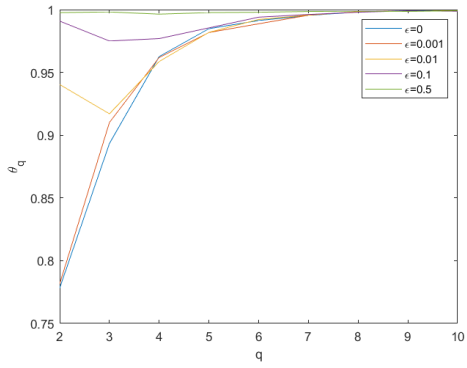
5.4.3 Influence of discrete noise for the DEI spectrum

A different scenario happens if we look at sequences of finitely many maps chosen in an i.i.d. way according to a Bernoulli process as we did in section (4.4.1). We choose the two maps $T_0 = 3x \pmod 1$ and $T_1 = 3x + b \pmod 1$. A combination of formulae (53) and (90) gives immediately the term $p_{q,0}$ entering the infinite sum defining the DEI:

$$p_{q,0} = \lim_{n\to\infty} \frac{\int_\Omega \rho_q(\Delta_n^q \cap T_{\xi^q}^{-1}\Delta_n^q) d\mathbb{P}^q(\xi)}{\rho_q(\Delta_n^q)}, \tag{99}$$



(a) $3x \mod 1$



(b) $\frac{1}{x} \mod 1$ (Gauss map)

Figure 19: Influence of the intensity of noise ϵ on the θ_q spectrum for maps of the circle perturbed by uniform noise. We used the Süveges estimate, as mentioned in the text.

where: ξ^q is the vector $\xi^q = (\xi_1, \cdots, \xi_q)$; $T_{\xi^q}^{-1} = T_{\xi_1}^{-1} \cdots T_{\xi_q}^{-1}$ and ρ_q is the stationary measure constructed as in [59]. We now split the integral over the cylinders $\xi^{q,0} = [\xi_1 = 0, \cdots, \xi_q = 0], \xi^{q,1} = [\xi_1 = 1, \cdots, \xi_q = 1]$ and their complement. Therefore we have

$$\frac{\int_\Omega \rho_q(\Delta_n^q \cap T_{\xi^q}^{-1}\Delta_n^q) d\mathbb{P}^q(\xi)}{\rho_q(\Delta_n^q)} = \mathbb{P}(\xi^{q,0}) \frac{\rho_q(\Delta_n^q \cap T_0^{-n}\Delta_n^q)}{\rho_q(\Delta_n^q)} + \mathbb{P}(\xi^{q,1}) \frac{\rho_q(\Delta_n^q \cap T_1^{-n}\Delta_n^q)}{\rho_q(\Delta_n^q)} + \tag{100}$$

$$\{\text{the integral computed over all the other cylinders of length } q\}.$$

The two fractions in (100) are equal to the $p_{q,0}^{(\text{unp})}$ of the unperturbed systems, which are the same by the particular choices of the maps and by formula (88).
All the other terms in (100) are zero. Let in fact consider a vector $\tilde{\xi}^q$ different from $(0, 0, ..., 0)$ and $(1, 1, ..., 1)$. Suppose for instance that the first coordinate $\tilde{\xi}_1^q$ of $\tilde{\xi}^q$ is zero and let $\tilde{\xi}_j^q \neq 0$ for some $1 < j \leq q$. If $x_q \in \Delta_n^q$, and since $T_0 x_1 = 3x_1 \pmod 1$ and $T_1 x_j = 3x_j + b \pmod 1$, we have that $d(T_0 x_1, T_1 x_i) \to_{n\to\infty} \min\{b, 1-b\}$ as $d(x_1, x_i) < e^{-u_n}$.
Therefore for n large enough, $\forall x_q \in \Delta_n^q$, $T^{\tilde{\xi}^q} x_q := T^{\tilde{\xi}_q^q} \cdots T^{\tilde{\xi}_1^q} x_q \notin \Delta_n^q$ and $\Delta_n^q \cap T_{\tilde{\xi}^q}^{-1} \Delta_n^q = \emptyset$. In conclusion, we have that

$$p_{q,0} = p_{q,0}^{(\text{unp})}(\rho^q + (1-\rho)^q), \tag{101}$$

where ρ is the weight of the Bernoulli measure. For similar reasons as described earlier, all the other values of $p_{q,k}$ are 0 and we have the result:

$$\theta_q = 1 - p_{q,0}^{(\text{unp})}(\rho^q + (1-\rho)^q). \tag{102}$$

We now test this last formula numerically with the maps $T_0(x) = 3x \pmod 1$ and $T_1(x) = 3x + b \pmod 1$, for different values of b and with $\rho = 0.5$. Very good agreement is found for different values of b (see figure (24)). We use trajectories of 5.10^7 points and the $0.995-$ quantile of the observable distribution as a threshold. Estimations are made using the Süveges estimate, for the same reasons as described earlier.

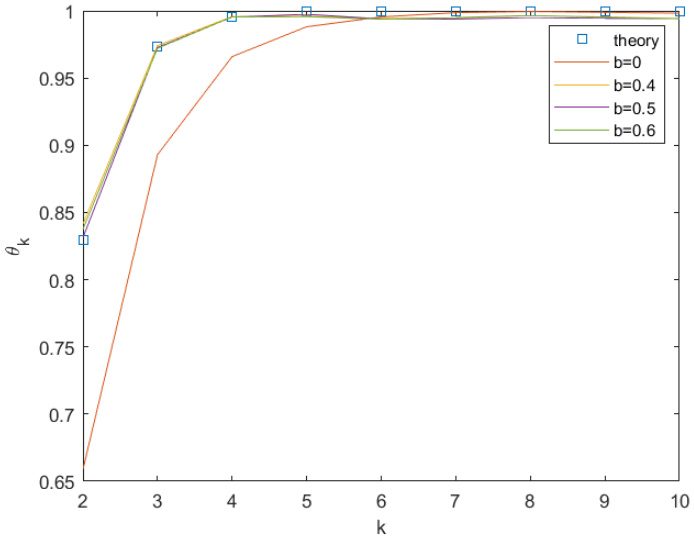


Figure 20: θ_q spectrum for the sequence of two maps described in the example, found using the Süveges estimate. The red curve is the value for $b = 0$, and for different choices of $b \neq 0$, the values of θ_q follow closely the predictions of formula (102) (blue squares).

6 Extreme value theory and the statistics of visits in small target sets

We have seen in section (3.3) that EVT allows to give asymptotic laws for the number of visits in the target sets Ω_n. We here investigate this matter for different target sets and for non-deterministic dynamical systems, for which only few results are known. This problem is of course of crucial importance in many physical applications, since it allows to foresee, for example, the number of occurrence of particular states of the system that are close to a particular, critical state. The results presented in the next subsections can be found in the publication *Caby et al. On the computation of the extremal index for time series [112].*

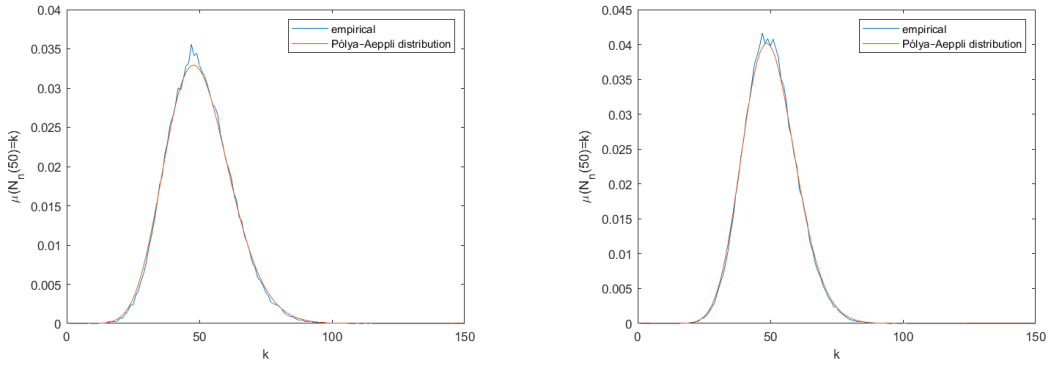
6.1 Statistics of returns for some non-deterministic stationary systems

We stressed in section (3.3) that for chaotic deterministic systems, the statistics of the number of visits into a ball of small radius around a target point z is given by a pure Poisson distribution for non-periodic z or by a Polyà-Aeppli distribution, depending on the extremal index associated to the standard observable for periodic z. It was already proven (see for example [85]) that when the dynamics is perturbed by a uniform noise, the extremal index associated to the standard observable immediately turns to 1, as periodic orbits are 'destroyed' by the perturbation, but there are only few results on the statistics of the number of visits in the case of random systems affected by other kinds of noise. In [70], the authors presented an example of partly random system for which the extremal index is 1 almost everywhere, and where a Poisson distribution for $N_n(t)$ holds for almost all target point z. In this example, a Polyà-Aeppli distribution holds for a set of particular target points of null measure. However, there are no general results concerning the statistics of visits inside these balls for stationary processes. We will here compute the statistics of the number of visits for two particular examples that are somehow interesting and for which we can compute explicitly the extremal index.

The first example is the composition of the two maps T_0 and T_1 presented in section (4.4.1) with $b = 1/2$, for which there are infinitely many non zero p_k terms in the spectral formula (20). The second example is the system with moving target presented in section (4.4.2). We compare in figure (21) the

empirical distribution of the number of visits $N_n(t)$ in the target set with the Polyà-Aeppli distribution given by the parameters t and the extremal index that we computed explicitly earlier for these processes. We recall that for the first example, the extremal index for the target point 0 is $2/3$ and in the moving target example, the extremal index is $1 - \frac{3}{2}(\frac{1}{4})^2 - 6(\frac{1}{4})^3(\frac{1}{2})^2 \approx 0.8828$. The results are depicted in figure (21) for $t = 50$ and compare very favorably with the Polyà-Aeppli distributions. This perfect agreement is stable against different values of t. We refer to the caption of the figure for the details of the numerical computation.

Although rigorous theoretical results are lacking, these numerical evidence seem to indicate that a geometric Poissonian distribution given by the extremal index seems to be a universal feature for this kind of observable. To corroborate this claim, we investigate in the next section whether a similar behavior occurs for natural, possibly noisy systems.



(a) (b)

Figure 21: For the composition of two maps with $b = 1/2$ (left) and the moving target example (right), comparison between the empirical distribution $\mu(N_n(50)) = k$ and the Polyà-Aeppli distribution of parameters $t = 50$ and respectively $\theta = 2/3$ and $\theta = 0.8828$. The empirical distributions are obtained by considering 10^5 runs for which is recorded the number of entrances in the target balls of radius $r = \exp(-\tau)$, where τ is the 0.99-quantile of the empirical distribution of the observable computed along a pre-runned trajectory of size 10^4, after $t/\mu(B(z,r))$ iterations of the map.

6.2 Statistics of returns for climate data

We study now the statistics of the number of visit for the atmospheric data presented in section (4.2). The target set considered in this analysis is a ball centered at a point z that corresponds to the mean sea-level pressure recorded on

January 5th, 1948. The radius $r = e^{-\tau}$, is fixed using as τ the $0.95-$quantile of the observable distribution. We consider the number of visits in an interval of length 365 days, that corresponds to the annual cycle. This time scale is relevant for the underlying dynamics of the atmospheric circulation. The parameter t in equation (29) is thus taken equal to $t = \lfloor 365\mu(B(z,r))\rfloor$. For the target set considered, we get $t = 18$, that is close to the mean value that we find for the number of visits that is approximately equal to $18,13$. In figure (22), we compare the obtained empirical distribution with the Polyà-Aeppli distributions of parameters t and the EI obtained using different estimates: Süveges, $\hat{\theta}_5$ and $\hat{\theta}_{10}$. The distribution using the Süveges estimate seems to fit the empirical data better than the other two. In fact, if we increase the order of the estimate $\hat{\theta}_m$, the value it attributes to the EI decreases and the associated Polyà-Aeppli distribution flattens. This is probably due to finite effects; indeed we observe that if the fixed threshold u in formula (37) is too small, the numerator approaches 0 as m increases, yielding bad estimates for the EI. Qualitatively similar results are found for different atmospheric states and different quantiles considered (although taking higher quantiles gives too few data for the statistics), suggesting that a Polyà-Aeppli distribution for the number of visits holds for this type of high dimensional observational data.

This analysis, applied systematically to a state of a complex system, can provide important information on the recurrence properties of the target state, namely how many visits to expect in a given time interval. The drawback is in the number of data required to provide a reliable estimates of the visits distribution (see figure (22)). Let us now move to a more exotic scenario and investigate what happens for non-stationary systems.

6.3 Statistics of returns for sequential systems

Results proving the existence of extreme value laws for non-stationary systems are lacking. Nevertheless, the methods of estimation of the extremal index described in the thesis can still be used to evaluate a quantity θ that would correspond to the extremal index in the eventuality that an extreme value law hold for this kind of systems. We can also alternatively define the extremal index by computing the statistics of the number of visit and evaluating the expectation in equation (32). We choose this second option and consider the motion given by the concatenation

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

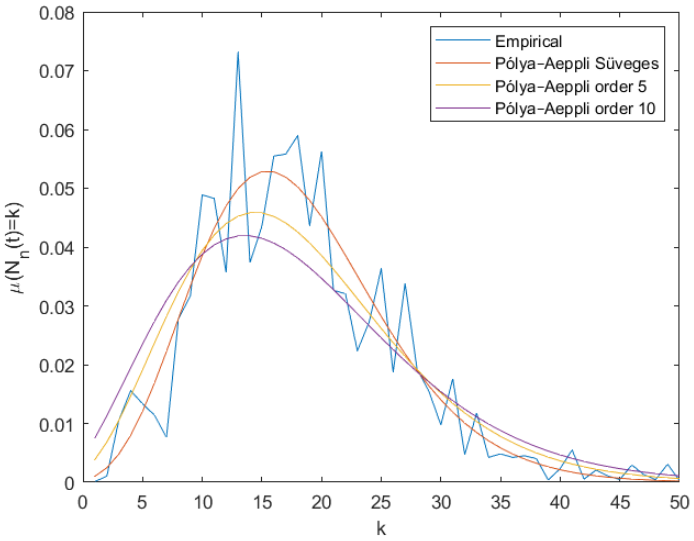


Figure 22: Comparison between the empirical distribution $\mu(N_n(t) = k)$ and the Polyà-Aeppli distributions of parameters $t = 18$ and the EI found with different estimates. The empirical distribution is obtained by considering 24396 intervals of length 365 and counting the number of visits inside the target ball described in the text.

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

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

where b_i is the i^{th} component of a vector $\bar{b}$ of size 10, with entries equally spaced between 0 and 1/2. We consider sequences of time intervals $[\tau k+1, (k+1)\tau]$, with $\tau = 10$, for $k = 0, 1, 2, \ldots$ in which the weights associated to ξ_i are equal to p_i^k. For every τ iterations, the weight associated to each ξ_i changes randomly, with the only constraint that they sum to 1. $\hat{\theta}_5$ is computed considering a trajectory of 5.10^7 points, and a threshold u corresponding to the $0.995-$quantile of the observable distribution. Using the same trajectory, we computed the empirical distribution of the number of visits in a ball centered at the origin and of radius $r = e^{-u}$ in intervals of time of length $\lfloor 2tr \rfloor$, with $t = 50$. This is indeed t times the Lebesgue measure (which is the invariant measure associated to all the maps f_i) of a ball of radius r centered in the origin. Choosing the same trajectory for the computation of θ and for the statistics of visits is of crucial importance here, because different probability laws imply large variations of θ depending on the trajectory considered. In figure (23), we observe again a perfect agreement between the empirical distribution of the number of visits and the Polyà-Aeppli distribution of parameters $t = 50$ and $\hat{\theta}_5$, which is equal to 0.92 for the trajectory presented in the figure (23). Although

different trajectories give different values for θ, this agreement is stable against different trajectories and different values of t and τ, suggesting that a geometric Poisson distribution for the number of visits given by θ is a universal feature, even for non-stationary scenario.

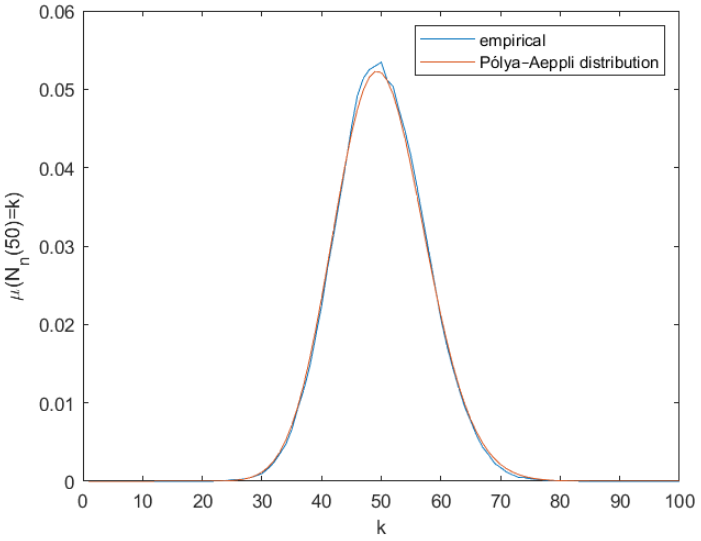


Figure 23: Comparison between the empirical distribution $\mu(N_n(50) = k)$ and the Polyà-Aeppli distribution of parameters $t = 50$ and $\hat{\theta}_5 \approx 0.92$, computed with a trajectory of length 5.10^7 of the sequential system described in the text. The procedure used to compute the empirical distribution is described in the text.

6.4 Statistics of returns in the neighborhood of the diagonal

This particular target set is of interest, since it allows to compute the statistics of close encounters of several independent trajectories of the system. We stated in section (3.3) and it was proven in [68] that the statistics of return in the set Δ_n^2 is given asymptotically by a compound Poisson distribution, different from Polyà-Aeppli. The example of the Markovian map presented earlier in equation (93), for which analytical computations are possible, allows us to check numerically this result. Consider the extremal index θ_2 associated to the observable ϕ_2. θ_2 is given by formula (95) and gives $\theta \approx 0.5926$.

