Extreme Value Theory of geophysical flows

Extremwerttheorie geophysikalischer Strömungen

Dissertation

Zur Erlangung des Doktorgrades der Naturwissenschaften im Department Geowissenschaften der Universität Hamburg

vorgelegt von

Davide Faranda

aus

Meteorologisches Institut - Universität Hamburg

Hamburg 01/11/2012

Als Dissertation angenommen vom Department Geowissenschaften der Universität Hamburg

Auf Grund der Gutachten von Prof. Dr. V. Lucarini

Hamburg, den .....
(Datum der vorläufigen Bescheinigung)

Prof. Dr. Jürgen Oßenbrügge
Leiter des Department Geowissenschaften

Introduction

> Le savant n'étudie pas la nature parce que cela est utile; il l'étudie parce qu'il y prend plaisir et il y prend plaisir parce qu'elle est belle. Si la nature n'était pas belle, elle ne vaudrait pas la peine d'être connue, la vie ne vaudrait pas la peine d'être vécue.
>
> Henri Poincaré, Science et méthode, 1908

In the past century, within the framework of the dynamical systems theory, a great scientific effort has been directed in introducing a great variety of tools useful to characterize the dynamics. They have been named *dynamical indicators of stability* due to the fundamental role that stability plays in determine chaotic or regular behaviours.
In the meanwhile, the fast growth of computer facilities has allowed for directly simulating complex natural phenomena. Nowadays, engineering software allows for saving money and time in the design of industrial objects, financial models help in taking real time decisions in the stock market, weather forecasts - practically impossible only fifty years ago - are used in the everyday life to plan agriculture and social activities. In most of the applications relying on the dynamical systems theory, the main focus has been reserved to an accurate simulation of the trajectories of the system in order to make them look coherent with the observed phenomena. Instead, less effort has usually been reserved in detecting and understanding the dynamical properties by looking at the stability of the solutions and at the geometry of the trajectories in the phase space. Moreover, most of the dynamical indicators introduced so far lack of versatility and cannot be computed by using the

trajectory of the dynamical systems but require accessory tools which can be practically inaccessible for high dimensional dynamical systems. In many cases, these indicators provide information only about the behaviour of the systems around their mean state whereas they are usually blind in detecting properties of events which happen with a very small probability: the so called extreme events which can be instead very relevant for characterizing several phenomena that may have a huge social or economical impact.

Nowadays, it is becoming clearer and clearer that, to advance in the understanding of complex natural phenomena, a deeper understanding of the dynamical properties of complex systems is needed. The results contained in this dissertation try to give a methodological and practical way for understanding relevant properties of dynamical systems by applying and devising new dynamical indicators. These indicators mainly rely on the results obtained for the so called Extreme Value Theory that can be used to deduce general and local properties of the physical measures associated to the trajectories in the phase space. The work of thesis is here presented as a collection of the articles [1, 2, 3, 4, 5, 6, 7, 8] which are reported in Chapters 2-9.

A first straightforward dynamical indicator - the Reversibility Error - is introduced in [1]. It relies on the ability of discerning between chaotic and regular behaviour of a system for which the dynamic is time-reversible. Although time reversibility is clearly a non-general properties of dynamical systems, the results shows how the numerical error grow in a different way for regular and chaotic dynamics giving an insight on the reliability of numerical simulations in several situations in which stability properties play a prominent role. The results described in [2, 3, 4, 5, 6, 7, 8], rely instead on the indicators built up for studying extreme events.
The Extreme Value Theory, originally introduced for series of independent variables, has been extended to dynamical systems in a nice theoretical framework which allows to infer stability and geometrical properties of the or-