Here, the block-size distribution π_k is not geometric and we get for the statistics of the number of visits a compound Poisson distribution which is not Polyà-Aeppli. To illustrate this, we show in figure (24) the distributions $\mu(N_n(t)) = k$ for $t = 50$ and different values of k for the real process associated to the Markovian map in-

troduced in section (5), and the distribution obtained by supposing a Polyà-Aeppli with parameter equal to 0.5926 which correspond to the dynamical extremal index of order 2 for this map (see formula (95)). The set Ω_n is in this context the tubular neighborhood of the diagonal Δ_n^2 defined in equation (69). For our numerical simulation we fix a small set Δ_n^2 by taking u_n in equation (69) equal to the $0.99-$quantile of the distribution of the observable computed along a pre-runned trajectory of 10^4 points. We easily compute that for a small enough neighborhood of the diagonal,

$$\mu_2(\Delta_n^2) \approx (h_1^2 + h_2^2 + h_3^2)\sqrt{2}\exp(-u_n). \tag{103}$$

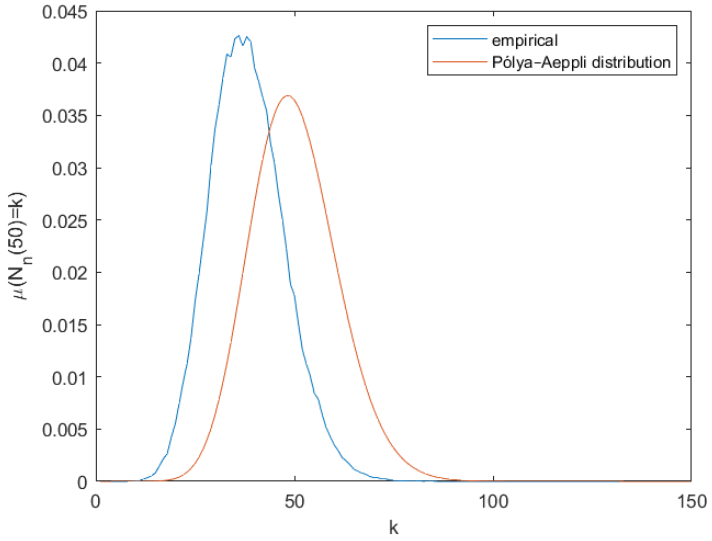


Figure 24: Comparison between the empirical distribution $\mu(N_n(50)) = k$ and the Polyà-Aeppli distribution of parameters $t = 50$ and $\theta = 0.5926$. The empirical distribution is obtained by considering 10^5 runs for which is recorded the number of entrances in the set Δ_n^2 , as defined in the text, after $t/\mu_2(\Delta_n^2)$ iterations of the map.

Figure (24) shows indeed an important difference between the obtained empirical distribution and the Polyà-Aeppli distribution of parameters $t = 50$ and $\theta_2 = 0.5926$, indicating that the distribution of the block size distribution π_l is not geometric in θ_2.

7 Generalized dimensions and large deviations of local dynamical quantities

In section (4.1), we described the set of tools used in the multifractal formalism. These tools, such as Legendre transforms and weighted integrals, are to be found in the much more general framework of large deviation theory. In fact, the mechanisms articulating this theory, of which we will give a brief review are also to be found in many physical situations, such as statistical physics, thermodynamics, and as we will see now, in the study of dynamical systems.

7.1 Elements of large deviations theory

The theory of large deviations is a very general framework which aims at studying the probability of occurrence of rare events of small probability, which are abnormal values taken by a random variable, that deviate from a typical value. This is somehow also the interest of extreme value theory, but the results are stated in a different formalism. In the study of chaotic dynamical systems, these rare events are atypical values taken by some local dynamical indicators such as finite resolution local dimensions, return times and hitting times.

Originally introduced in insurance mathematics by Crámer [35], the theory gained solid and general mathematical formulation thanks to the works of Varadhan [36], Gärtner and Ellis [108] between others. It is a particularly suited formalism to work with thermodynamical quantities and it is used in a wide variety of fields such as statistical physics, finance and engineering. We will now give a concise introduction to this matter (a more detailed presentation of the theory can be found in [88] and a basic introduction in [52]).

Let X_n be a random variable whose realizations depend on the parameter n and such that $X_n \to s_{typ}$ as $n \to \infty$. The fluctuations of X_n around the typical value s_{typ} are generally modulated by a function of the parameter n and vanish as this parameter is pushed to the limit. The study of these fluctuations is the core of large deviation theory. This motivates the following definition.

Definition 15 *We say that the random variable X_n, with associated probability density function p_n satisfies a large deviations principle if there exists a non-degenerate function Q such that following limit exists:*

$$\lim_{n\to\infty} -\frac{1}{n}\log p_n(s) = Q(s). \tag{104}$$

In this case, Q is called the rate function and it measures the intensity of the deviations from the typical value.

The last definition implies the scaling

$$p_n(s) \underset{n\to\infty}{\sim} e^{-nQ(s)}. \tag{105}$$

Q is generally a convex function defined in an interval (s_{min}, s_{max}) and has a minimum of 0 that is reached at the typical value s_{typ}, so that at the limit of large n, p_n converges to the Dirac distribution $\delta_{s_{typ}}$: when $n \to \infty$, the probability that X_n is equal to s_{typ} tends to 1, and to 0 for all the other values. We see from the scaling (105) that when the convergence is not achieved, the probability of deviating with an intensity ε from the typical value decreases exponentially with ε, at a rate that is given by the function Q.

A typical example of random variable satisfying a large deviation principle is the empirical mean S_n of a cluster of n observations of an i.i.d. Gaussian random variable X_i. As $n \to \infty$, it is well known from the law of large numbers that S_n tends in probability to the expected value of X_0, and the asymptotic deviations from this typical value are given by the central limit theorem. In this simple case, it is immediate to see that S_n follows a large deviation principle with a rate function equal to $\frac{(s-\mu)^2}{2\sigma^2}$, where μ is the expected value and σ the standard deviation of the X_i.

The rate function is the central object in the theory, since it gives an approximation of the distribution of X_n for large n. Depending on the studied quantities, the rate function can take various forms that can sometimes be computed explicitly. To do so, one can use the central result of the theory, the so called Gärtner-Ellis theorem [108] that establishes a profound connection between the rate function and an other object that we define now.

Definition 16 *We call the scaled cumulant generating function γ_q associated with X_n the following function (if it exists):*

$$\gamma(q) := \lim_{n\to\infty} \frac{1}{n}\log \mathbb{E}[e^{nqX_n}]. \tag{106}$$

In the context of dynamical systems, $\gamma(q)$ is often called the free energy *associated*

to the process X_n, in analogy with thermodynamics.

This function is in many applications computable numerically, if not analytically (as we will see later). Its importance lies in its relation with the rate function Q of the process , enlightened by the following theorem:

Theorem 6 *(Gärtner and Ellis [108]) If the scaled cumulant generating function $\gamma(q)$ associated to the process X_n is differentiable, then:*

1- X_n satisfies a large deviations principle.

2-The associated rate function Q is the Legendre transform of γ:

$$Q(s) = \sup_{q \in \mathbb{R}} \{qs - \gamma(q)\}. \tag{107}$$

This important result will be the principal tool used in the following sections, where we derive some large deviations principles for various local quantities associated with the dynamics, in connection with the generalized dimensions of the invariant measure.

7.2 Generalized dimensions and large deviations of the local dimension

The similarity between the objects used in large deviations theory and in the multifractal formalism in striking. For example, the spectrum of generalized dimensions and the singularity spectrum are also related via a Legendre transform. In fact we will see now that the spectrum of generalized dimensions is deeply connected to the free energy associated to the finite resolution local dimension. Indeed, let us consider a chaotic dynamical system (M, T, μ). The value taken by the local dimension at resolution $r = e^{-n}$, $d_n = d(., r)$, can be seen as a random variable depending on the parameter n. If we assume that the limit exists, we have that the scaled cumulant generating function writes:

$$\begin{aligned} \gamma(q) &= \lim_{n\to\infty} \frac{1}{n} \log \int_M e^{nqd_n} d\mu = \lim_{r\to 0} \frac{-\log \int_M r^{-qd(x,r)} d\mu(x)}{\log r} \\ &= \lim_{r\to 0} \frac{-\log \int_M \mu(B(x,r))^{-q} d\mu(x)}{\log r} = -\tau(-q+1). \end{aligned} \tag{108}$$

We recall that the function $\tau(q)$ is defined in the multifractal formalism as $(q-1)D_q$. If moreover $\tau(q)$ is differentiable, then by applying the Gärtner-Ellis theorem, we have that $d(.,r)$ satisfies the large deviation relation:

$$\mu(\{x \in M, d(x,r) \in [s, ds)\}) = r^{Q(s)}, \tag{109}$$

where

$$Q(s) = \sup_{q\in\mathbb{R}}\{-qs + \tau(-q+1)\} = \sup_{q\in\mathbb{R}}\{qs + \tau(q+1)\}. \tag{110}$$

The existence of the limit and the differentiability of τ was proven rigorously for a class of systems called conformal repellers in [80], but the universality of the large deviation principle along with numerical evidence that we will present now suggest that it could hold for a much larger class of dynamical systems. This relation throws a new light on the considerations of section (4.3): although the local dimension is constant $\mu - a.e.$, when considering the finite resolution local dimensions, fluctuations from the typical value D_1 are observed, that may be caused by the presence of points of different scaling hidden is the neighborhood of the target point z (the ball of center z and of radius r).
We now test this relation for systems that are not conformal repellers, by comparing the rate function associated to the motion on the Sierpisnki Gasket as defined in the section (5) obtained via two different methods. First, as the analytical value for D_q is known from formula (83), the rate function can be evaluated using equation (110). We then compare it to the empirical rate function obtained using the following procedure:

1. For a fixed radius r, compute for a large number l of points $z \in M$ the quantity $d(z,r)$ via the EVT method described in section (4), and discard all eventual values greater than the dimension of the ambient space (these values are generally the result of numerical effects, often associated to border points or pathological points of the attractor).

2. Record d_{min} and d_{max} the minimum and maximum of the obtained values and chop the interval $[d_{min}, d_{max}]$ into n_{bin} segments of the form $I_k = [s_k, s_{k+1}]$ with $s_1 = d_{min}$ and $s_{k+1} = s_k + \frac{d_{max}-d_{min}}{n_{bin}}$. Normalize the result in order to obtain an empirical histogram distribution $f(s_k)$ of $d(.,r)$, the fraction of

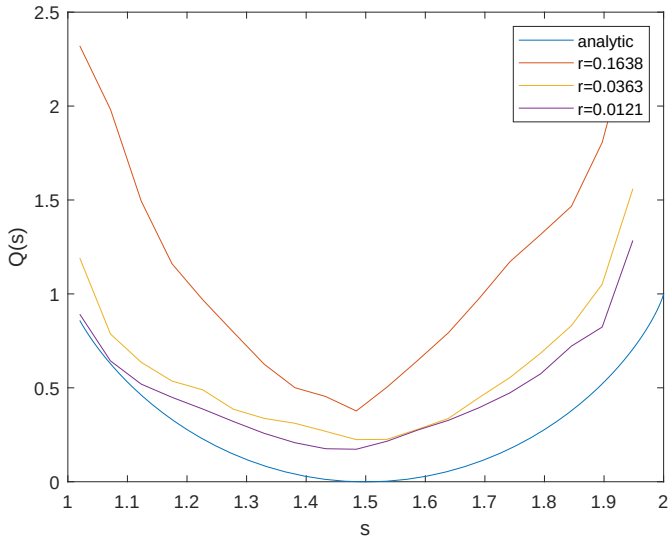


Figure 25: Rate functions for the motion on a Sierpinski gasket, computed analytically for the blue curve and applying the aforementioned procedure from $1,000$ sampling points, each of which required a trajectory consisting of 10^6 iterates.

points for which $d(.,r)$ falls in I_k.

3. denote $\bar{d}$ the mean of this distribution and for each bin $[s_k, s_{k+1}]$ compute the quantity $\hat{Q}$:

   If $s_{k+1} < \bar{d}$, $\hat{Q}(s_k) = \frac{\log f(s_{k+1})}{\log r}$.

   If $s_k \geq \bar{d}$, $\hat{Q}(s_k) = \frac{\log f(s_k)}{\log r}$.

4. Apply this procedure for smaller and smaller values of r.

By applying this procedure to the motion on the Sierpinski gasket, we observe in figure (25) a convergence as r tends to 0 of the empirical rate function $\hat{Q}$ to the predicted curve. We point out that for numerical constraints, we cannot push the convergence too far and that the smallest value of r that we took in this study is 0.0121.
The dynamics on the Sierpinski gasket defines a system that is not a conformal repeller. Nonetheless, we observe a convergence of the empirical rate functions to the theoretical curve, indicating that such a large deviation principle governed

by the generalized dimensions of the system seems to hold in more general contexts. The results of section (4.3) also seem to support this fact. Indeed, the link between the distribution of finite resolution local dimensions and the generalized dimensions was made for a larger class of dynamical systems. Let us now investigate the existence of large deviations principles associated to other local dynamical quantities.

7.3 Generalized dimensions and large deviations of first return/hitting times

We have seen in section (7.2) that the large deviations of the finite resolution local dimensions are somehow governed by the generalized dimensions of the system. We will see now that the multifractal nature of the measure also influences the behavior of dynamical quantities such as first return time and first hitting time of the dynamics into small target balls. We now define these quantities.

Definition 17 *We call the first Hitting time of the point $x \in M$ into the ball centered at $z \in M$ and of radius r the following quantity:*

$$\mathscr{H}_{B(z,r)}(x) = \min\{n > 0, T^n(x) \in B(z,r)\}. \tag{111}$$

Whenever the point x belongs to $B(z,r)$, this quantity is called the first return time of x into this ball.

The first hitting time is a quantity of many interest in physical applications, since its statistics allow to foresee the evolution of a system near a particular state of interest. We have already pointed out in section (3) the relation between the existence of an Extreme value law for a suitable observable and the existence of hitting time statistics, made by the Freitas Freitas Todd theorem [10].
A fundamental result concerning return times statistics is their relation, for ergodic dynamical systems, to the measure of the target set. This result is known as the Kac theorem.

Theorem 7 *(Kac [41]) If the invariant measure μ is ergodic, then:*

$$\int_{B(z,r)} \mathscr{H}_{B(z,r)}(x) d\mu|_{B(z,r)}(x) = \frac{1}{\mu(B(z,r))}. \tag{112}$$

Here, $\mu|_{B(z,r)}$ is the restriction of μ to $B(z,r)$, defined for all measurable set A by:

$$\mu|_{B(z,r)}(A) = \frac{\mu(A)}{\mu(B(z,r))}. \tag{113}$$

This theorem is central in ergodic theory and is used to prove plethora of results. For example, by combining it with the Ornstein Weiss Lemma [42], it can be proven in some quite general setting that for μ almost all $z \in M$,

$$\lim_{r\to 0} \frac{\log \mathscr{H}_{B(z,r)}(z)}{-\log r} = D_1. \tag{114}$$

We refer to [77, 78] for the detailed conditions of application and a proof of this result. Just like for the finite resolution local dimensions, it has been proven recently in [80] that for conformal repellers, the quantity

$$h_r(z) := \frac{\log \mathscr{H}_{B(z,r)}(z)}{-\log r} \tag{115}$$

follows a large deviation principle. The quantity deviates from its typical value D_1 with a rate function Q^* that is a bit more complicated than Q but is again determined by the spectrum of generalized dimensions of the system [80]. We have that for any interval I:

$$\mu(\{z \in M, h_r(z) \in I\}) \sim r^{\inf_{s\in I} Q^*(s)}. \tag{116}$$

This result suggests that the multiplicity of scaling behaviors of the measure influences the statistics of return times. We investigate now whether an analogous result exists for the first hitting time and consider the quantity

$$h_r(z,x) := \frac{\mathscr{H}_{B(z,r)}(x)}{-\log r}. \tag{117}$$

It is not surprising that for systems with fast enough decay of correlations, the origin of the starting point x is not of fundamental importance, since the system loses memory over time. We therefore expect to find a result analogous to the one for the return time. It was indeed proven in [84] that for $\mu \times \mu-$ almost all points $(x,z) \in M^2$, we have:

$$\lim_{r\to 0} h_r(z,x) = D_1. \tag{118}$$

More generally, the relation between hitting time behavior and the local dimension has been explored in a series of papers [109, 110, 111]. It is now natural to ask whether the quantity $h_r(z,x)$ enjoys large deviations from its typical value D_1 and

what is its eventual associated rate function $\hat{Q}$. The usual procedure to answer this question is to start by computing the free energy $R(q)$ associated to this process. Again, taking $n = -\log r$ in definition (16), we have that:

$$\begin{aligned} R(q) &= \lim_{r\to 0} \frac{1}{-\log r} \log \int_M \int_M e^{-\log r q h_r(z,x)} d\mu(x) d\mu(z) \\ &= \lim_{r\to 0} \frac{1}{-\log r} \log \int_M \int_M e^{q \log \mathscr{H}_{B(z,r)}} d\mu(x) d\mu(z) \\ &= \lim_{r\to 0} \frac{1}{-\log r} \log \int_M \int_M \mathscr{H}^q_{B(z,r)} d\mu(x) d\mu(z). \end{aligned} \tag{119}$$

Now that we have found a general expression for the free energy R(q), an important step to go toward the desired large deviation result is to prove a technical result: lemma (1). This is done under the following set of assumptions:

- **A-1: Exponential distribution of hitting times with error**. There is a constant $C > 0$ such that for μ-a.e. $z \in M$ and $t > 0$ we have

$$\left| \mu\left(\left\{ x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x) \geq \frac{t}{\mu(B(z,r))} \right\} \right) - e^{-t} \right| \leq C\delta_r \max(t,1)\ e^{-t} \tag{120}$$

  where

$$\delta_r = O\left(\mu(B(z,r))|\log \mu(B(z,r))|\right). \tag{121}$$

  In particular, for $t > \mu(B(z,r))|\log \mu(B(z,r))^C|$ we have:

$$\mu\left(\left\{ x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x) \geq \frac{t}{\mu(B(z,r))} \right\} \right) = \exp[-t(1+O(\delta_r)](1+O(\eta_r)), \tag{122}$$

  with $\eta_r = O(\mu(B(z,r)))$, while for $t \leq \mu(B(z,r))|\log \mu(B(z,r))^C|$ we have[1]

$$\mu\left(\left\{ x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x) \geq \frac{t}{\mu(B(z,r))} \right\} \right) \geq 1 - \frac{t}{C}. \tag{123}$$

  Notice that the above implies that both η_r and δ_r depend on the point $z \in M$.

- **A-2: Exact dimensionality.** The measure μ is exact dimensional.

- **A-3: Uniform bound for the local measure.** There exists $d^* > 0$ such that for all $z \in M$ we have

$$\mu(B(z,r)) \leq r^{d^*}. \tag{124}$$

[1]In the proof of lemma 1 we set the constant exponent C to unity because its value is irrelevant for the proof.

- **A-4: Existence and regularity of the correlation integrals.** For all $q \in \mathbb{R}$ the limit defining $\tau(q)$, equation (12), exists. Moreover the function $\tau(q)$ is real analytic for all $q \in \mathbb{R}$, $\tau(0) = -D_0$, $\tau(1) = 0$, $\tau'(q) \geq 0$ and $\tau''(q) \leq 0$. In particular $\tau''(q) < 0$ if and only if μ is not a measure of maximal entropy.

This last assumption is a strong one and it has been proven to hold for conformal mixing repellers endowed with Gibbs measures in [24]. This condition has also been assumed in [80], to prove the large deviation result for the return times (116). For the same class of conformal repellers Assumption A-3 holds too (see Lemma 3.15 in [80]). Assumption A-1 was proven for a large class of mixing systems with exponential decay of correlations in [37]. However, we consider these *ideal* systems interesting models to establish rigorous results that might also hold in more general settings.

Lemma 1 *If the system verifies the assumptions A-1 to A-4, then we have that the free energy $R(q)$ associated to $h_r(.,.)$ is equal to:*

- $$R(q) = -\tau(1-q) = qD_{1-q} \text{ if } q > -1. \tag{125}$$
- $$R(q) = D_2 \text{ if } q \leq -1. \tag{126}$$

In order to prove with result, we will need a further lemma, whose demonstration is a standard exercise.

Lemma 2 *Consider a function f from M to the integer numbers larger than, or equal to one: $f : M \to \mathbb{N}_+$. Let $0 < A \leq 1$ and define*

$$I(q) = \int_M f^q(x) d\mu(x). \tag{127}$$

Then, when $q > 0$,

$$I(q) = 1 - \lim_{t\to\infty} t^q \mu\left(\{x \in M \text{ s.t. } f(x) > t\}\right) + \frac{q}{A^q} \int_A^\infty t^{q-1} \mu\left(\left\{x \in M \text{ s.t. } f(x) > \frac{t}{A}\right\}\right) dt. \tag{128}$$

On the other hand, when $q < 0$,

$$I(q) = -\frac{q}{A^q}\int_A^\infty t^{q-1}\mu\left(\left\{x \in M \ s.t. \ f(x) < \frac{t}{A}\right\}\right)dt. \tag{129}$$

Proof of Lemma 1. We follow the scheme of the proof of Theorem 3.1 in [47], by translating its argument from cylinders to balls. To achieve so, we apply lemma 2 to $f(x) = \mathscr{H}_{B(z,r)}(x)$ and consider the integral

$$I(q-1,z,r) = \int_M \mathscr{H}_{B(z,r)}(x)^{q-1}d\mu(x), \tag{130}$$

which is of the form studied in lemma 2. There are three cases to consider.

1. Case 1: $q > 1$, that is, $q - 1 > 0$ that allows to apply formula (128). Because of assumption **A-1**, equation (120), the limit in equation (128) is null. Moreover, using again equation (120) for $t > 1$,

$$e^{-t}(1-C\delta_r t) \leq \mu\left(\left\{x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x)) \geq \frac{t}{\mu(B(z,r))}\right\}\right) \leq e^{-t}(1+C\delta_r t), \tag{131}$$

   we can bound the integrand at right hand side in equation (128):

$$I(q-1,z,r) \geq \frac{q-1}{\mu(B(z,r))^{q-1}}\left[\int_1^\infty t^{q-2}e^{-t}dt - C\delta_r\int_0^\infty t^{q-1}e^{-t}dt\right].$$

   The last two integrals are convergent. Moreover, observe that equation (121) and **A-3** imply that δ_r is uniformly bounded from above. Taking r small enough the term into brackets becomes positive and larger than a quantity independent of r. A similar reasoning yields an upper bound for $I(q-1,z,r)$ of the same form. Since the double integral $\int_M\int_M \mathscr{H}_{B(z,r)}(x)^{q-1}d\mu(x)d\mu(z)$ is the single integral of $I(q-1,z,r)$ with respect to $d\mu(z)$, the equality (125) follows.

2. Case 2: $0 < q < 1$. Since $q - 1 < 0$ we employ equation (129):

$$I(q-1,z,r) = \frac{1-q}{\mu(B(z,r))^{q-1}}\int_{\mu(B(z,r))}^\infty t^{q-2}\mu\left(\left\{x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x) < \frac{t}{\mu(B(z,r))}\right\}\right)dt. \tag{132}$$

   We use again equation (131) to get

$$I(q-1,z,r) \geq \frac{1-q}{\mu(B(z,r))^{q-1}}\int_1^\infty t^{q-2}(1-e^{-t}(1+C\delta_r t))dt = \frac{1-q}{\mu(B(z,r))^{q-1}}S(z,r).$$

In the above we have put

$$S(z,r) = \int_1^\infty t^{q-2}(1-e^{-t})dt + C\delta_r \int_1^\infty t^{q-1}e^{-t}dt,$$

The term $S(z,r)$ is again composed of a constant (the first convergent integral) and of a vanishing quantity (when $r \to 0$), which puts us in the position of using the previous technique to obtain a first inequality between the two terms of equation (125).

To prove the reverse inequality we begin by observing that for $t < 1$ condition **A-1** simply becomes $|\mu\left(\left\{x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x)) \geq \frac{t}{\mu(B(z,r))}\right\}\right) - e^{-t}| \leq O(\delta_r)$; then we start from equation (132) and we part the integral in two, the first from $\mu(B(z,r))$ to one and the second from one to infinity:

$$I(q-1,z,r) = \frac{1-q}{\mu(B(z,r))^{q-1}}(J_1 + J_2). \tag{133}$$

We use again equation (131) to get:

$$J_1 = \int_{\mu(B(z,r))}^1 t^{q-2}\mu\left(\left\{x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x) < \frac{t}{\mu(B(z,r))}\right\}\right)dt \leq \tag{134}$$

$$\int_{\mu(B(z,r))}^1 t^{q-2}(1-e^{-t}+O(\delta_r))dt \leq \int_0^1 t^{q-2}(1-e^{-t})dt + O(\delta_r)\int_{\mu(B(z,r))}^1 t^{q-2}dt. \tag{135}$$

The first integral in the above is a positive constant. Let us consider the second term:

$$O(\delta_r)\int_{\mu(B(z,r))}^1 t^{q-2}dt \leq \frac{O(\delta_r)}{1-q}\mu(B(z,r))^{q-1} = O\left(\mu(B(z,r))^q|\log\mu(B(z,r)))\right|,$$

where the last equality follows from $\delta_r = O\left(\mu(B(z,r))|\log\mu(B(z,r)))\right|$, equation (121). Therefore, when r tends to zero, this term vanishes.

The case of J_2 is easier:

$$J_2 = \int_1^\infty t^{q-2}\mu\left(\left\{x \in M \text{ s.t. } \mathscr{H}_{B(z,r)}(x) < \frac{t}{\mu(B(z,r))}\right\}\right)dt \leq \int_1^\infty t^{q-2}dt = C_2, \tag{136}$$

which is again bounded by a constant.

3. Case 3: $q < 0$. We use again equation (133), with J_1 and J_2 defined in equations (134) and (136), respectively. For the latter integral the inequality (136)

still holds, with a different constant C_2. To deal with the integral J_1 we further split its domain into the intervals $[\mu(B(z,r)), \mu(B(z,r))|\log\mu(B(z,r))|]$ and $[\mu(B(z,r))|\log\mu(B(z,r))|, 1]$, thereby defining the integrals $J_{1,1}$ and $J_{1,2}$, respectively.

At this point we use equation (123) to estimate from above the integrand of $J_{1,1}$ and equation (122) to do the same for $J_{1,2}$. Putting the two estimates together, we obtain

$$J_1 \leq C_3 \int_{\mu(B(z,r))}^{1} t^{q-1} dt \leq \frac{C_3}{|q|} \left[\mu(B(z,r))^q - 1\right] \leq \frac{C_3}{|q|} \mu(B(z,r))^q, \qquad (137)$$

where C_3 is another constant independent of r and z.

To get a lower bound we write

$$J_1 = \int_{\mu(B(z,r))}^{1} t^{q-2} \mu\left(\mathscr{H}_{B(z,r)} < \frac{t}{\mu(B(z,r))}\right) dt \geq \int_{\mu(B(z,r))}^{1} t^{q-2} \mu(\mathscr{H}_{B(z,r)} \leq 1)\, dt$$

$$= \mu(B(z,r)) \int_{\mu(B(z,r))}^{1} t^{q-2} dt = |q-1|^{-1} \mu(B(z,r))^q \left[1 - \mu(B(z,r))^{|q-1|}\right]$$

$$\geq |q-1|^{-1} \mu(B(z,r))^q \left[1 - r^{d^*|q-1|}\right].$$

In the last step we have used equation (124), so that the term in the square brackets is positive and uniformly bounded for r small enough.

By collecting all the preceding estimates, we get the desired result for all q. ■

One of the assumptions required to prove the previous lemma was the existence and differentiability of the function τ defined in section (4.1). Even with this condition, the free energy function R is not differentiable in -1 and so we cannot apply the usual procedure of the Gärtner-Ellis theorem to obtain a large deviation principle. However, as R is differentiable and convex on the interval $(-1, +\infty)$, a local version of the theorem can be applied (as reported in the lemma XIII.2 in [89]) and yields the following theorem:

Theorem 8 *If the system satisfies the assumptions A1-A4, then for all* $s \in [D_1, D_{-\infty})$, h_r *follows the large deviations principle:*

$$\lim_{r\to 0} \frac{1}{\log r} \log \mu \times \mu(\{h_r(z,x) > s\}) = \hat{Q}(s), \qquad (138)$$

where $\hat{Q}(s) = \sup_{q\in\mathbb{R}} \{qs + \tau(1-q)\}$.

We find, again a large deviation principle associated to the hitting time in which the rate function is entirely determined by the generalized dimensions of the system. Although the result has been proven under the quite restrictive conditions $A1-A4$, we will present in the next section numerical results suggesting that it could hold in more general situations. These large deviations results reveal once again that the multifractal nature of the invariant measure of the system determines some pre-asymptotic dynamical behaviors of the system and that generalized dimensions can be used as a bridge between what we called at the end of section (4.3) the penultimate attractor and the ultimate fine structure of the attractor.

8 Further methods to compute the generalized dimensions using scaling relations

We have seen throughout the thesis that generalized dimensions are involved in different scaling relations. One can take advantage of these relations to compute numerically the generalized dimensions of a system. We will present briefly some of these methods already found in the literature and also introduce a further method using the results of the previous section relating hitting times and generalized dimensions.

8.1 Using the Correlation Integral

Let us describe the so called *correlation integral* method, discussed in details in [26]. This standard technique takes advantage of the scaling implied by definition (12), that is:

$$\int_M \mu(B(z,r))^{q-1} d\mu(z) \sim_{r\to 0} r^{D_q(q-1)}. \tag{139}$$

The idea is to estimate the quantity in the left hand side of the above relation, called the correlation integral and denoted $\Gamma_\mu(q,r)$. Then, the generalized dimensions of order q can be extracted by finding the scaling of $\Gamma_\mu(q,r)$ when r tends to 0. We will discuss in section (8.4) how to do so. To evaluate $\Gamma_\mu(q,r)$, we start by estimating, for a target point z_i chosen at random with respect to the invariant measure, the quantity $\mu(B(z_i,r))$. This is achieved by choosing a generic starting point x_0, running a long trajectory of the system $(x_0, Tx_0, ..., T^{n-1}x_0)$ and by computing the quantity:

$$\hat{B}_n(z_i,r) = \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{B(z_i,r)}(T^j x_0), \tag{140}$$

that is an estimate of the measure of a ball of radius r centered at z_i. Once these measures are evaluated for a sufficient number m of target points z_i, $\Gamma_\mu(q,r)$ is estimated via a weighted Birkhoff sum:

$$\Gamma_\mu(q,r) \approx 1/m \sum_{i=1}^{m} \hat{B}(z_i,r)^{q-1}. \tag{141}$$

This method is particularly suited to compute the positive q part of the spectrum. Indeed, for negative q, some balls of exceedingly small estimated measure can

make the weighted Birkhoff sum in relation (141) explode. Boxes of unnaturally small measure are often obtained numerically when considering points z_i located at border regions of the attractor, where much larger statistics should be used to obtain reliable results. To overcome this issue, several variant methods, still based on this general idea were proposed and gave satisfactory results [29],[39]. A further method based on the computation of return times was proposed. We will present it now.

8.2 Using return times

The Kac theorem (7) makes a link between measures of balls and return time statistics inside these balls. This profound relation led to the development of a method to compute generalized dimensions, particularly suited for negative q. The method, presented in [45] (see also [46] for a detailed review) takes advantage of the relation

$$\Psi_\mu(q,r) := \int_M \mathscr{H}^{1-q}_{B(x,r)}(x)\, d\mu(x) \sim_{r\to 0} r^{D_q(q-1)}. \tag{142}$$

This relation is obtained using Kac theorem and replacing the measure of balls in equation (12) by the return time of the center into its ball. In fact, a rigorous result would be given in terms of the expected value of the return time to the ball $B(x,r)$. As it is more difficult to implement, the method computes only the return time of the center of the ball and rely on an agreement between this quantity and the expectation of return times as the radius of the ball goes to zero. Although there is no general proof of this scaling, numerical results suggest the validity of relation (142). A discussion of this matter and a rigorous proof in a specific example is available in [46].

The *return time integral* $\Psi_\mu(q,r)$ can be evaluated via a weighted Birkhoff sum over the return times of the ball centers. It remains now, as for the correlation integral, to evaluate the scaling as r goes to 0 of $\Psi_\mu(q,r)$ to extract the generalized dimension of order q (it works only for negative q). We will discuss how to do so in part (8.4). It is interesting to remark that this method, as the one presented earlier using EVT in section (5) is also based on close recurrence properties of the system. The interesting difference is that they allow to evaluate complementary parts of the spectrum of generalized dimensions. We now present an analogous method, based on relation (125), that has the advantage of using hitting times, for which data, obtained using a single trajectory, are much more numerous than for return times.

8.3 Using hitting times

We start from the expression of the free energy in the first equation of lemma (1). Combining it with equations (12) and (119), the simple substitution $q \to 1-q$ yields the following scaling for $q < 2$:

$$\Upsilon_\mu(q,r) = \int_M \int_M \mathscr{H}_{B(z,r)}(x)^{1-q} d\mu(x) d\mu(y) \sim_{r\to 0} r^{D_q(q-1)}. \tag{143}$$

By estimating the *Hitting time double Integral* defined by the left hand side of the above relation, one can extract estimates of D_q for $q < 2$. To do so, we proceed in three steps. We start evaluating the integral over x for a fixed target point z_i by the Birkhoff sum:

$$\hat{H}_{n,q,r}(z_i) = 1/n \sum_{j=1}^{n} \mathscr{H}_{B(z_i,r)}(x_j)^{1-q},$$

where the x_j's are n points chosen at random with respect to the invariant measure. In order to optimize the computations, we select the points x_j to be the j^{th} iterates of a generic point x_0. By doing so for a long enough trajectory, we are assured that these points are sampled with respect to the physical invariant measure of the system. To determine the number of points n considered in the Birkhoff sum, we choose to stop the computation as soon as the trajectory of x_0 has entered $B(z_0, r)$ l times, where l is sufficiently large. In so doing, enough hitting time statistics should be considered in the sum. l is now a function of n.
Repeating this computation for m generic target points z_i on the attractor (that can also be taken along a trajectory of the system), a second step requires to perform another Birkhoff sum of the previous quantities, now over the target points to estimate the double hitting time integral:

$$\Upsilon_\mu(q,r) \approx 1/m \sum_{i=1}^{m} \hat{H}_{n,q,r}(z_i).$$

Since the scaling of $\Upsilon_\mu(q,r)$ is given by the generalized dimension of order q of the system, it remains (as for the correlation integral and the return time integral) to estimate this scaling for small values of r. We present in the following section a technique to do so.

8.4 Numerical extraction of the generalized dimensions

All three methods presented before are characterized by a scaling of the type

$$I_\mu(q,r) \sim_{r\to 0} r^{D_q(q-1)}, \tag{144}$$

where $I_\mu(q,r)$ can denote any of the three integrals presented in equations (139), (142), (143). A naive method to extract D_q would be to take r as small as possible and hope that the ratio between the logarithm of the integral estimate and the logarithm of r be not far from its limit value. In practice, the convergence of this ratio can be quite slow and this approach can give very bad estimates. A more reliable method consists in computing the integral for various values of r equally spaced in logarithmic scale, and finding a least square fit of the logarithm of the integral with respect to $\log r$.
In the numerical simulations that are presented below, the same trajectories are considered for the evaluation of $\Psi_\mu(q,r)$, $\Gamma_\mu(q,r)$ and $\Upsilon_\mu(q,r)$, so that the results can be compared fairly.

We first test this method of extraction on the Arnol'd cat map, defined by equation (74), that constitutes an archetypal chaotic dynamical system for which it is easily seen that assumptions $A1-A4$ are satisfied. We recall that this map has an invariant measure which is Lebesgue on the torus, and has therefore a constant spectrum of generalized dimensions equal to 2. The results are displayed in the left window of figure (26), where the scaling of the logarithms of the hitting time double integral and the correlation integral are shown, for various values of q. A visual investigation confirms the linear dependence of these quantities on $\log r$. To analyze the scaling more in detail, we define, for a parameter $\rho<1$ that gives the spacing of the values of r in logarithmic scale, the quantity:

$$\sigma_q(r) = \frac{1}{q-1}\frac{\log I_\mu(q,\rho r) - \log I_\mu(q,r)}{\log(\rho r) - \log r}. \tag{145}$$

This quantity is constant in the case of a linear dependence, as it is just the slope between successive values of r. On the other hand, non constant values of $\sigma_q(r)$ indicate that the asymptotic linear scaling is not yet achieved. The values of $\sigma_q(r)$ against r, for values of q between -1 and 2 are presented in figure (26). Unlike the case of the correlation integral (green curves), where $\sigma_q(r)$ are almost constant and equal to the predicted value of 2 (note that an upward shift of 0.05 of the curves is operated for clarity), the curves for the hitting time double integral for q

close to 2 are far from constant and the value obtained for $q = 2$ for the smallest radius $r = 1.77.10^{-3}$ is still far from 2 (around 1.83). For negative q however, the convergence for the hitting time double integral is obtained much more rapidly and match perfectly the predicted result. To deal with the problem for values of q close to 2, a further extrapolation can be considered. The convergence being quite slow, we expect a behavior that is typically of the type:

$$\sigma_q(r) = D_q + B/\log(r). \tag{146}$$

A fit that involves now two parameters can be performed and a better estimate of D_q can be found. The blue line in the right panel of figure (26) is the best fit obtained for $q = 2$, which gives a value of $D_2 = 2.006$ that matches the theoretical value up to three digits. For smaller values of r, the results get less precise. This effect is shown more in details in figure (27), in which the results for the return time integral method are also reported. This last integral is estimated with the same data as for the hitting time integral, but only first return statistics are considered. We see from this plot that all three methods provide reliable estimates correct up to three digits, at least for a suited range of r.