bits. In the last decade a renewed interest in the Extreme Value Theory has emerged because of new mathematical tools recently developed to connect the classical theory of extremes of independent and identically distributed variables (i.i.d.) to the output of dynamical systems. The traditional Extreme Value theory states a sort of central limit theorem for extreme values: under rather general assumptions about the tail decay of the data parent distribution, the maxima (minima) extracted from a series of i.i.d. variables converge to the so-called Generalized Extreme Value (GEV) distribution [9]. The selection of maxima (minima) is made through the so-called Block Maximum approach procedure, which consists in taking partial maxima in subsets of the original series obtained by dividing it into bins of equal length. A new mathematical theory has been developed to show that extreme values related to the output of a dynamical systems may have as limiting distribution a GEV if it fulfils chaotic constraints [10]. The series of i.i.d. variables is here replaced by the time series of some observable defined on the attractor. In a totally new fashion, the theory then allows to connect the properties of extremes to geometrical properties, e.g. the local dimension of the attractor, and dynamical properties of the system, such as the hitting time statistics which connects the return times of a trajectory in a subset of the phase space with the chaotic properties of the system [10].

The mathematical theory developed so far ensures that such properties exist when the asymptotic behaviour is considered: we expect to observe convergence when both the number of maxima and the length of the bin in which the maxima are taken, approach infinity. Naturally, when moving from theory to practice, the first question to answer is whether this asymptotic behaviour can be observed even at finite time. The need of building a bridge between the abstract statements of the theory and the world of practical applications has triggered my research activity on Extreme Value Theory in dynamical systems during my PhD program in the SICSS graduate school at Klimacampus - Institute of Meteorology, University in Hamburg[1]. First of all we have

[1]The Phd program started in October 2010 at Reading University - Department of

explored the existence of a limiting extreme value distribution at finite times for simple low-dimensional discrete maps [2]. We compared the analytical results computed using the theory with numerical experiments in which we studied the convergence to the GEV limiting distribution using the block-maxima approach, pointing out in which cases we obtain a robust estimation of parameters. In particular we found that a good agreement between theoretical and experimental parameters is achieved when the number of maxima is order or 1000, and each maximum is taken over more than 1000 observations. Even though we have tested the algorithm on a restricted selection of low-dimensional chaotic maps, these properties seem quite robust and they can be interpreted as lowest numerical boundaries needed to obtain a reliable numerical inference of extremes behaviour in dynamical systems. These ranges have been first proposed in investigation performed on intermidiate complexity models featuring several degrees of freedom [11]. In regular maps, for which mixing properties do not hold, we showed that the fitting procedure to the classical Extreme Value Distribution fails as expected. However, the empirical distributions we have obtained can be explained starting from analytical results valid for finite time on quasi-periodic orbits [12]. The robustness of the algorithm just described has allowed us to extend the simulations to classical chaotic discrete maps that possess singular invariant measures [4]. In these systems, for selected observables, a GEV distribution has been observed as the best continuous approximation to the histograms obtained numerically and can be theoretically justified by linking the parameters of the fitted distribution to the local dimension exponent characterizing the attracting set. By connecting the results obtained for chaotic and regular dynamics, we have also used the Extreme Value Theory results as indicators of stability by studying systems that feature the coexistence of regular and chaotic motions: by excluding the regular orbits which do not show a GEV distribution for extremes, we have been able to depict the dynamical structure of the Standard map [3].

Mathematics and Statistics and have been transferred to Hamburg in September 2011.

Besides the extreme events already described, complex dynamical systems, ranging from ecosystems to financial markets and the climate, can have tipping points at which a sudden shift to a contrasting dynamical regime may occur [13]. In dynamical systems theory there is a clear distinction between what we have defined as an extreme event and the so called tipping points or critical transitions. A whole branch of dynamical system theory is dedicated to the study of these kind of phenomena and it is known as bifurcation theory. Unfortunately, understanding critical transitions associated with bi-stability in noisy high-dimensional systems remains a tough scientific challenge: such transitions are difficult to classify or even model with the theoretical or numerical tools that have been proven successful in low-dimensional systems [7]. In the past decades a renewed interest in this topic has emerged and many "indicators of criticality" have been developed to identify early warnings of abrupt transitions to different dynamical states: some of these indicators are based on the modifications of the auto-correlation properties of particular observables when the system is pushed towards a transition, others on the fact that an increase of the variance is observed when moving towards tipping points[13, 14]. Although the variance is a very straightforward quantity to compute, it does not provide all the information that we are able to extract by analysing the complete data distribution. In particular it does not reveal whether fluctuations increase towards maximum or minimum values, or whether the right and the left tails of the distribution are symmetrical or not. All these properties can be explored via Extreme Value analysis as explained in [8].

This dissertation is organized as follows: in the first chapter is presented a wide outline of the theoretical and numerical tools used in the papers enclosed: we introduce some elements of the extreme value theory for i.i.d. variables, the indicators of stability already introduced in the literature such as Recurrences, Hitting Time Statistics, Correlation function. We end the outline by stating the main theorem which connects the Hitting Time Statis-

tics to the emergence of extreme value laws in dynamical systems. In Chapter 2-9 the papers are attached, whereas in the last pages conclusions and suggestions for extending the results presented in this dissertation are drawn.

Contents

Introduction **i**

1 Outline of the research **1**

1.1 Traditional Extreme Value Theory 1

1.1.1 The Extreme Value laws 2

1.1.2 Obtaining the limiting distributions 4

1.1.3 Gnedenko's results 5

1.1.4 GEV inference criteria 8

1.1.5 The Kolmogorov-Smirnov test 11

1.1.6 The Peak over Threshold approach: 13

1.1.7 Maximum likelihood estimation for the GPD 14

1.2 Recurrences and Hitting time statistics 15

1.2.1 Poincaré Recurrence Theorem. 16

1.2.2 Return Time Statistics 19

1.2.3 Mixing Systems 20

1.2.4 Hitting Time Statistics 22

1.3 Reversibility Error and Correlations 26

1.3.1 Reversibility Error 26

1.3.2 Correlations function 27

1.3.3 Correlations and Recurrences 28

1.4 Hitting time statistics and Extreme Value Theory 31

1.4.1 Freitas-Freitas-Todd Theorem 33

1.5 Methodological notes and informatics tools 38

2 Article: Analysis of Round Off Errors with Reversibility Test as a Dynamical Indicator 41

3 Article: Numerical Convergence of the Block-Maxima Approach to the Generalized Extreme Value Distribution 57

4 Article: Generalized Extreme Value Distribution as dynamical indicators of stability 85

5 Article: Extreme value theory for singular measures 103

6 Article: Universal Behaviour of Extreme Value Statistics for Selected Observables of Dynamical systems 123

7 Article: Extreme Value Statistics for dynamical systems with noise 137

8 Article: Bistable systems with stochastic noise: virtues and limits of effective one-dimensional Langevin equations 169

9 Article: Extreme Value Theory provides early warnings for critical transitions 185

Conclusions 201

Bibliography 206

Chapter 1

Outline of the research

1.1 Traditional Extreme Value Theory

The study of extreme events is of great interest in different disciplines. It has been applied successfully to the study of extreme floods [15], amounts of large insurance losses [16], extreme earthquake, meteorological and climate events (see [17] for an extended review on extremes and climate).
The extreme value theory originates to find up methods to model and measure events which occur with very small probability. In this section we explain the classical theory, devised to study the stochastic behaviour of the extremes of independent and identical distributed (hereinafter i.i.d.) variables. Traditionally, two approaches have been used to study the Extreme Value Theory:

- The Block Maxima approach: the distribution of a series of maxima (minima) converges to the so-called Extreme Value Laws for extremes taken by dividing the data series into an asymptotically infinite number of bins each containing an asymptotically infinite number of observations.

- The Peak over Threshold approach: the distribution of excesses over a given threshold converge to the so called Generalized Pareto distribution in the limit of an infinitely high (or low) threshold.