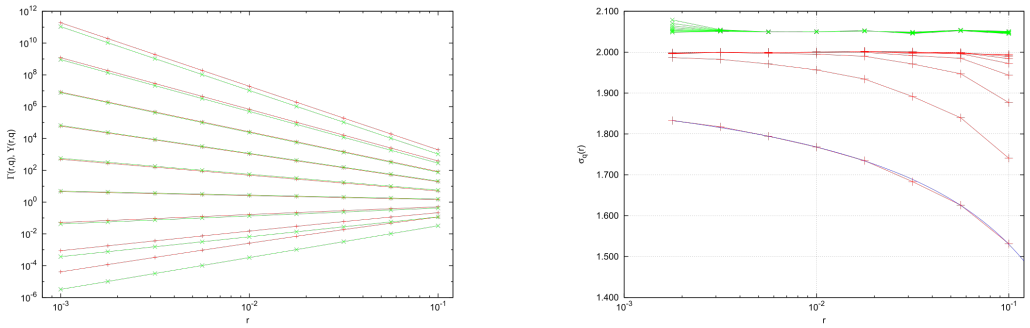


Figure 26: Left panel: correlation integral $\Gamma_\mu(r, q)$ (green lines) and hitting time integral $\Upsilon_\mu(q, r)$ (red lines) evaluated numerically by the procedure described in the text, with $l = 32$, $m = 256,000$. Lines join values with the same q, ranging from $q = -1$ (highest curve) to $q = 2$ (lowest). Right panel: slopes $\sigma_q(r)$ extracted from $\Upsilon_\mu(q, r)$ in the left panel, following equation (145) (red). Values of q range from $q = -1$ (highest curve) to $q = 2$ (lowest). The blue curve is the fit given by equation (146) with $D_2 = 2.006$ and $B = 1.095$. Finally, values of $\sigma_q(r)$ extracted from $\Gamma_\mu(q, r)$ are plotted in green, shifted upwards by .05.

The scaling (143) on which the hitting time double integral method is based has been proven to hold for systems satisfying assumptions $A1 - A4$, which are quite restrictive. We can ask now whether the relation still holds for a larger class of

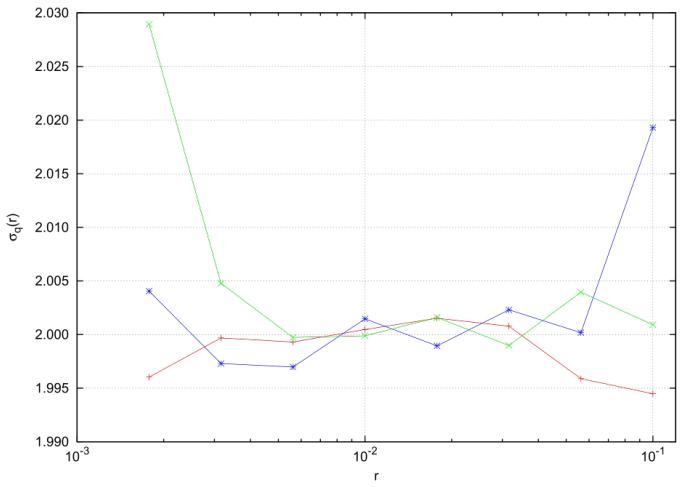


Figure 27: As in figure (26): data for $q = -1$. Also plotted (blue) are the data obtained by the first return time integral $\Phi_\mu(q, r)$, equation (142).

dynamical systems. Lacking a proof of this fact, we investigate numerically the validity of this relation, and so the domain of applicability of the technique for the case of the Hénon system, presented in section (5.1). This system does not satisfy assumption $A4$ in particular, as its spectrum of generalized dimension contains a point $q_T \approx 2.3$ where it is not differentiable, due to the non-uniform hyperbolicity of the system [6].

In figure (28), the variation of $\sigma_q(r)$ with respect to both q and r is plotted. For the Hénon physical measure, fits for the correlation integral are known to be extremely difficult as they depend strongly on the range of r considered [46]. This issue is apparent in figure (28), right panel, where the values of $\sigma_q(r)$ change significantly in a non predictable way for a fixed q as r changes. Despite this fact, a striking concordance is found between the values of $\sigma_q(r)$ for the hitting time double integral and for the correlation integral. This suggests that the method can be applied also in this context, so that equation (125) may hold in more general context than the set up under which it was proven rigorously.
In figure (29), we plot the best fits performed over the whole range of r displayed in the left panel of figure (26). The agreement between the two method is striking for values of q ranging from -4 to 1. The curve $D_2/(q-1)$, which follows from equation (126) and describes the scaling of the hitting time integral for q larger than or equal to two is also plotted (in blue). As in the case of Arnol'd cat, estimates for values

of q close to 2 need to be accelerated by the fitting of the kind of equation (146). For larger values of q, the scaling follows the predicted curve remarkably well.

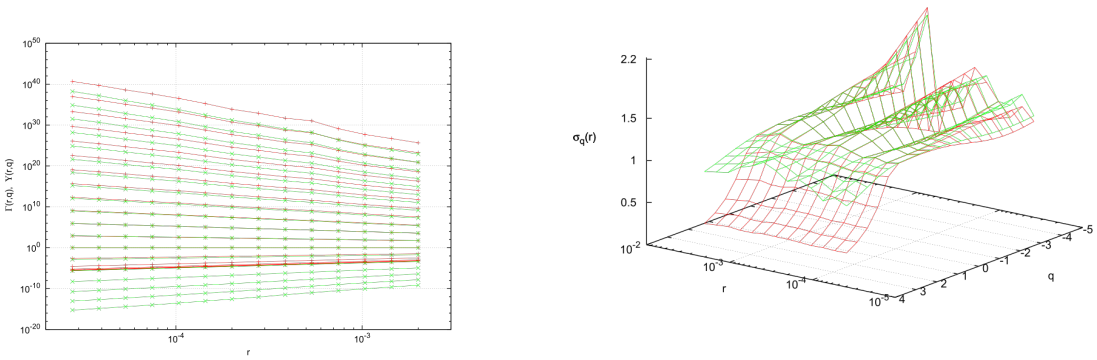


Figure 28: Left panel: correlation integral $\Gamma_\mu(r,q)$ (green lines) and hitting integral $\Upsilon_\mu(q,r)$ (red lines) evaluated numerically by the procedure described in the text with $l = 64$, $m = 256,000$. Lines join values with the same q, ranging from $q = -5$ (highest curve) to $q = 4$ (lowest). Right panel: slopes $\sigma_q(r)$ versus r and q extracted from $\Gamma_\mu(q,r)$ (green) and $\Upsilon_\mu(q,r)$ (red) in the left panel, following equation (145).

These experimental data leads us to conclude that the theoretical method to determine generalized dimensions implied by equation (125) has a practical value, especially for negative q. For values of q close to 2, reliable estimates may still be obtained by accelerating the convergence by suitable techniques.

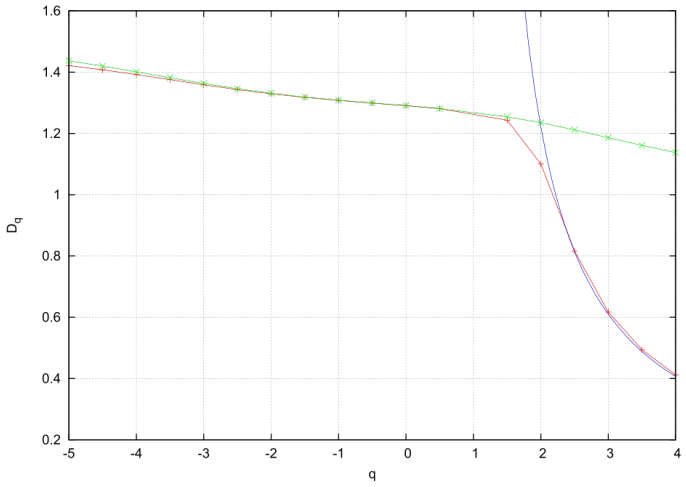


Figure 29: Generalized dimensions obtained by fitting the data in figure (28), left panel. Dimensions obtained from the correlation integral are plotted in green, from the hitting time integral in red. Plotted in blue is the curve $D_2/(q-1)$, implied by equation (126).

9 An application of EVT to the study of a real world multi-dimensional dynamical system; collective events in neural networks

This chapter is largely taken from the publication *T. Caby, G. Mantica, Extreme value theory of evolving phenomena in complex dynamical systems: firing cascades in a model of neural network [57].*

9.1 Motivations

Until here, we presented applications of Extreme Value Theory to the study of chaotic dynamical systems, making use of quite simple observables, in the sense that they can be computed at a given moment in time and they characterize regions with simple geometrical structures, such as balls or neighborhoods of the diagonal in the product space. This allowed us to obtain some meaningful information on the system. The classical theory of EVT however (the one originally developed in statistics) has been fruitfully applied to phenomena like earthquakes, floods, epidemics [114, 115], that all involve complex, many dimensional interactions that operate for a certain period and are otherwise silent. For such systems, the ex-

treme events can develop over long temporal scales and the regions of phase space characterized by the extreme events can have very complex, possibly fractal geometry. As a first step to account for such situations, [139] introduced observables (or, as we will call now *intensity functions*) that depend on the distance from an uncountable set of points populating a fractal set $K \subset M$. The authors revealed the role of fractal quantities like Minkowski dimension and Minkowski content of the fractal set K on the extreme value distribution of the observable. These results apply for general (*i.e.* not necessarily Lebesgue and possibly singular) invariant measures. Recently, this analysis gained full mathematical rigor in [132], where the authors proved rigorously that such extreme value laws hold and pointed out the role of the extremal index as an indicator of the compatibility of the dynamics with the geometry of the fractal set K.

Suppose that, unlike what we have studied so far, the intensity of a phenomenon can*not* be determined instantly at time n, when the dynamics visits the phase–space point z_n, because this phenomenon develops over a certain time span, so that its true intensity can only be assessed when it terminates. This is the case of neuronal avalanches, when the electrical activity of the brain is sustained for a few milliseconds. But it is perhaps easier, to focus ideas, to consider a geophysical analogy [123], to which the theory to be exposed herein can be equally applied.

An earthquake can be triggered by a single seismic event, a rupture, which propagates and multiplies in space and time. The total energy released by the quake is the intensity of the event, which can be assessed only after the *shock* terminates. Notice at this point that even if this event takes place continuously over a finite time–span, it is usually ascribed in geophysical catalogs to the moment of the first, triggering rupture. To put this observation into the theoretical framework of dynamical extreme values, the intensity function must take the form $\phi : M^m \to \mathbb{R}$, in the sense that the event at time n has an intensity $I(n, z_n)$ of the form $I(n, z_n) = \phi(z_n, z_{n+1}, \ldots, z_{n+m-1})$, where m is the duration of the phenomenon. We will now consider such a situation. A first observation must be made at this point. Since we are in the presence of a deterministic dynamics, $z_{n+j} = Ti^j(z_n)$, the intensity function takes on the special form

$$\phi(z_n, z_{n+1}, \ldots, z_{n+m}) = \phi(z_n, T^1(z_n), \ldots, T^{m-1}(z_n)) = \phi(z_n). \tag{147}$$

Intensity can therefore be thought of as depending only on the initial point z_n via a global function $\phi : M \to \mathbb{R}$. At first sight, one might therefore think that this approach is void of any generalization. As a matter of facts, because of this recourse

to future points of the evolution, the function ϕ is presumably very complex and its *phase–space portrait*, *i.e.* its level sets, might feature a fractal, hierarchical organization, thereby providing a concrete dynamical example of the theory of [139] mentioned before. On the other hand, when the dynamical system is chaotic, the function ϕ in formula (147) cannot be given a more efficient computational scheme other than what expressed by the central expression in equation (147): the best one can do is to follow the dynamics as it evolves and record its data. Think again of the earthquake example: while it develops, it is *practically impossible* to forecast its intensity. On a more theoretical vein, recall the aphorism by Joe Ford: *a chaotic system is its own fastest computer and most concise description* [118].
There is a second consideration that renders the generalization (147) far from trivial: the time–span of the phenomenon, m, is not known *a priori*, nor it is constant over different events: the integer m in equation (147) *depends* on z_n. Because of the reasons just outlined there is no significantly better way to compute m than to follow the dynamics until m reveals itself: that is, the phenomenon under investigation—the quake—terminates.
We will in this chapter propose a theoretical framework to extend the dynamical treatment of extreme events to this kind of systems. We shall do this by focusing on a specific example, a model of neural network that has been introduced to describe various phenomena occurring in the electrical activity of the cerebral cortex [116, 117, 119, 120]. This model draws its origin, object and motivation from a vast body of existing research involving mathematics, physics and physiology.
It is not intended to reveal new phenomena in the last field, although we will provide a new conceptual approach to a much studied problem, the statistical description of firing cascades, better known as neuronal avalanches [133, 140, 128, 136, 135]. Rather, the model just mentioned is particularly suited to introduce a formal advancement in the dynamical theory of extreme values, which is relevant to the neuronal case, among many others. In fact, at difference with the standard theory, we will consider extreme events in complex, many dimensional systems that are not restricted to a specific instant in time, but which evolve in a finite time interval.

9.2 Neuronal avalanches

It is now possible to return from the geophysical analogy to the paradigmatic example described here: *neuronal avalanches*. It is observed that electrical stimuli originate (in the presence of, or without a sensorial stimulus) and propagate in the

cerebral cortex to involve the *firing* of thousands of neurons. In fact, a neuron may either be silent, or *spike*, which means, depolarize and send an electric stimulus via synaptic connections to other neurons. It is established that this activity composes a fundamental part of the computation performed by the brain [145, 141, 130]. It can be recorded by suitable devices [128] and it appears as bursts of firings, at times well separated between each other, but more often adjacent. In certain instances this activity may be due to synchronized oscillations of large number of neurons [122], but more complicated patterns are frequently observed, somehow intermediate between the former and homogeneous, uncorrelated firing of individual neurons. A substantial amount of literature (see *e.g.* [133, 140, 128, 136, 135]) has shown that this regime can be described as *critical*, characterized by power laws and scaling relations. We focus our attention on both the critical and the sub-critical regime, but we introduce a seemingly new operational definition of avalanche.

This definition is akin to what Rangan and Young [124, 125] have termed *multiple firing events*, as *the maximal sequence of firings that are caused by a single neuron.* By simulating each neuron in a network (typically, several thousands of these are considered) as characterized by microscopic variables (voltages and conductances) evolving according to coupled ODEs, and by the technical expedient of setting certain synaptic time constants to zero, Rangan and Young have been capable of determining exactly which neuron causes the firing (also called the *spike*) of a post-synaptic neuron.

The situation in the model that we will study is different and more general, in that the combined electrical activity of many different pre–synaptic neurons can cause the spike of a neuron. Nonetheless, by studying the detailed dynamics of a single, unperturbed neuron—in particular, the *basin of attraction* of the spiking event—we will be able to introduce a rule according to which the firing of a particular neuron can be identified as the *precipitating cause* of the spike of another. We shall then define a *firing cascade* as the *tree* that mathematically encodes such chain of excitations. We believe that this approach is general enough to be applicable to a wide set of complex systems composed of many interacting components.

Typical characteristics of these trees/cascades are intensity functions H like their intensity (number of nodes/number of firings), duration, and degree (number of levels in the tree). Extreme value theory, as described in the previous section, will be applied to these functions, both from the point of view of exceedances and of block maxima (concepts to be made precise below), thereby providing a concrete example of the abstract theory. We first describe in part (9.3) the dynamical model

of neural network that we adopt, borrowed from [120]. It consists of a collection of two–dimensional maps, each representing an individual neuron, initially proposed by Rulkov [116, 117]. In parts (9.4) and (9.6) the theory follows closely, with minor modifications, the original references [116, 117, 119, 120]. This is a necessary introduction, not only to set notation and define concepts, but also to investigate a few details of dynamical import: in particular, the *basin of attraction of the spiking event*, described in part (9.5), which is instrumental to define causality interactions. Neurons do not work in isolation, but are linked in a sparse network, where different neurons interact only when the spiking state is reached. In part (9.6) we describe the model of network connections, following [119, 120], and then in part (9.7) the collective dynamics that takes place on it.

We start in part (9.7.1) from the problem of the existence of an invariant measure that we heuristically replace by Birkhoff time averages, having nonetheless in mind that this procedure might be dubious due to possible long transients in the dynamics [137]. Part (9.7.2) contains the fundamental definition of *firing cascades as trees of causal excitations*, as well as its practical implementation, which is based on the concept of basin of attraction of the spiking event. This permits to define and measure important quantities like the intensity S of a cascade, its degree/generation level G and its time span T. In part (9.7.3) the statistics of these quantities is analyzed, showing a transition to critical behavior as the intensity of coupling among different neurons is increased; the exponents of the power–laws describing the above quantities obey critical scaling theory. From the point of view of extreme value theory, this analysis can be classified as statistics of exceedances.

Finally, in part (9.8) we further widen the scope of our investigation: we show the emergence of Extreme Value Laws in the statistics of extrema in time–intervals of finite length. This provides scaling relations that also permit to compute the probability that all cascades in a certain time interval have intensity less than a given threshold.

The conclusions briefly summarize the work, and an Appendix describes in an elementary way the stability regions, in parameter space, of the fixed point of Rulkov map.

9.3 The dynamical model: a sparse network of Rulkov maps

A variety of models have been proposed to study numerically neuron dynamics. Many of these are random. Purely dynamical models, on the other hand, can be

continuous time models, involving systems of differential equations [129]. They are able to reproduce a whole variety of observed biological patterns [124, 125, 126], but are expensive from a computational point of view. On the contrary, Rulkov 2d-iterative neurons model [116, 117] is a good compromise between capacity of well reproducing real biological behaviors, such as irregular spiking and bursts of spikes [117, 119] and fast computations, two features that are crucial for analysis. In this section, we briefly describe the dynamics of a single neuron, which is effected by Rulkov map. We follow the implementation presented in references [119, 120] that also contains illustrative pictures. We need only to mention that, at difference with these references, we do *not* allow a random external input: our dynamical system is fully deterministic.

Rulkov map describes the evolution of a fast variable x_n (n is integer time) that models the membrane potential of a neuron, and a slow control variable y_n that, although it is void of physiological meaning, is required to pilot the evolution of the former, so to reproduce some patterns observed in real experiments. This pair of variables describes the state of every neuron in the system. We label neurons by the variable $i = 1, \ldots, N$, so that the phase space of the system is $M = \{(x^i, y^i), i = 1, \ldots, N\}$. Recall that n labels time; Rulkov map is then defined as follows:

$$\begin{cases} x^i_{n+1} = F(x^i_n, y^i_n, I^i_n, x^i_{n-1}) \\ y^i_{n+1} = y^i_n - \mu(x^i_n + 1) + \mu\sigma + \mu I^i_n. \end{cases} \tag{148}$$

The variable I^i_n is the synaptic input, a *current*, acting on neuron i at time n, defined below in equation (155). Different neurons are coupled only via this term. It is convenient to first discuss the dynamics when this external input is set to zero and each neuron evolves independently of the others.