1.1.1 The Extreme Value laws

Let X_1, X_2, X_n be i.i.d. variables with the same cumulative distribution function:

$$F(x) = \int_{-\infty}^{x} f(t)dt \tag{1.1}$$

being $f(t)$ the associated probability density function. Define $M_n = \max\{X_1, ..., X_n\}$. We are interested in studying if it exists an asymptotic distribution $G(x)$ such that:

$$G(x) = \Pr[M_n \leq x]$$

For example X_i can be a dataset of independent observations taken daily for n years. In the ith year we will extract the year maximum M_i, assuming that the block sizes are quite large and the maxima in different blocks are independent realisations [18]. For any finite value x , the maximum will exceed x as N increases, with $1 - [F(x)]^N$ tends to unit if $F(x) < 1$ or to zero if $F(x) = 1$. In principle we obtain a degenerate distribution for $[F(x)]^N$ as $N \to \infty$ unless we introduce some sequences of transformed and reduced values a_n and b_n so that $(a_n M_n + b_n)$ is renormalized for each n [19] . By denoting $G(x)$ this reduced distribution, it follows that $G(x)$ must satisfy:

$$[G(x)]^N = G(a_N x + b_N). \tag{1.2}$$

Gnedenko [20] showed that, under general assumptions on the nature of $F(x)$, the series of $M_{x,n}$ converges to one of the three Extreme Value Laws (EVLs):

- Gumbel cumulative distribution (type 1):

$$G_1(x) = \exp\left\{-e^{(x-\mu)/\sigma}\right\} \tag{1.3}$$

- Frechet cumulative distribution (type 2):

$$\begin{cases} G_2(x) = 0 & x < \mu \\ G_2(x) = \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\} & x \geq \mu \end{cases} \tag{1.4}$$

- Weibull cumulative distribution (type 3):

$$\begin{cases} G_3(x) = \exp\left\{-\left(-\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\} & x \leq \mu \\ G_3(x) = 1 & x > \mu \end{cases} \tag{1.5}$$

  where $\mu, \sigma > 0, \xi \in \mathbb{R}$ are parameters.

It is possible to represent the three types distributions using a single family of generalized distribution called Generalized Extreme Value distribution:

$$G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{1.6}$$

which holds for:

$$1 + \xi(x-\mu)/\sigma > 0 \tag{1.7}$$

and where:

- $\mu \in \mathbb{R}$ is the location parameter
- $\sigma > 0$ the scale parameter
- $\xi \in \mathbb{R}$ the shape parameter. also called the tail index, indicates the thickness of the tail of the distribution.

When $\xi \to 0$, the distribution $G(x)$ corresponds to a Gumbel type (Eq. 1.3) . When the index is negative, it corresponds to a Weibull (Eq. 1.5); when the index is positive, it corresponds to a Frechet (Eq. 1.4).
The leading idea in extreme value theory is analogous to the one used in central limit theorem: here, nonetheless, we are dealing with two different limits as we require at the same time the selection of authentic extremes and a sufficient amount of maxima to compute the statistics. This issue is widely described in the dynamical system framework in [2].

1.1.2 Obtaining the limiting distributions

We give an intuitive idea of how the EVLs can be obtained by direct computation, dedicating the next section to the theoretical statements.
Using condition 1.2 it is possible to obtain Gumbel distribution by taking $a_N = 1$. We report the proof in the case of type 1. Since $a_N = 1$ equation 1.2 becomes:

$$[G(x)]^N = g(x + b_N) \tag{1.8}$$

the latter equation must satisfy Eq. 1.2 so that:

$$[G(x)]^{NM} = [G(x + b_N)]^M = G(x + b_N + b_M) \tag{1.9}$$

$$[G(x)]^{NM} = G(x + b_{NM}) \tag{1.10}$$

Using Eqs. 1.9-1.10 we can see that:

$$b_{NM} = b_N + b_M \tag{1.11}$$

so that:

$$b_n = \sigma \log(N), \qquad \text{with } \sigma \text{ costant} \tag{1.12}$$

By taking logarithms of Eq. 1.9 twice we have:

$$\log N + \log\{-\log G(x)\} = \log\{-\log G(x + b_N)\} \tag{1.13}$$

plugging Eq. 1.12 in the latter we obtain:

$$\log N + \log\{-\log G(x)\} = \log\{-\log G(x + \sigma \log N)\} \tag{1.14}$$

Let us introduce:

$$h(x) = \log\{-\log G(x)\} \tag{1.15}$$

then:

$$h(x) = h(0) - \frac{x}{\sigma} \tag{1.16}$$

Now, since $h(x)$ decreases as x increases, σ is positive:

$$-\log G(x) = \exp\left[-\frac{x-\sigma h(0)}{\sigma}\right] \tag{1.17}$$

Eventually, by denoting $\mu = \sigma \log(-\log G(0))$:

$$-\log G(x) = \exp\left(-\frac{x-\mu}{\sigma}\right) \tag{1.18}$$

which demonstrates our statement since it is the logarithm of the type 1 distribution in Eq. 1.3.

Types 2 and 3 can be obtained by taking $a_N \neq 1$. In this case

$$x = a_N x = b_N \qquad \text{if} \quad x = b_N(1-a_N)^{-1} \tag{1.19}$$

See [21] for a complete proof.

1.1.3 Gnedenko's results

We state here the well known theoretical results originally framed by Gnedenko [20] under the assumptions described in Section 1.1.1. By rewriting $P\{a_n(M_n - b_n) \leq x\}$ as $P\{M_n \leq u_n\} = F^n(u_n) = \{1-(1-F(u_n))\}^n$ where $u_n = u_n(x) = x/a_n + b_n$ is a normalized sequence. Such type of normalized sequence converges to one of the three types of distribution described in Eqs 1.3-1.5 (where $\mu = 0$ and $\sigma = 1$ as they have already being renormalized by using a_n and b_n) if one of the following necessary and sufficient conditions established by [20] is satisfied:

Let us define the right endpoint x_F of a cumulative distribution function $F(x)$ as:

$$x_F = \sup\{x : F(x) < 1\} \tag{1.20}$$

Theorem (Gnedenko): *Necessary and sufficient conditions for the distribution function F of the random variables of the i.i.d. sequence $\{\xi_n\}$ to belong to each of the three types are:*

- *For the Frechet (type 2) distribution, $x_F = \infty$:*

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha}, \qquad \alpha > 0, \forall x > 0. \tag{1.21}$$

- *For the Weibull (type 3) distribution, $x_F < \infty$:*

$$\lim_{x\to 0^-} \frac{1-F(x_F - xh)}{1-F(x_F - h)} = x^{\alpha}, \qquad \alpha > 0, \forall x > 0 \tag{1.22}$$

- *For the Gumbel (type 1) distribution, there exists some positive function $G(t)$ such that:*

$$\lim_{t\to X_F} \frac{1-F(t+xG(t))}{1-F(t)} = e^{-x}, \qquad \forall x \in \mathbb{R} \tag{1.23}$$

It may in fact be shown that $\int_0^\infty (1-F(u))du < \infty$ when a type 1 limit holds, and one appropriate choice of g is given by

$$G(t) = \int_1^{x_F} \frac{1-F(u)}{1-F(t)} du, \quad t < x_F$$

Gnedenko [20] has shown that these conditions are necessary and sufficient to obtain the asymptotic EVLs. It is evident that, using these , different asymptotic EVLs may hold when maxima and minima are taken into account. Besides the theorem, the following corollary is useful to determine the normalizing sequences a_n and b_n to be used:

Corollary (Gnedenko): *The normalizing sequences a_n and b_n in the convergence of normalized maxima $P\{a_n(M_n - b_n) \le x\} \to G(x)$ may be taken as:*

- *Type 2:* $a_n = \gamma_n^{-1}, \quad b_n = 0;$

- *Type 3:* $a_n = (x_F - \gamma_n)^{-1}, \quad b_n = x_F;$
- *Type 1:* $a_n = [g(\gamma_n)]^{-1}, \quad b_n = \gamma_n;$

with $\gamma_n = F^{-1}(1 - 1/n) = \inf\{x; F(x) \geq 1 - 1/n\}$.