9.4 Unperturbed dynamics of the single neuron

In this subsection, accounting for the fact that I^i_n is set to zero and neurons do not interact, we drop the neuron index i for clarity. The function F in equation (148) is defined as follows:

$$F(x_n, y_n, I_n, x_{n-1}) = \begin{cases} \frac{\alpha}{1-x_n} + y_n + \beta I_n & \text{if } x_n \leq 0 \\ \alpha + y_n + \beta I_n & \text{if } x_n > 0 \text{ and } \{x_n < \alpha + y_n + \beta I_n \text{ and } x_{n-1} \leq 0\} \\ -1 & \text{if } x_n > 0 \text{ and } \{x_n \geq \alpha + y_n + \beta I_n \text{ or } x_{n-1} > 0\} \end{cases} \tag{149}$$

The parameters $\alpha, \beta, \mu, \sigma$ appearing in equations (148),(149) are all positive.
The parameter β modulates the relative influence of the synaptic input I_n with respect to y_n in equation (149). We choose $\beta = 0.133$. Notice that a memory effect is present in the above definition, via the value of the variable x at time $n-1$.
The parameter μ is taken to be small ($\mu = 10^{-3}$), so that the variable y evolves slowly with respect to time, following the second equation in (148). Also, its values cover a limited range. The combination of μ with α and σ determines the structure of motions, as discussed in detail in [116, 117]. In Appendix we sketch a simple analysis of the stability of the fixed point $(\sigma - 1, \sigma - 1 + \frac{\alpha}{\sigma - 2})$. Stoop *et al.* [120] choose $\alpha = 3.6$, so that only σ is left free to vary. As shown in Appendix, under these conditions the map admits a bifurcation at $\sigma = \sigma_{cr} = 2 - \sqrt{\frac{\alpha}{1-\mu}}$. Parameter space is accordingly parted in four regions, I to IV, plotted in figure (31) in the Appendix. Choosing $\sigma = 0.09$ puts the parameter-space point in region II, implying an attracting fixed point, while $\sigma = 0.103$ leads to region III, where the fixed point is unstable. This permits to define two kinds of neurons: Intrinsically Non-Spiking Neurons (INSN) (parameters in region II, motion in the left frame in figure (30)) and Intrinsically Spiking Neurons (ISN) (region III, right frame). These latter are the initiators of the dynamics of the network. The different behaviors are apparent in figure (30) where we plot motions in the single neuron phase space (as in figure (1) in [120], the y axis is in abscissa and the x axis in ordinate).
A neuron is said to *fire* when its membrane potential x_n reaches a local positive maximum (depolarization) and then suddenly drops to the hyperpolarized state $x = -1$: in physiology one talks of increasing polarization when the potential becomes more negative. At this moment a signal is transmitted to the other neurons, as described in the next two sections. It is instructive to follow the motion in the right panel of figure (30). Three cycles are plotted (two are almost coincident and require attention to be distinguished), each of which contains a spike. Start from the rightmost part of a cycle and follow points rotating counter-clockwise, so that y_n initially decreases and x_n increases. This happens because of the first alternative in equation (149). The map (148) is defined in such a way that when x becomes positive at time n (x_{n-1} being negative: draw a horizontal line at ordinate

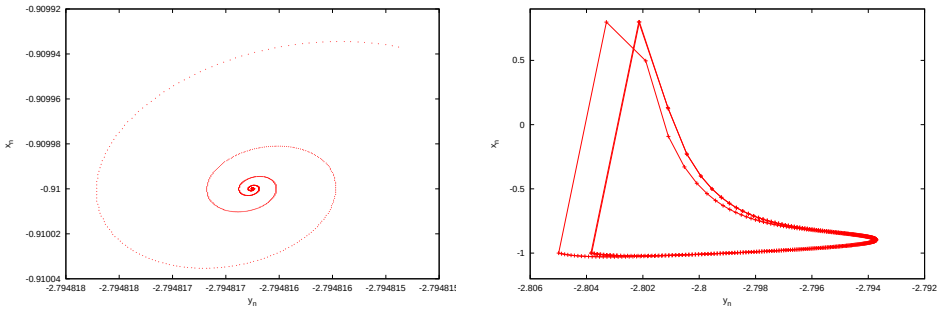


Figure 30: Typical trajectories in the single neuron phase–space (y, x) of an isolated INSN (left) and of an isolated ISN (right). Lines between points are merely to guide the eye. Motions happen counter-clockwise in both frames. Notice the different range of the axes; in the INSN case the figure is zoomed about the fixed point of the map.

$x = 0$ in the figure for visualization), either x_n is larger than $\alpha + y_n + \beta I_n$, so that at the next step the membrane potential drops to the negative value $x_{n+1} = -1$ (third alternative), or else it further increases to the value $x_{n+1} = \alpha + y_n + \beta I_n$ (second alternative), before achieving the value $x_{n+2} = -1$ at the following iteration, again in force of the third alternative. This motivates the introduction of the spike variable ξ_n that is always null, except at those times when the neuron potential reaches the hyperpolarized state, in which case it takes the value one. It can be formally defined as follows:

$$\xi_{n+1} = \begin{cases} 1 \text{ if } F(x_n, y_n, I_n, x_{n-1}) = -1, \\ 0 \text{ otherwise.} \end{cases} \tag{150}$$

The spike variable ξ_n will be used later in part (9.6) to define the dynamical action of synaptic coupling. Let us now study the stability of the fixed point of Rulkov map. It is easy to derive the fixed point of Rulkov map and its stability. Consider again an isolated neuron, and assume that we are in the first alternative of equation (149), which yields the map

$$\begin{cases} x & \to & \frac{\alpha}{1-x} + y \\ y & \to & y - \mu(x+1) + \mu\sigma. \end{cases} \tag{151}$$

From this, the unique fixed point $(\bar{x}, \bar{y})$ easily follows:

$$\begin{cases} \bar{x} & = & \sigma - 1 \\ \bar{y} & = & \sigma - 1 - \frac{\alpha}{2-\sigma} \end{cases} \tag{152}$$

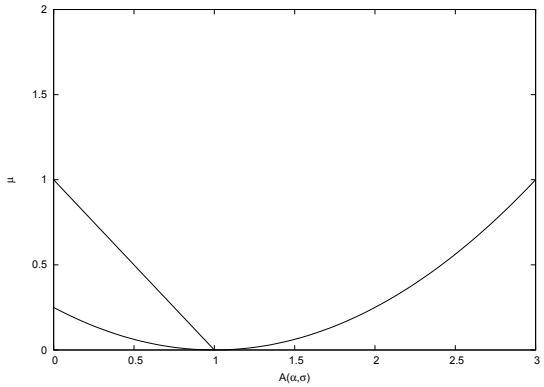


Figure 31: Stability regions of the Rulkov map (152) in parameter space. Moving from the bottom left corner $(0, 0)$ to the bottom right $(3, 0)$ one encounters regions I to IV in order. See text for further details.

This result is consistent with equation (149) provided that $\bar{x} < 0$, that is, $\sigma < 1$. The differential of the transformation, *i.e.* the linear map from the tangent space at the fixed point to itself is

$$J = \begin{bmatrix} \frac{\alpha}{(2-\sigma)^2} & 1 \\ -\mu & 1 \end{bmatrix} \tag{153}$$

Putting for simplicity $A(\alpha, \sigma) = \frac{\alpha}{(2-\sigma)^2}$ the linear map J has real eigenvalues provided that $A(1 - (\alpha, \sigma))^2 \geq 4\mu$. In the opposite case, complex conjugate pairs appear. In this case, the square modulus of the eigenvalues equals the determinant of J and is

$$\text{Det } (J) = A(\alpha, \sigma) + \mu.$$

This situation is pictured in figure (31): $A(\alpha, \sigma)$ and μ are always positive and are taken as coordinates in the (A, μ) plane. When the corresponding point lies below the parabola $(1 - A(\alpha, \sigma))^2 = 4\mu$, J has two real eigenvalues. If moreover $A < 1$ (region I) their absolute value is less than one and the (hyperbolic) fixed point is attractive, while it is repulsive when $A > 1$ (region IV). In the region above the parabola and below the line $A(\alpha, \sigma) + \mu = 1$ (region II) J has a pair of complex conjugate eigenvalues of modulus less than one, so that the fixed point is still attractive, and trajectories spiral down to it. When finally $A(\alpha, \sigma)$ is above both the parabola and the straight line (region III) two eigenvalues of modulus larger than one appear: the fixed point is unstable, and trajectories spiral out of it.

9.5 Perturbed dynamics of the single neuron: basin of attraction

The synaptic input I_n received by a neuron can either be stochastic or deterministic, as in our case. How this latter is generated in our model is inessential for the present discussion and will be explained in the next section; suffices here to observe that I_n enters equation (149) linearly. In particular, a positive current I_n increases the potential x_n (hence, it depolarizes the neuron) and can therefore be associated with an *excitatory* action, while the reverse is true for a negative current I_n. Notice however that this rule of thumb has non–trivial exceptions that we will describe momentarily.

Figure (32) shows how the dynamics of both ISN and INSN observed in figure (30) is perturbed by I_n. Typical trajectories are shown, in which sharp angles appear when I_n is non–null. Notice however that time being discrete, lines in the figure join distinct points–albeit typically very close, so to simulate a continuous trajectory that does not exist. In the left frame one can observe how INSN, silent when isolated, can be triggered to spike by the input I_n. The dynamics of ISN (right panel) can also be perturbed, altering their regular spiking regime. In fact, consider figure (33), left panel, where the potential x_n of an ISN neuron (red curve) is plotted versus time n, together with the synaptic input I_n (green curve). The synaptic current I_n is positive, stemming from an excitatory action. Three excitations are visible in the range displayed. Prior to these the ISN neuron spikes regularly with a period of about 242 iterations. The first excitation has the effect of anticipating the successive spike by 35 iterations. In the right panel, which shows the same motion in the single neuron phase–space, regular oscillations appear as a thick limit cycle. The first spike is evident in the thread which emerges from the cycle at the right of the picture, and moves initially by increasing both x_n and y_n, according to the fact that the excitation I_n is positive: see equation (149). Returning to the left panel, we notice two more excitatory inputs to the neuron, which have the paradoxical effect of suppressing further spikes! This can be explained by observing, in the right panel, two more sharp angles in the trajectory, the first starting form the bottom left of the picture. In both cases x_n and y_n increase, but this has the effect of pushing the trajectory towards the fixed point of the map, which is unstable but from whose neighborhood the motion requires a long time to escape.

We have so discovered that excitatory actions can sometimes produce an inhibition of spikes, due to the particular nature of the phase–space of the system. For a similar reason, inhibitory actions can sometimes induce spikes. This paradox leads

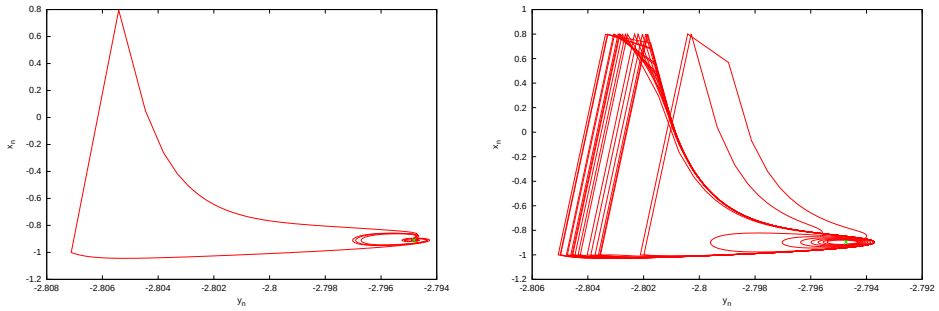


Figure 32: Typical trajectories in the single neuron phase–space (y, x) of INNS (left) and ISN (right) neurons embedded in a network. Marked by a symbol are the fixed points of the Rulkov map. Motions takes place counter-clockwise in both frames.

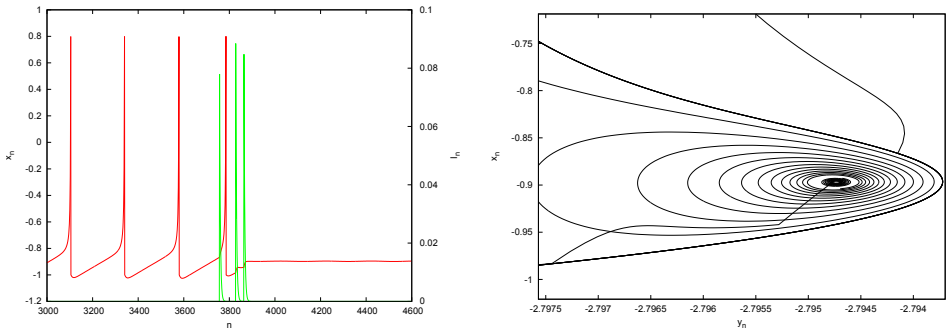


Figure 33: In the left frame, the potential x_n of an ISN neuron (red curve) is plotted versus time n, together with I_n (green curve). The right frame plots the corresponding phase–space trajectory. See text for further discussion.

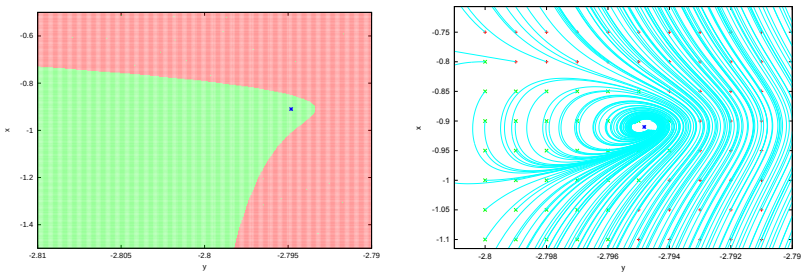


Figure 34: In the left frame, basin of attraction of the spiking event, in red, and basin of attraction of the fixed point in green. The fixed point is marked in blue. The right frame shows the detail with trajectories plotted.

to a deeper understanding of the dynamics of this system.
In fact, consider an INSN, like those pictured in the left panels of figures (30) and (32). Let (y, x) be the initial condition of the motion and let $I_n = 0$ (isolated neuron). Even if finally attracted by the fixed point, some trajectories will spike before that happens. In the left frame of figure (34) we color in red such initial conditions, and in green those that head directly to the fixed point. A blue asterisk marks the fixed point. A limit cycle separates the two regions. In the right frame, the phase space portrait near the stable fixed point is displayed. It is then clear that an inhibitory action (reducing both x_n and y_n) can move a point from the green to the red region and that an excitatory kick can do the opposite, but it is equally clear that this occurs in a restricted region of phase space. In fact, increasing x typically moves a phase–space point from the green to the red region, *i.e.* from the basin of the fixed point to that of the spike. In particular, when the potential x of a NISN is larger than a certain threshold x_{th}, the paradoxical effects just described are not possible. In fact, all initial conditions with x_n larger that x_{th} lie in the red region and lead to firing, even in the absence of further excitation: the effect of the current I_n that puts a NINS in the red region can therefore only be reversed by an inhibitory input. This fact will be of paramount importance in the following. In the range of parameters of our experiments we can safely take $x_{th} = -0.7$.

9.6 Network topology and synaptic coupling

According to the way in which they act upon the others, neurons can be classified either as *excitatory* or *inhibitory*. Following [120] we consider a network of size N with a ratio of inhibitory/excitatory neurons r, so that the first $n_{ex} = \lfloor Nr \rfloor$ neurons are excitatory and the remainder $N - n_{ex} = n_{in}$ are inhibitory. Furthermore, as

described in the previous section, neurons can either be intrinsically spiking or not, so that each group (excitatory and inhibitory) can also be parted in two: out of the n_{ex} (n_{in}) excitatory (inhibitory) neurons m_{ex} (m_{in}) are chosen to be intrinsically spiking.

Neurons are connected so that each of them is acted upon by p_{ex} excitatory neurons and p_{in} inhibitory neurons, chosen at random in the network—with no regard on whether ISN or NISN, with the only constraint that no neuron can act upon itself. The six integer numbers n_{ex}, m_{ex}, n_{in}, m_{in}, p_{ex}, p_{in}, plus the connection diagram fully specify the network. Analytically, this is formalized by the adjacency matrix ω_{ij}, defined by

$$\omega_{ij} = \begin{cases} 0 & \text{if neuron } j \text{ does not act on neuron } i, \\ 1 & \text{if neuron } j \text{ has excitatory action on neuron } i, \\ \psi & \text{if neuron } j \text{ has inhibitory action on neuron } i. \end{cases} \tag{154}$$

The factor $\psi > 0$ serves to differentiate the intensity of inhibitory with respect to excitatory action. We choose $\psi = 3$.

Coupling in the network is driven by action potentials generated by firing neurons. Following again [120] the spike of a presynaptic neuron j determines the input I_n^i received by the neuron i according to the linear difference equation:

$$I_{n+1}^i - \eta I_n^i = W \sum_{j=1}^{N} \omega_{ij}(\chi_{ij} - x_n^i)\xi_n^j. \tag{155}$$

W is a crucial intensity parameter, varying which we can increase the effect of synaptic excitation: recall that I_n^i enters the dynamical equation of evolution of neuron i, equations (148),(149). The form of equation (155) can be explained in physiological terms. First observe that the homogeneous equation (with null right hand side) implies an exponential decay: $I_n \sim \eta^n$ (we choose $\eta = 0.75$). The spike variables ξ_n^j at r.h.s. are the source of the currents I_n^i. In equation (155) we notice that when the synapse reversal potential χ_{ij} is larger than the potential x_n^i, its activation (that happens when $\xi_n^j = 1$) increases the synaptic input I_n^i and also, because of equation (149), the membrane potential x_{n+1}^i. This is equivalent to a depolarization of the neuron and the synapse acts as excitatory. In the opposite case, the activation hyperpolarizes the neuron and the action is inhibitory. The numerical values of the above parameters are chosen as in [119, 120], where their physiological relevance is discussed: the reversal potential is null for excitatory connections (*i.e* $\chi_{ij} = 0$ when neuron j is excitatory) and equal to $\chi_{ij} = -1.1$

when neuron j is inhibitory, for all i. This ends the description of the *laws of motion* of the network.

9.7 Network dynamics

We can now begin to study the dynamical behavior of the full network. We start by introducing a basic problem: the existence of an invariant measure.

9.7.1 Invariant measure

Equations (148),(149),(150),(154) and (155) define a multi-dimensional dynamical system in the phase–space $M = \mathbb{R}^{2N}$ with variables $(x^i, y^i), i = 1, \ldots, N$. Let $z = (x, y)$ be such $2N$–component phase–space vector and let T be the transformatiion. A fundamental ingredient of dynamical theory are the so–called invariant measures, that is, positive Borel measures that are preserved by the dynamical evolution. Lacking a theoretical control of these measures in many–dimensional systems like the one under consideration, we substitute phase–space averages by Birkhoff sums over trajectories. In measure–preserving systems, the two coincide almost everywhere (with respect to the initial point of a trajectory, under the invariant measure).
More precisely, if we consider an intensity function $\phi(z)$ (for simplicity of notation, from here on we write ϕ for ϕ, see equation (147)), we construct its distribution function $F_\phi(h)$ as follows: we start from a random point $z^* \in M$. We iterate the dynamics T for an initial number of iterations n_0 sufficiently large to let the system reach a sort of dynamical equilibrium. At that moment we let $z_0 = T^{n_0}(z^*)$ be the initial point of a trajectory and we compute the sample values $\{\phi(T^n(z_0)), n = 0, \ldots, L-1\}$, by which the distribution function F_ϕ is

$$F_\phi(h) = \lim_{L\to\infty} \frac{1}{L} \#\{\phi(T^n(z_0)) \leq h,\ 0 \leq n < L\}. \tag{156}$$

Clearly, in numerical experiments the limit is replaced by the value obtained for a large but finite sample length L. We shall also use the complementary distribution function $E_\phi(h) = 1 - F_\phi(h)$: this is the measure of the tail of the distribution of values of ϕ:

$$E_\phi(h) = \mu(\{z \in M \text{ s.t. } \phi(z) > h\}) \tag{157}$$

The study of coupled map lattices has revealed puzzling phenomena like stable motion with exponentially long chaotic transients [137, 138], so that the numerical

procedure just explained might be theoretically flawed. Nonetheless, the result of [120] on small networks ($N = 128$) revealing a transition to positive Lyapunov exponents lead us to believe that convergence in equation (156) can be readily obtained.

9.7.2 Firing cascades as trees of causal excitations

A crucial difference between conventional coupled map lattices and the network under study is that in this latter interaction between different maps happens only at specific instants of time, while for most of the time individual systems evolve independently of each other. When it comes to physiological networks, the first and foremost consequence of this fact is that information can be stored, elaborated and transmitted [130]. In a dynamical perspective this calls attention to collective behaviors like synchronization [122, 127]. Here we focus on a more elusive set of phenomena: neuronal avalanches, which we prefer to call firing cascades.

Most of the studies on the collective behavior of neural networks define avalanches as the uninterrupted activity of the network during a certain time interval [133, 135]. Statistical properties of these collective events have been extensively studied in a large series of works, both experimental, see for example [133, 140, 128, 136, 135], and theoretical, with the aid of model systems [124, 125, 126, 131, 141, 142, 143, 144, 145, 146], but statistical analysis alone does not clarify the dynamical detail of the activity network and, in particular, of the impulse propagation. Moreover, it suffers from a degree of arbitrariness in the definition of avalanche. In fact, coarse–graining of time is usually employed, for spiking events appear at particular instants of time and must be grouped to define the quiescence periods separating different avalanches. Figure (35) shows a conventional *raster plot* of spikes, depicted as red points at coordinates (n, i) where n is the time at which neuron i spikes. One observes regions with different density of red spikes. The usual procedure [133] is to partition the time axis into different bins, so that a sequence of adjacent, non–empty bins constitutes a cascade. The arbitrariness in the choice of the bin length is apparent, and its influence on the resulting statistical properties has been well described in [133]. Moreover, different regions of the network even if uncorrelated but spiking at close instants may be classified in the same cascade. In short, the true dynamics of the network is overlooked in this approach. This is justified in experimental data, when no access to the network of connections is available.

Therefore, we pursue here a different approach, based upon the detailed knowledge

of the dynamics presented in the previous sections. We have already remarked that different types of neuron are present in the network: in particular, ISN are the origin of propagating excitations. When one of these neuron spikes, say neuron j, it acts upon its post-synaptic neurons via equation (155). As observed in part (9.3), if such postsynaptic neuron is also an ISN, this action will alter its dynamics, anticipating or delaying its next spike. Otherwise, if the post-synaptic neuron i is of the kind NISN, this excitation might eventually drive it to spike, this occurrence following from the dynamics generated by equation (148). The induced spiking, when it takes place, is not instantaneous: according to the coordinates (x_n^i, y_n^i) at the time of excitation, the motion requires a certain amount of time to exit the region close to the fixed point and spike. It is nonetheless possible to estimate with a certain advance if this event will take place, using the notion of basin of attraction of the spiking event and the threshold x_{th} discussed in part (9.3). Recall that this is defined as the set of initial conditions of the *isolated non–spiking* neuron which lead to a spike, before returning to approach the fixed point. When x_n is pushed by a synaptic excitations into this basin (red in figure (34)) the NISN will fire, regardless of further excitatory actions: such spike can only be avoided by inhibitory actions.

This reasoning permits to draw a *causal* relationship between the spiking of a NISN neuron and *a single originating excitation* from another neuron. Even if it is clear that in a complex network many pre–synaptic neurons can contribute to set the dynamics of i on course to firing, we stipulate to ascribe the cause of the spike of neuron i to a single neuron, i', the one which is the last to act on i, at time n', before x^i enters the basin of attraction of the spiking event, to eventually fire at time n. Excitations following this time are not considered in this count, because they are not necessary to induce a spike, which would happen nonetheless. Equally not considered are impulses preceding such precipitating event—the straw that broke the camel's back—because if taken alone they might not be enough to cause excitation. As a simpler rule, that can also be applied in experiments, we approximate the basin of attraction by the fact that x^i exceeds a threshold value x_{th}. This procedure can be formally summarized as follows: we say that the spike of neuron i' at time n' causes the spike of neuron i at time n, and we write $(n', i') \to (n, i)$ if

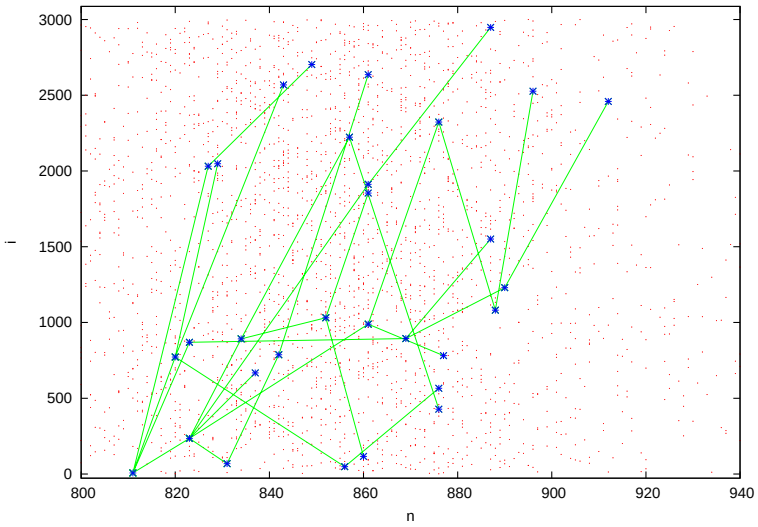


Figure 35: Tree construction. Red dots mark neuron firing events (n, i); green lines join causally related events (blue symbols). See text for description.

$$
\begin{aligned}
(n', i') \to (n, i) \iff \quad & \xi^i_n = 1,\ \xi^{i'}_{n'} = 1, \\
& \exists\ n^*,\ n' \le n^* < n, \text{ such that the following holds:} \\
& x^i_m < x_{th} \text{ for } n' \le m < n^* < n, \\
& x^i_m \ge x_{th} \text{ for } n^* \le m \le n, \\
& \xi^j_m = 0 \text{ for } n^* \le m \le n,\ \forall j \text{ such that } \omega_{i,j} > 0.
\end{aligned}
\tag{158}
$$

Following this rule we can therefore create trees of excitations, which start at the firing of a single ISP neuron and propagate through the network. Any such tree is defined to be a *firing cascade*. Return for illustration to figure (35), which also shows a realization of this construction. Green lines join causally related spikes $(n', i') \to (n, i)$ (blue symbols), on the background of red dots that mark all spiking events. The picture shows how the initial excitation of a spiking neuron ($i' = 7$ at $n' = 811$) propagates to non–spiking neurons ($i = 772$ at $n = 820$, $i = 235$ and $i = 869$ at $n = 823$, $i = 2032$ at $n = 827$) that, on their turn, stimulate further NISNs to fire (for a total of 31 spikes in the figure). Clearly, not all *physical* connections are dynamically activated in a spike—for otherwise the cascade will have no end. In the example, neuron $i' = 7$ is also presynaptic to $i = 908$, $i = 1115$ and $i = 2677$ that remain silent.

9.7.3 Statistical analysis of firing cascades

The dynamical approach described in the previous section permits to define precisely and measure accurately various quantities usually defined via time binning. These are:

- S, the size of a cascade, that is, the total number of spikes in a cascade—equally, the number of nodes in a causal tree;
- T, the time span of a cascade: the length of the time interval between the first spike and the last;
- G, the generation number, which is the number of levels of the tree—also, the length of the longest chain of successive excitations.

We are so equipped with the tools to analyze the statistics of cascades in the neuronal network: the functions H under considerations will be either one of S, T, G. As mentioned in part (9.7.1), we replace phase–space averages by Birkhoff sums. We consider time-segments of length $L = 4 \cdot 10^6$ iterations, following an initial transient (discarded) consisting of 6000 iterations. Occasionally we raise this number to account for the fact that long transient can be observed in extended systems [137]. Using a parallel architecture, we accumulate the results of 32 realizations of the initial conditions. We performed an extensive investigation, of which we report here a typical instance.

Example 1 *$N = 3000$ neurons, with $n_{ex} = 2400$, $m_{ex} = 2$, $n_{in} = 600$, $m_{in} = 1$, $p_{ex} = 4$, $p_{in} = 2$, Physical parameters are: $\beta = 0.133$, $\alpha = 3.6$, $\sigma_{NINS} = 0.09$, $\sigma_{ISN} = 0.103$, $\mu = 10^{-3}$, $\psi = 3$, $\chi exc = 0$, $\chi_{inh} = -1.1$.*

Since this model neither includes pruning nor adaptation, to achieve criticality one needs to tune the parameters of the network. Following [120], we vary the coupling intensity W in equation (155). We first compute the *average number of events per-unit time* τ^{-1}, observing a rather sharp transition as we increase W: when $W = 0.084$ one has $\tau^{-1} = 3.6 \cdot 10^{-2}$, which increases to $\tau^{-1} = 5.68 \cdot 10^{-1}$ for $W = 0.087$ and $\tau^{-1} = 5.65$ for $W = 0.09$. At the largest coupling, roughly five neurons (out of the total 3000) spike on average at each single (discrete) time. This transition is clearly reflected in the statistics: in figure (36) we plot the fraction $\rho_S(S)$ of firing cascades with S spikes and its (complementary) distribution function $E_S(S)$ defined in equation (157). Two values of the coupling constant W are plotted in figure (36): $W = 0.087$ and $W = 0.09$. The former is visibly an

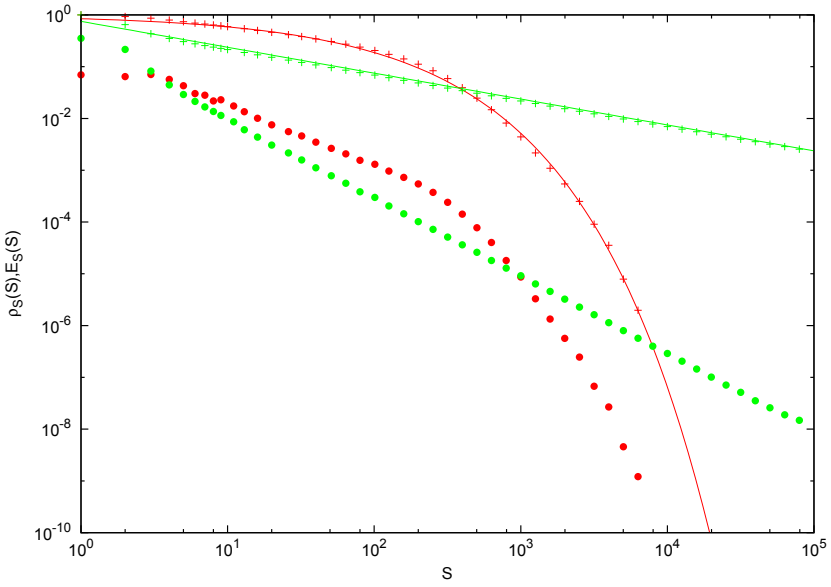


Figure 36: Cascade statistics in the case of example (1). Fraction of cascades with S spikes, $\rho_S(S)$, versus S, for $W = 0.087$ (red bullets) and $W = 0.09$ (green bullets) and the complementary distribution functions $E_S(S)$ (red and green crosses, respectively). The fitting curves (red and green) are described in the text.

under-threshold case, the latter displays a full-fledged cascade which extends over orders of magnitude. In the first case the distribution $E_S(S)$ is well fitted by a stretched exponential: $E_S(S) \simeq \exp\{-AS^B\}$ with $A \simeq -.1694$, $B \simeq .497$ (red curve). In the second, the power-laws $\rho_S(S) \simeq C'S^{-\beta_S}$ and $E_S(S) \simeq CS^{-\beta_S+1}$ are well satisfied, with exponent $\beta_S = 1.5$ ($C = 0.75$, green line).

In the case of large coupling, $W = 0.09$, we plot in figure (37) the statistical distribution $\rho_G(G)$ of generation levels G and of time span T, $\rho_T(T)$. For both of these quantities we observe a power–law decay with exponents β_G and β_T close to two. In the same figure we also plot the inverse, for graphical convenience, of the quantity $\bar{S}(G)$, the average number of spikes in a cascade with G levels, which approximately follows a power–law with the same exponent: $\bar{S}(G) \sim G^\gamma$, $\gamma \simeq 2$.

The observed exponents $\beta_S = 3/2$ and $\beta_T = 2$ are to a good approximation equal to those appearing in physiological experiments [133, 135] and in various models of neuronal networks [134, 131]. To the contrary, these values differ from those obtained in [120] for the system that we are investigating, but with the conventional definition of cascade. It is also to be quoted the measurement of these exponents in high resolution experimental data [136], still within the time–binning approach. Critical scaling theory predicts [121, 136] that $\beta_T - 1 = (\beta_S - 1)\gamma'$, where γ' is defined in terms of the scaling of the average of S at fixed time length T, versus

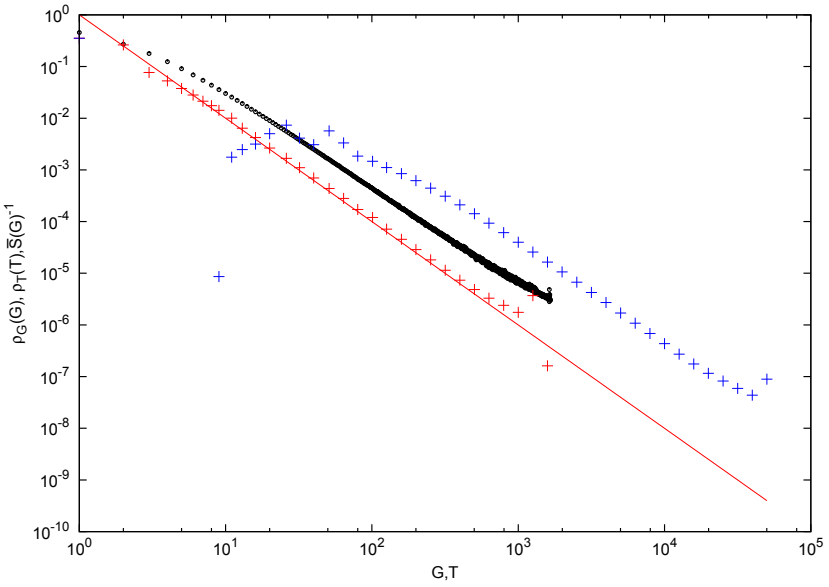


Figure 37: Fraction $\rho_G(G)$ of cascades comprising G generations (red crosses) and fraction $\rho_T(T)$ of cascades extending over T units of time (blue crosses), for $W = 0.09$. The red line is the power–law decay $\rho_G(G) = G^{-2}$. Also plotted is the inverse of the quantity $\bar{S}(G)$ versus G (black bullets).

T. The exponents derived in [120] verify this scaling relation. Since in our model the time span of a cascade T and its generation level G are proportional, we can identify the exponents γ and γ', providing the relation

$$\rho_S(S) \sim S^{\beta_S}, \quad \beta_S = \frac{\beta_G + \gamma - 1}{\gamma}, \tag{159}$$

From the data of figure (37) we see that approximately $\gamma = 2$, so that the relation (159) for the triple $\beta_S = 3/2$, $\beta_G = \gamma = 2$ is rather well satisfied by the numerical data.

9.8 Statistics of extreme events in a finite time–interval

The excitation of intrinsically spiking neurons propagates sometimes to a large number of neurons and generates a cascade. In the previous section we have studied the relative frequency of these cascades versus their intensity, in a theoretically infinite time–interval. It is now interesting to consider the probability that none of these cascades, within a certain *finite* observational interval, exceeds a certain intensity threshold—equivalently, the probability that the *largest* cascade in a time–interval is smaller than such threshold. This investigation belongs to the realm of extreme value theory and goes under the name of statistics of *block*

maxima.

Let in fact n be the time at which an excitatory ISN fires and generates a cascade (possibly, and frequently, limited to such single spike). To this value we associate an intensity, $I(n)$, which can be chosen to be the size of the cascade S, its level G, or its time span T. If more than a cascade is originated at time n, $I(n)$ is the maximum of the values corresponding to such cascades. To the contrary, if no cascade starts at time n (*i.e.* no excitatory ISN spikes at time n) $I(n)$ is null. Since the system is deterministic, $I(n)$ is actually a function of the phase–space point z at which the trajectory is located when one starts recording these intensities:

$$I(z,n) = \phi(z_n) = \phi(T^n(z)) \tag{160}$$

Following standard usage we then define the block maxima function $\phi_L(z)$, for various values of L, as the largest intensity of a cascade that originates in the time–window $n = 0, \ldots, L-1$:

$$\begin{aligned} \phi_L(z) &= \max\{I(z,0), I(z,1), \ldots, I(z,L-1))\} \\ &= \max\{\phi(z), \phi(T(z)), \ldots, \phi(T^{L-1}(z))\}. \end{aligned} \tag{161}$$

The cascade with such maximum intensity is therefore appropriately called the *extreme event* in the time window under consideration. The wider is the time–window, the larger is presumably its intensity. An *Extreme Value Law* for $\phi_L(z)$ is an asymptotic form, for large L, of the statistical distribution of this variable. Let us now show that it applies to the present case.

As indicated, $\phi_L(z)$ depends on the position z in phase–space M and it is therefore convenient to consider the distribution of the values of $\phi_L(z)$ with respect to an invariant measure μ. As discussed before, we assume that the distribution function $F_{\phi_L}(h) = \mu(\{z \in M \text{ s.t. } \phi_L(z) \leq h\})$ can be obtained numerically by Birkhoff time averages over trajectories of the system.

Suppose that different cascades arise independently of each other, that is, correlation of the values of $\phi(T^j(z))$ for different j decays quickly when their time difference increases. This might also due to the fact that different cascades involve separate sets of neurons. Under this assumption the probability that $\phi_L(z)$ is less than a certain value h, which is clearly equal to the probability that the intensity $\phi(T^n(z))$ of *all* cascades in the time interval $n \in [0, L-1]$ is less than h, becomes a *product* of the probabilities of these individual events. If we let λ be the density

of cascades per unit time (not to be confused with the density of spikes per unit time), we expect a number λL of cascades to originate in a time–window of length L, so that

$$F_{\phi_L}(h) \sim (F_\phi(h))^{\lambda L}. \tag{162}$$

Next, taking logarithms, and letting h grow, so that $F_\phi(h)$ approaches one, we obtain

$$\begin{aligned} \log(F_{\phi_L}(h)) &\sim \lambda L \log(F_\phi(h)) \\ &\sim \lambda L (F_\phi(h) - 1) \\ &= -\lambda L E_\phi(h). \end{aligned} \tag{163}$$

Clearly, this is a heuristic estimate, but it yields the same results predicted by rigorous theory, which requires appropriate conditions [**?**] to dispose of the correlations of the motion that might render the independence assumption incorrect. We can verify a posteriori the fact that these correlations fade in the limit of large L by a direct evaluation of equation (163), when the asymptotic relation symbol is replaced by a much more stringent equality.

To do this, let us consider again the case of example (1) when $W = 0.09$. Dividing the total number of cascades by the total time span of the computation (averaged over 32 different runs) yields the value $\lambda = 8 \cdot 10^{-3}$, which is much smaller than $\tau^{-1} = 5.65$ and reveals that spikes are organized in cascades of average size $(\lambda\tau)^{-1}$, that is, roughly 22 spikes per cascade, in this case. In figure (38) we plot $-\log(F_{\phi_L}(h))/L$ versus h, for various values of L, for the three functions $\phi = G, S, T$. In all cases the data points collapse on a single curve, coinciding with $\lambda E_\phi(h)$, as predicted by equation (163). Observe that the value of λ employed in the figure does not come from a fit of the data, but it is the one just obtained by counting the number of cascades. The extreme value law given by equation (163) appears therefore to apply, at least in this regime.

It is to be remarked that in certain cases an additional positive factor θ, smaller than one, appears in front of $\lambda E_\phi(h)$. This factor is called the extremal index: it implies a slower decay of $F_{\phi_L}(h)$, originating from the tendency of extreme events to appear in clusters. Rigorous theory has studied to a large detail this phenomenon in simple systems [132], which has also been detected in many dimensional systems [105]. The data of figure (38) indicate a unit value of this parameter.

In the critical regime the three quantities $\phi = G, S, T$ are characterized by a power–law decay: $E_\phi(h) \simeq A_\phi h^{-\beta_\phi}$, where A_ϕ are suitable constants and the exponents β_ϕ have been discussed in the previous section. Inserting this information in equation

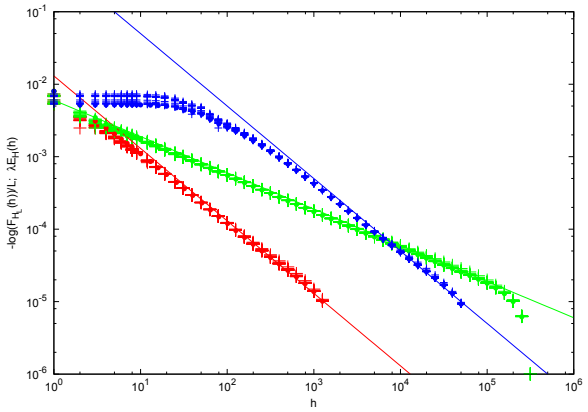


Figure 38: Extreme value law for the case of example 1, when $W = 0.09$. Plotted are the values of $-\log(F_{\phi_L}(h))/L$ (crosses) and $\lambda E_\phi(h)$ (bullets) versus h, and H is G (red), S (green) and T (blue). Values of L range from 10 to 10^4. Also reported are the power law decay laws: $0.013 \cdot h^{-1}$ (red), $0.006 \cdot h^{-1/2}$ (green) and $0.5 \cdot h^{-1}$ (blue).

(163) yields the relation

$$F_{\phi_L}(h) \simeq \exp\{-\lambda L A_\phi h^{-\beta_\phi}\}. \tag{164}$$

In figure (39) we plot the numerically determined function $F_{\phi_L}(h)$ when $\phi = S$ for values of L and h ranging over orders of magnitude and their difference with the analytical relation (164). This latter is remarkably small in all the range of data considered.

We can perform the same analysis in the subcritical case, example (1) and $W = 0.087$. In figure (40), which is the analogue of figure (38) for this case, we plot again $-\log(F_{\phi_L}(h))/L$ and $\lambda E_\phi(h)$, for $\phi = G, S, T$. Collapse of the data points is again observed, but for larger values of h than in the previous case. The asymptotic law for the distribution of the variable $H_L(h)$ becomes here

$$F_{\phi_L}(h) \simeq \exp\{-\lambda L \exp\{A_\phi h^{B_\phi}\}\}. \tag{165}$$

The experimental function $F_{\phi_L}(h)$ for $\phi = S$ is plotted versus L and h in the left frame of figure (41). Difference with the theoretical expression (165) is plotted in the right frame. As expected from figure (40), convergence improves as L and h increase.

Within the framework of Extreme Value Theory, the two cases just discussed differ because they imply different *normalizing* sequences $h_L(t)$, which we now define, while providing the same limiting distribution $F(t) = e^{-t}$. In fact, we observe that

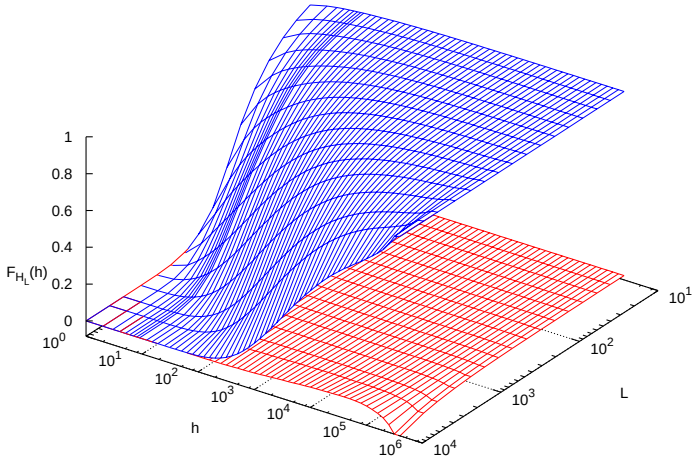


Figure 39: The experimental function $F_{\phi_L}(h)$ for $\phi = S$ versus L and h in the case of example (1) and W=0.09 (blue graph). Difference between the experimental function and the theoretical expression (164) (red graph).

equations (162) and (163) imply that

$$F_{\phi_L}(h) \sim \exp\{-\lambda L E_\phi(h)\}. \tag{166}$$

Let t be a positive, real variable. The sequence of levels $h_L(t)$, $L = 1, 2, \ldots$, parameterized by t, is defined so to satisfy the limit behavior

$$\lambda L E_\phi(h_L(t)) \to t, \quad \text{as } L \to \infty. \tag{167}$$

As a consequence of the above,

$$\mu(\{z \in M \text{ s.t. } \phi_L(z) \leq h_L(t)\}) = F_{\phi_L}(h_L(t)) \to F(t) = e^{-t}, \text{as } L \to \infty. \tag{168}$$

The two previous equations are the core results of EVT, when added to the fact that *three* different limiting distributions $F(t)$ can be achieved after further rescaling, as we do below in the critical case. They show that $h_L(t)$ is a quantile of the asymptotic distribution of the statistical variable $\phi_L(z)$, whose analytical form involves the ratio t/L and the inverse of the complementary distribution of the function ϕ:

$$h_L(t) \sim E_\phi^{-1}(\frac{t}{\lambda L}). \tag{169}$$

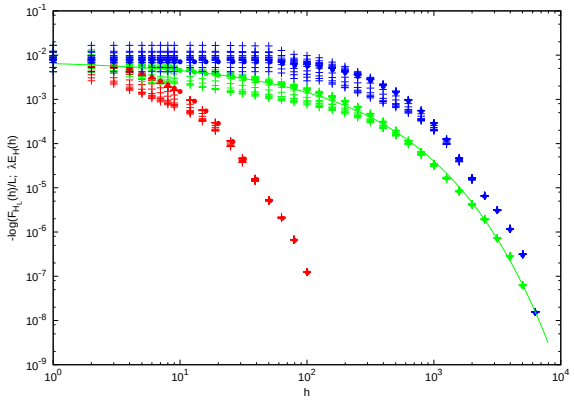


Figure 40: Extreme value laws in the case of example (1), when $W = 0.087$. The same data as in figure (38) are plotted: the values of $-\log(F_{\phi_L}(h))/L$ (crosses) and $\lambda E_\phi(h)$ (bullets) versus h, when ϕ is G (red), S (green) and T (blue). Values of L range from 10 to 10^4. Asymptotic coalescence of the data is observed. The green curve is the stretched exponential $E_S(S) \simeq \exp\{-AS^B\}$ with $A \simeq -0.1694$, $B \simeq 0.497$ already plotted in figure (36) and described in part (9.7.3).

In the subcritical case these quantiles grow as the power of a logarithm:

$$h_L(t) \sim \left(\log(\frac{\lambda L}{t})\right)^{\frac{1}{B_\phi}}. \tag{170}$$

At criticality a power–law behavior is present:

$$h_L(t) \sim (\frac{\lambda L}{t})^{\frac{1}{\beta_\phi}} \tag{171}$$

and any q moment of the statistical distribution of $\phi_L(z)$ scales as $(\frac{\lambda L}{t})^{\frac{q}{\beta_\phi}}$. In this latter case, the previous relations can be put in one of the three standard forms [113]. In fact, simple algebra permits to rewrite

$$\mu(\{z \in Z \text{ s.t. } (\lambda L)^{-1/\beta_\phi}\phi_L(z) \leq u\}) \to e^{-u^{\beta_\phi}}, \tag{172}$$

as $L \to \infty$, which is the expression of a Fréchet distribution.
We have in this chapter poposed an extension of the dynamical theory of extreme events to complex, multi–dimensional systems. We have applied it to a model of neuronal network, that was studied in full detail. This analysis attains wide generality because the example is characterized by features that are frequently encountered: it is composed of many almost identical subsystems, linked in a sparse, directed network; each individual system integrates the input from the network

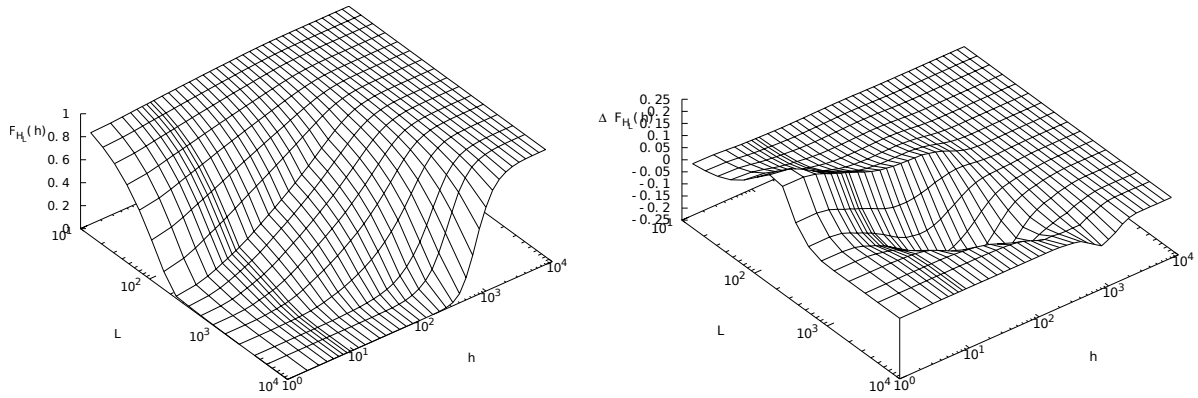


Figure 41: The experimental function $F_{\phi_L}(h)$ for $H = S$ versus L and h in the case of example (1) and W=0.087 (left graph). Difference between the experimental function and the theoretical expression (165) (right graph).

and in turn returns a signal to it; signal propagation can develop a full chain of excitations. This has lead us to introduce a novel approach to define neuronal avalanches, whose statistical characteristics were studied within the framework of extreme value theory. The same definition of avalanche can be obviously applied to other systems in the class just described. The presented results are in line with previous theoretical models and experimental findings on neuronal networks, although it is now intended to bring any novel contribution to this field. While the set-up that we propose to extend the dynamical extreme value theory to complex phenomena, evolving in time, is fully rigorous, our results are formal, even if confirmed by numerical experiments. This poses a new challenge to the mathematical analysis.

10 Conclusions

We have, throughout the thesis, contributed on several aspects to the active dynamical theory of extreme events.

We reviewed and discussed pre-existing methods that are widely used in climate science and physics and proposed some sophistication of these techniques. First, by introducing a novel estimator of the extremal index which contains more information on the dynamics, revealing, for example, periodic phenomena. We then provided a deeper understanding of the nature of the computed quantities, whose wide distribution originates from the multifractal structure of the attractor. This relation provides a new method to compute the singularity spectrum of a system from the distribution of local dimensions (that are computable by EVT techniques). The influence of noise of diverse types on the local dimensions and on the extremal index was investigated, both from an analytical and numerical point of view, and a numerical study of the statistics of the number of visits in the neighborhood of a particular point, or in other target sets of interest was performed. These results could find application in future works in climate science or other area in physics.

On a more theoretical level, we explored the link between the asymptotic statistics of the closest encounter of multiple trajectories in a highly chaotic system and the generalized dimensions of its measure. The asymptotic distribution for q orbits turns out to be an extreme value law whose parameters are given by the length of the trajectory, the generalized dimension of order q of a system and an extremal index that was called the dynamical extremal index of order q. This latter quantity allows to define naturally a novel spectrum of quantities, easily computable from (extensive) data, that is an indicator of the hyperbolicity of the system. This interesting connection could motivate further investigation.

We then discussed the role of the generalized dimensions on the distribution of diverse local quantities. This influence is apparent in the rate function governing the large deviations of these quantities. Such results were known for local dimensions, return times and hitting times in cylinders. We here extended these results to hitting times in a spherical neighborhood of a point. This has potential for physical applications, since knowing the generalized dimensions of the system could, thanks to this relation, help to predict the occurence of some events.

Finally, we extended the scope of the dynamical theory of extreme values to cover phenomena that do not happen instantaneously, but evolve over a finite, albeit unknown at the onset, time interval. We considered complex dynamical systems, composed of many individual subsystems linked by a network of interactions. As a specific example of the general theory, a model of neuronal network, introduced to describe the electrical activity of the cerebral cortex, is analyzed was in detail: on the basis of this analysis, we proposed a novel definition of neuronal cascade, a physiological phenomenon of primary importance. We showed the emergence of extreme value laws for the statistics of these cascades, both from the point of view of exceedances (that satisfy critical scaling theory) and of block maxima. The observables considered here are diverse metrics associated to these cascades: size, length and depth. They characterize complex null sets likely to have a fractal structure, which would then constitute a first physical application of the recently developed dynamical theory of extreme events in fractal landscape.

References

[1] H. Poincaré, Sur le problème des trois corps et les équations de la dynamique, Acta mathematica, 13, 1890, 1-270

[2] E.N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. 20(2) (1963) 130–141.

[3] E. Fermi, P. Pasta, S. Ulam, M. Tsingou, Studies of the nonlinear problems (1955)

[4] G.D. Birkhoff, Proof of the ergodic theorem, Proc. Natl. Acad. Sci. USA, 17 (12) (1931) 656660

[5] A. Boyarsky, P. Gora, Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension, Birkhäuser, Basel, 1997.

[6] E. Ott, Chaos in dynamical systems, Chapter 9, Cambridge University press (1993).

[7] V. Araújo, M.J. Pacifico, Three-dimensional flows, Springer, (2010).

[8] Y. Sinai, Introduction to ergodic theory, Princeton Univ. Press, Princeton, New Jersey, 1977, 144 pp.

[9] A. Katok, B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems, Encyclopedia of Mathematics and Its Applications 54, Cambridge University Press, 1995.

[10] A.C.M. Freitas, J.M. Freitas and M. Todd, Hitting time statistics and extreme value theory, M. Probab. Theory Relat. Fields (2010) 147: 675. https://doi.org/10.1007/s00440-009-0221-y

[11] A.C.M. Freitas, J.M. Freitas, F.B. Rodrigues, J. Valentim Soares, Rare events for Cantor target sets, arXiv:1903.07200

[12] P. Grassberger, Generalized dimension of strange attractors, Phys. Lett. A 97 (1983) 227–230.

[13] J. Ford, The Fermi-Pasta-Ulam problem: Paradox turns discovery, Phys. Rep., Volume 213, Issue 5, 1992, Pages 271-310

[14] B. Mandelbrot, Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence In Statistical Models and Turbulence, 331-351, Murray Rosenblatt and Charles Van Atta New York Springer.

[15] U Frisch and G Parisi, Fully developed turbulence and intermittency, Proc. Int. Summer school on Turbulence and Predictability in Geophysical Fluid Dy namics and Climate Dynamics, 84–88, Amsterdam 1985.

[16] J. Levy Vehel and P. Mignot, Multifractal Segmentation of Images, Fractals (1994)

[17] P. Grassberger, I. Procaccia, Characterization of strange sets, Phys. Rev. Lett. 50 (1983) 346–349.

[18] T.C. Halsey, M. Jensen, L. Kadanoff, I. Procaccia, B. Shraiman, Fractal measures and their singularities: The characterization of strange sets, Phys. Rev. A 33 (1986) 1141–1151.

[19] U. Frisch, G. Parisi, Turbulence and predictability of geophysical fluid dynamics, M. Ghil, R. Benzi and G. Parisi eds., p. 84, North Holland, 1985.

[20] R. Benzi, G. Paladin, G. Parisi, A. Vulpiani, On the multifractal nature of fully developed turbulence and chaotic systems, J. Phys. A: Math. Gen. 17 (1984) 3521.

[21] D. Ruelle, Thermodynamic Formalism: The Mathematical Structures of Classical Equilibrium Statistical Mechanics, Addison-Wesley, Reading, 1978

[22] N. Haydn, S. Vaienti, Limiting entry times distribution for arbitrary sets, preprint

[23] Y. Pesin, Dimension theory in dynamical systems, The University of Chicago Press, Chicago, 1997.

[24] 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.

[25] R.H Riedi, Introduction to Multifractals, 1999

[26] H. Kantz, T. Schreiber, Nonlinear time series analysis, Cambridge University Press, Cambridge, 1997.

[27] R. Badii, G. Broggi, Measurement of the dimension spectrum $f(\alpha)$: Fixed-mass approach, Phys. Lett. A. 131 (1988) 339–343.

[28] R. Badii, A. Politi, Statistical description of chaotic attractors – the dimension function, J. Stat. Phys. 40 (1985) 725–750.

[29] R. Pastor-Satorras, R.H. Riedi, Numerical estimates of the generalized dimensions of the Hénon attractor for negative q, J. Phys. A: Math. Gen. 29 (1996) L391–L398.

[30] G. Mantica, L. Perotti, Extreme value laws for fractal intensity functions in dynamical systems: Minkowski analysis

[31] M.R. Leadbetter, On extreme values in stationary sequences. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 28, 289303, (1973/74).

[32] M.R. Chernick, T. Hsing, W.P. McCormick, Calculating the extremal index for a class of stationary sequences. Adv. in Appl. Probab., 23 (4), 835850, (1981).

[33] A.C.M. Freitas and J.M. Freitas, On the link between dependence and independence in extreme value theory for dynamical systems. Statist. Probab. Lett., 78 (9), 10881093, (2008) doi:10.1016/j.spl.2007.11.002. URL http://dx.doi.org/10.1016/j. spl.2007.11.002.

[34] A.C.M. Freitas, J.M. Freitas, M. Todd, Speed of convergence for laws of rare events and escape rates. Stochastic Process. Appl., 125 (4), 16531687, (2015) doi:10.1016/j.spa.2014.11.011. URL http://dx.doi.org/10.1016/j. spa.2014.11.011.

[35] H. Crámer, On a new limit theorem of the theory of probability, Uspekhi Matematicheskikh Nauk, 166–178 (1944).

[36] Asymptotic probability and differential equations, Comm. Pure Appl. Math., 261–286, (1966).

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

[38] G. Keller and C. Liverani, Rare events, escape rates and quasistationarity: some exact formulae, J. Stat. Phys. 135 (2009), 519-534.

[39] R.H. Riedi, An Improved Multifractal Formalism and Self-Similar Measures, J. Math. Anal. Appl. 189 (1995) 462–490.

[40] M. Alber, J. Peinke, Improved multifractal box-counting algorithm, virtual phase transitions, and negative dimensions, Phys. Rev. E 57 (1998) 5489–5493.

[41] M. Kac, Probability and related topics in physical sciences, Wiley, New York, 1959.

[42] D. Ornstein, B. Weiss, Entropy and data compression, IEEE Trans. Inf. Th. 39 (1993) 78–83.

[43] M.H. Jensen, L.P. Kadanoff, A. Libchaber, I. Procaccia, J. Stavans, Global universality at the onset of chaos: results of a forced Rayleigh–Bénard experiment, Phys. Rev. Lett. 55 (1985) 2798–2801.

[44] T.C. Halsey, M.H. Jensen, Hurricanes and Butterflies, Nature 428 (2004) 127.

[45] N. Haydn, J. Luevano, G. Mantica, S. Vaienti, Multifractal properties of return time statistics, Phys. Rev. Lett. 88 (2003) 224502.

[46] G. Mantica, The Global Statistics of Return Times: Return Time Dimensions Versus Generalized Measure Dimensions, J. Stat. Phys. 138 (2010) 701–727.

[47] J.R. Chazottes, E. Ugalde, Entropy estimation and fluctuations of Hitting and Recurrence Times for Gibbsian sources, Disc. Cont. Dyn. Sys. B 5(3) (2005) 565–586.

[48] L. Barreira, Dimension and recurrence in hyperbolic dynamics, Progress in Mathematics 272, Birkäuser Verlag, Basel, 2008.

[49] M.F. Barnsley, S.G. Demko, Iterated function systems and the global construction of fractals, Proc. R. Soc. London A 399 (1985) 243–275.

[50] D. Faranda, S. Vaienti, Correlation dimension and phase space contraction via extreme value theory, Chaos 28 (2018) 041103.

[51] T. Caby, D. Faranda, G. Mantica, S. Vaienti, P. Yiou, Generalized dimensions, large deviations and the distribution of rare events

[52] H. Touchette, A basic introduction to large deviations: Theory, applications, simulations, Modern Computational Science 11: Lecture Notes from the 3rd International Oldenburg Summer School, BIS-Verlag der Carl von Ossietzky Universitat Oldenburg, 2011

[53] K. T. Alligood, T. D. Sauer, J. A. Yorke, Chaos: An Introduction to Dynamical Systems, Springer, 1996

[54] D. Faranda, V. Lucarini, P. Manneville and J. Wouters, On using extreme values to detect global stability thresholds in multi-stable systems: the case of transitional plane Couette flow. Chaos Solitons and Fractals, 64. 26 - 35, (2014)

[55] V. Lucarini, D. Faranda, J. Wouters and T. Kuna, Towards a general theory of extremes for observables of chaotic dynamical systems. Journal of Statistical Physics, 154 (3). pp. 723-750, (2014)

[56] V. Lucarini, R. Blender, C. Herbert, F. Ragone, S. Pascale, and J. Wouters, Mathematical and physical ideas for climate science. Reviews of Geophysics, 52 (4). pp. 809-859, (2014)

[57] T. Caby, G. Mantica, Extreme value theory of evolving phenomena in complex dynamical systems:firing cascades in a model of neural network, submitted to Physica E (2019).

[58] L. Arnold, Random Dynamics, Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[59] H. Aytac, J.M. Freitas, S. Vaienti, Laws of rare events for deterministic and random dynamical systems, Trans. Amer Math. Soc., 367, no. 11, 8229-8278, 2015.

[60] R.C. Bradley, Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions, Probab. Surveys, Volume 2, (2005), 107-144.

[61] J.P. Conze, A. Raugi, Limit theorems for sequential expanding dynamical systems on $[0, 1]$, Ergod. Theory Relat. Fields. Contemp. Math. 430, (2007), 89-121.

[62] N. Moloney, D. Faranda, Y. Sato, *An overview of the Extremal Index*, Submitted to Chaos, 2018

[63] D. Faranda, J.-M. Freitas, V. Lucarini, G. Turchetti, S. Vaienti, Extreme Value Statistics for Dynamical Systems with Noise, Nonlinearity, 26, 2597-2622, (2013).

[64] C. Gonzalez-Tokman, Multiplicative ergodic theorems for transfer operators: towards the identification and analysis of coherent structures in non-autonomous

dynamical systems, Contributions of Mexican mathematicians abroad in pure and applied mathematics. Guanajuato, Mexico: American Mathematical Society. Contemporary Mathematics, 709, 31-52 (2018).

[65] A. C. M. Freitas, J. M. Freitas and M. Todd. The extremal index, hitting time statistics and periodicity. Adv. Math. 231 (5) (2012) 26262665.

[66] A.C.M. Freitas, J.M. Freitas, S. Vaienti, Extreme value laws for nonstationary processes generated by sequential and random dynamical systems, Annales de l'Institut Henri Poincaré, 53, 1341–1370, (2017).

[67] M. Abadi, A.C.M. Freitas, J.M. Freitas, Dynamical counterexamples regarding the Extremal Index and the mean of the limiting cluster size distribution, arxiv 1808.02970

[68] N. Haydn, S. Vaienti,The compound Poisson distribution and return times in dynamical systems. Probab. Theory Related Fields, 144 (3-4), 517542, (2009).

[69] J. Beirlant, Y. Goegebeur, J. Segers, J. L. Teugels, D. De Waal (Contributions by), C. Ferro (Contributions by), *Statistics of Extremes: Theory and Applications*.

[70] A.C.M. Freitas, J.M. Freitas, M.Magalhaes, S. Vaienti, Point processes of non stationary sequences generated by sequential and random dynamical systems, in preparation.

[71] N. Haydn, M. Nicol, A. Torok, S. Vaienti, Almost sure invariance principle for sequential and non-stationary dynamical systems, Trans. Amer Math. Soc., 36, (2017), Pages 5293-5316.

[72] D. Faranda, Y. Sato, G. Messori, N. R. Moloney,P. Yiou, Minimal dynamical systems model of the northern hemisphere jet stream via embedding of climate data, Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2018-80

[73] G.L. O'Brien, Extreme values for stationary and Markov sequences, Annals of Probability, 15, 281291, (1987).

[74] V. Barros, L. Liao and J. Rousseau, On the shortest distance between orbits and the longest common substring problem, Adv. Math., 334(2019) , 311-339.

[75] A. Pikovsky, A. Politi, Lyapunov Exponents: A Tool to Explore Complex Dynamics, Cambridge University Press, Cambridge, 2016.

[76] L.S. Young, Dimension, entropy and Lyapunov exponents, Ergodic Theory and Dynamical Systems 2(1) (1986) 109–124.

[77] B. Saussol, S. Troubetzkoy, S. Vaienti, Recurrence, dimensions and Lyapunov exponents, J. Stat. Phys. 106 (2002) 623–634.

[78] B. Saussol, S. Troubetzkoy, S. Vaienti, Recurrence and Lyapunov exponents, Moscow Math. Journ. 3 (2003) 189–203.

[79] J. Beirlant et al., Statistics of Extremes, theory and Applications, John Wiley and Sons Ltd.

[80] A. Coutinho, J. Rousseau, B. Saussol, Large deviation for return times, Nonlinearity 31(11) (2018) 51625179.

[81] D.J. Feng, J. Wu, The Hausdorff dimension sets in symbolic spaces, Nonlinearity 14 (2001) 81–85.

[82] B. Saussol, J. Wu, Recurrence spectrum in smooth dynamical systems, Nonlinearity 16 (2003) 1991–2001.

[83] L. Olsen, First return times: multifractal spectra and divergence points, Disc. Cont. Dyn. Sys. 10 (2004) 635–656.

[84] S. Galatolo, Dynamical Systems Dimension and waiting time in rapidly mixing systems, Math. Res. Lett. 14(5) (2007) 797805.

[85] D. Faranda, A.C. Moreira Freitas, J. Milhazes Freitas, M. Holland, T. Kuna, V. Lucarini, M. Nicol, M. Todd, S. Vaienti, Extremes and Recurrence in Dynamical Systems, Wiley, New York, 2016.

[86] M. Carvalho, A.C. Moreira Freitas, J. Milhazes Freitas, M. Holland, M. Nicol, *Extremal dichotomy for uniformly hyperbolic systems.* Dyn. Syst. 30(4) (2015)

[87] M. Viana, Lectures on Lyapunov exponents, vol.145, Cambridge University Press, 2014.

[88] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Springer Verlag, Berlin, 2010.

[89] H. Hennion, L. Hervé, Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness, Lecture Notes in Mathematics 1766, Springer, Berlin, 2001.

[90] D. Faranda, S. Vaienti, Extreme value laws for dynamical systems under observational noise, Phys. D, 2014.

[91] R. Artuso, P. Cvitanović, B. G. Kenny, Phase transitions on strange irrational sets, Phys. Rev. A 39 (1989) 268–281.

[92] G. Paladin, S. Vaienti, J. Stat. Phys. 57, 289 (1989).

[93] V.I. Arnold, A. Avez, Ergodic Problems of Classical Mechanics, The Mathematical Physics Monograph Series, W. A. Benjamin Inc., New York/Amsterdam, 1968.

[94] M. Hénon, A two-dimensional mapping with a strange attractor, Comm. Math. Phys. 50(1) (1976) 69–77.

[95] A. Arneodo, G. Grasseau, E.J. Kostelich, Fractal dimensions and f(α) spectrum of the Hénon attractor, Phys. Lett. A 124 (1987) 426–432.

[96] R. Domínguez-Tenreiro, L. J. Roy, V. J. Martnez, On the Multifractal Character of the Lorenz Attractor, Progress of Theoretical Physics, Volume 87, Issue 5, 1 May 1992, Pages 11071118.

[97] S. Coles, An Introduction to Statistical Modeling of Extreme Values, Springer Series in Statistics, Springer-Verlag, London, 2001.

[98] Mathworks, Generalized extreme value parameter estimates. https://fr.mathworks.com/help/stats/gevfit.html

[99] D. Faranda, H. Ghoudi, P. Guiraud, S. Vaienti, Extreme Value Theory for synchronization of Coupled Map Lattices, Nonlinearity 31 (2018) 3326–3358.

[100] M. Süveges, Likelihood estimation of the extremal index, Extremes 10 (2017) 41–55.

[101] P.C. Hemmer, The exact invariant density for a cusp-shaped return map, J. Phys. A: Math. Gen. 17 (1984) L247–9.

[102] R. Kistler et al., The NCEPNCAR 50-year reanalysis: monthly means CD-ROM and documentation, Bull. Amer. Meteorol. Soc. 82(2) (2001) 247–268.

[103] R. Lozi, Un attracteur étrange du type attracteur de Hénon, J. Physique, 39 (C5), 9-10 (1978).

[104] B.J. Hoskins, I.N. James, Fluid dynamics of the mid-latitude atmosphere, John Wiley and Sons, United States, 2014.

[105] D. Faranda, G. Messori, P. Yiou, Dynamical proxies of North Atlantic predictability and extremes, Sci. rep. 7 (2017) 41278.

[106] R. Caballero, D. Faranda, G. Messori, A dynamical systems approach to studying midlatitude weather extremes, Geophys. Res. Lett. 44(7) (2017) 3346–3354.

[107] M. Hénon, A two-dimensional mapping with a strange attractor, Communications in Mathematical Physics. 50 (1): 6977 (1976).

[108] J. Gärtner, On large deviations from the invariant measure. Th. Prob. Appl., 22:24-39 (1977).

[109] S. Galatolo, Dimension via waiting time and recurrence, Math. Res. Lett. (2005).

[110] S. Galatolo, Dimension and hitting time in rapidly mixing systems Math. Res. Lett. 14 (2007), no. 5, 797805.

[111] S. Galatolo, J Rousseau , B Saussol, Skew products, quantitative recurrence, shrinking targets and decay of correlations Erg. Th. Dyn. Sys 35 (2015), no. 6, 18141845.

[112] T. Caby, D. Faranda, S. Vaienti, P. Yiou, *On the computation of the extremal index for time series*, arXiv:1904.04936

[113] B.V. Gnedenko, *Sur la distribution limite du terme maximum d'une serie aleatoire, Annals of Mathematics* **44**(3), 423–453 (1943).

[114] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer, London (2001).

[115] M. Thomas, M. Lemaitre, M. L. Wilson, C. Viboud, Y. Yordanov, H. Wackernagel and F. Carrat, *Applications of Extreme Value Theory in Public Health, PLoS One* **11**(7), e0159312 (2016).

[116] A.L. Shilnikov, N.F. Rulkov, *Origin of chaos in a two-dimensional map modelling spiking- bursting neural activity*, *Int. J. Bifurcat. Chaos* **13**, 3325–3340 (2003).

[117] N. F. Rulkov, *Modeling of spiking-bursting neural behavior using two-dimensional map*, *Phys. Rev. E* **65**, 0419222002 (2002).

[118] J. Ford, *Directions in classical chaos*, in *Directions in Chaos*, H. Bai-lin Ed., World Scientific, Singapore (1987).

[119] K. Kanders, R. Stoop, *Phase Response Properties of Rulkov Model Neurons*, in *Emergent Complexity from Nonlinearity, in Physics, Engineering and the Life Sciences*, G. Mantica, R. Stoop, S. Stramaglia, Eds., Springer Proceedings in Physics **191** (Springer International Publishing AG) (2017).

[120] K. Kanders, T. Lorimer, R. Stoop, *Avalanche and edge-of-chaos criticality do not necessarily co-occur in neural networks, Chaos* **27**, 047408 (2017).

[121] J. P. Sethna, K. A. Dahmen, and C. R. Myers, *Crackling noise, Nature* **410**, 242–250 (2001).

[122] J.J. Hopfield and A.V. Herz, *Rapid local synchronization of action potentials: toward computation with coupled integrate-and-fire neurons, Proc. Natl. Acad. Sci. U.S.A.* **92**, 6655 (1995).

[123] A.V. Herz and J.J. Hopfield, *Earthquake cycles and neural reverberations: collective oscillations in systems with pulse-coupled threshold elements, Phys. Rev. Lett.* **75**, 1222–1225 (1995).

[124] A. V. Rangan, L. S. Young, *Emergent dynamics in a model of visual cortex*, *J. Comput. Neurosci.* **35**, 155–167 (2013).

[125] A. V. Rangan, L. S. Young, *Dynamics of spiking neurons: between homogeneity and synchrony*, *J. Comput. Neurosci.* **34**, 433–460 (2013).

[126] L. Chariker, L. S. Young, *Emergent spike patterns in neuronal populations*, *J. Comput. Neurosci.* **38**, 203–220 (2015).

[127] A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences*, Cambridge University Press, Vol. 12 (2003).

[128] D. Plenz *et al.*, *Multi-electrode array recordings of neuronal avalanches in organotypic cultures*, *Journal of Visualized Experiments*, **1**:pii, 2949 (2011).

[129] A. L. Hodgkin, A. F. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve, J Physiol.* **117**, 500–544 (1952).

[130] W. L. Shew and D. Plenz, *The Functional Benefits of Criticality in the Cortex, The Neuroscientist* **19**(1), 88–100 (2013).

[131] L. de Arcangelis and H.J. Herrmann, *Learning as a phenomenon occurring in a critical state, Proc. Natl. Acad. Sci. U.S.A.* **107**, 3977–3981 (2010).

[132] D. Azevedo, A. C. Moreira Freitas, J. Milhazes Freitas and F. B. Rodrigues *Clustering of extreme events created by multiple correlated maxima, Phys. D* **315**, 33–48 (2016).

[133] J. Beggs and D. Plenz, *Neuronal avalanches in neocortical circuits*, *J. Neurosci.* **23**, 11167–11177, (2003).

[134] Z. Lu, S. Squires, E. Ott and M. Girvan, *Inhibitory neurons promote robust critical firing dynamics in networks of integrate-and-fire neurons, Phys. Rev. E* **94**, 062309 (2016).

[135] A. Mazzoni, F. D. Broccard, E. Garcia-Perez, P. Bonifazi, M. E. Ruaro and V. Torre, *On the dynamics of the spontaneous activity in neuronal networks*, *PLoS ONE* **2**, e439 (2007).

[136] N. Friedman, S. Ito, B. A. W. Brinkman, M. Shimono, R. E. L. DeVille, K. A. Dahmen, J. M. Beggs and T. C. Butler, *Universal critical dynamics in high resolution neuronal avalanche data*, *Phys. Rev. Lett.* **108**, 208102 (2012).

[137] A. Politi, R. Livi, G.L. Oppo and R. Kapral, *Unpredictable behaviour in stable systems, Europhys. Lett.* **22**, 571 (1993).

[138] R. Zillmer, R. Livi, A. Politi and A. Torcini, *Desynchronization in diluted neural networks, Phys. Rev. E* **74**, 036203 (2006).

[139] G Mantica, L. Perotti, *Extreme value laws for fractal intensity functions in dynamical systems: Minkowski analysis*, *J. Phys. A: Math. Theor.* **49**, 374001 (2016).

[140] J. Beggs and D. Plenz, *Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures*, *J. Neurosci.* **24**, 5216–5229 (2004).

[141] W. L. Shew, H. Yang, T. Petermann, R. Roy and D. Plenz, *Neuronal avalanches imply maximum dynamic range in cortical networks at criticality*, *J. Neurosci.* **29**, 15595–15600 (2009).

[142] D. B. Larremore, M. Y. Carpenter, E. Ott and J. G. Restrepo, *Statistical properties of avalanches in networks, PRE* **85**, 066131 (2012).

[143] C. W. Eurich, J. M. Herrmann and U. A. Ernst, *Finite-size effects of avalanche dynamics, Phys. Rev. E* **66**, 066137 (2002).

[144] A. Levina, J. M. Herrmann and T. Geisel *Dynamical synapses causing self-organized criticality in neural networks, Nat. Phys.* **3**, 857 (2007).

[145] C. Haldeman, J. M. Beggs, *Critical branching captures activity in living neural networks and maximizes the number of metastable states, Phys. Rev. Lett.* **94** , 058101 (2005).

[146] L. de Arcangelis, C. Perrone-Capano and H. J. Herrmann, *Self-organized criticality model for brain plasticity, Phys Rev. Lett.* **96**, 028107 (2006).