The complete proof of both the theorem and the corollary can be found in [9].
As we have already pointed out, the shape parameters has a very important role in discriminating the type of distribution observed. In the next paragraphs is given a brief overview of the different limiting EVLs

The Gumbel distribution (type 1)

From Eq. 1.6, we obtain a Gumbel distribution (equation 1.3) whenever $\xi \to 0$. The Gumbel probability density function has a shape skewed to the left and, therefore, the location parameter μ is equal to the mode but it differs from median and mean. As σ increases, the probability density function spreads out and becomes shallower. According to the condition stated in Eq. 1.23, the Gumbel distribution is the asymptotic EVLs if the parent distribution shows an exponential tail decay which includes the normal, log-normal, gamma or exponential types. Since the data of flood, earthquake occurrences and many other shows an asymptotic Gumbel distribution, it is often named in the scientific literature as the "extreme value distribution". This prominent role will be motivated in the dynamical systems set up by showing why the Type 1 distribution is more likely to be observed than the others [3, 4].

The Frechet distribution (type 2)

To obtain an asymptotic Frechet EVL (Eq. 1.4) the parent distribution must be a so called fat-tailed distribution which means that the tail decay is a power law (like it happens for Cauchy or T-Student distributions). It is usu-

ally observed for dataset which have a lower threshold such as precipitations, winds, incomes, whose series cannot assume negative values.

The Weibull distribution (type 3)

When the tail index is negative ($\xi < 0$) the GEV distribution is a type 3 of Weibull. In this case the parent distribution has a a finite upper endpoint $x_{up} = \mu + \sigma/(-\xi)$. It is also called a "reversed" Weibull distributions pointing out that it was first defined in a specular way with a finite lower endpoint. The Weibull distribution is widely used to fit minima distribution, temperature and sea-level data and whenever the observations used have bounded value in the direction of the tail considered.

In table 1.1.3 we summarize the limiting Extreme Value distribution (where $\mu \to 0$ and $\sigma \to 1$ which is obtained starting from a different standardized initial distribution.

Parent Dist.	**Lim. Dist.for Maxima**	**Lim. Dist.for Minima**
Exponential	Gumbel (Type 1)	Weibull (Type 3)
Gamma	Gumbel (Type 1)	Weibull (Type 3)
Normal	Gumbel (Type 1)	Gumbel (Type 1)
Log-normal	Gumbel (Type 1)	Gumbel (Type 1)
Uniform	Weibull (Type 3)	Weibull (Type 3)
Pareto	Frechet (Type 2)	Weibull (Type 3)
Cauchy	Frechet (Type 2)	Frechet (Type 2)

1.1.4 GEV inference criteria

We present here some methods used in order to infer the GEV distribution parameters. We focus our attention on the maximum likelihood criteria and the L-moments estimation since they have been applied to all the numerical investigations carried out in the enclosed papers.

As pointed out in Section 1.4.1 we can construct a sequence of maxima (minima) by subdividing the available data X into N bins of equal length L and by extracting the maximum from each bin i: $M_{N,i}$. The block length is a critical parameter to choice fairly between bias and variance in the parametric estimates.
The procedure relies on the maximization of the log likelihood function:

$$l(\mu, \sigma, \xi) = \prod_{i=1}^{L} \ln(G'(M_{N,i}; \mu, \sigma, \xi)) \tag{1.24}$$

where $G'(x; \mu, \sigma, \xi))$ is the derivative of $G(x; \mu, \sigma, \xi))$. Using Eq. 1.6, the log likelihood function can be rewritten as:

$$-m \ln(\sigma) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{L} \left\{ \ln\left[1 + \xi\left(\frac{M_{N,i} - \mu}{\sigma}\right)\right] - \left[1 + \xi\left(\frac{M_{N,i} - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \right\} \tag{1.25}$$

if $\xi \neq 0$, and as:

$$-m \ln(\sigma) - \sum_{i=1}^{L} \left\{ \left(\frac{M_{N,i} - \mu}{\sigma}\right) - \exp\left[-\left(\frac{M_{N,i} - \mu}{\sigma}\right)\right] \right\} \tag{1.26}$$

if $\xi = 0$. We can obtain a profile likelihood of μ, ξ or σ by setting the other two parameters to their maximum likelihood estimates $\tilde{\mu}, \tilde{\xi}, \tilde{\sigma}$ in Eqs. 1.25 or 1.26. For example, to compute the profile likelihood for the parameter ξ, we can construct the graph:

$$(x, y) = (\xi, l(\tilde{\mu}, \tilde{\sigma}, \xi)) \tag{1.27}$$

giving a section of the likelihood surface as viewed from the axis. The intersections of the horizontal line with the profile likelihood graph allows for estimatng a confidence interval:

$$y = \tilde{\xi} - 0.5 q_{0.95} \tag{1.28}$$

where $q_{0.95}$ is the 95% quantile of the χ^2 distribution with 1 degree of freedom. In our numerical analysis we have used the Matlab function *gevprobability density function* and *gevcdf* which return 95% confidence intervals for the parameter estimates.

Whenever the probability density function is not absolutely continuous the maximum likelihood estimation may fail as the minimization procedure may be not well defined causing unexpected divergences of the parameters. In all these situation is better to rely on a L-moments estimation , which, being based upon the computation of integrals rather than upon the solution of a variational problem, is more robust. The L-moments are analogous to ordinary moments, but are computed from linear combinations of the data values, arranged in increasing order. For a random variable X, the r th population L-moment is:

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \mathrm{E}X_{r-k:r}, \tag{1.29}$$

where $X_{k:n}¿$ denotes the k^{th} order statistic (k^{th} smallest value) in an independent sample of size n from the distribution of X and E denotes the expected value. In particular, the first four population L-moments are

$$\lambda_1 = \mathrm{E}X \tag{1.30}$$

$$\lambda_2 = (\mathrm{E}X_{2:2} - \mathrm{E}X_{1:2})/2 \tag{1.31}$$

$$\lambda_3 = (\mathrm{E}X_{3:3} - 2\mathrm{E}X_{2:3} + \mathrm{E}X_{1:3})/3 \tag{1.32}$$

$$\lambda_4 = (\mathrm{E}X_{4:4} - 3\mathrm{E}X_{3:4} + 3\mathrm{E}X_{2:4} - \mathrm{E}X_{1:4})/4. \tag{1.33}$$

The first two of these L-moments have conventional names:

$$\lambda_1 = \text{mean, L-mean or L-location}, \tag{1.34}$$

$$\lambda_2 = \text{L-scale}. \tag{1.35}$$

The L-scale is equal to half the mean difference.
Asymptotic approximations to sampling distributions are better for L-moments than for ordinary moments [22]. Whenever we have used this inference, we have computed the 95% confidence intervals of the parameters using a bootstrap procedure.

1.1.5 The Kolmogorov-Smirnov test

There exist many methods to check the goodness of fit to the GEV distribution. In our numerical investigation we have often used the Kolmogorov-Smirnov test [23] because it avoids the discretization of the null hypothesis unlike the χ^2 which instead requires to group the data and consider a weaker discretized null hypothesis.
We start with a sample of variable $X_1, ..., X_n$ belonging to some distribution $\mathbb{P}$ and we test the hypothesis that $\mathbb{P}$ is equal to a particular distribution $\mathbb{P}_0$ obtained, for example, by a fit of data. There are two possible hypotheses:

$$H_0 : \mathbb{P} = \mathbb{P}_0 \qquad H_1 : \mathbb{P} \neq \mathbb{P}_0 \tag{1.36}$$

Let us denote, as usual, by $F(x)$ the true underlying distribution function of data. While $F(x)$ is defined in Eq. 1.6, we can also define an *empirical* cdf by:

$$F_n(x) = \mathbb{P}_n(X \leq x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x) \tag{1.37}$$

that counts the proportion of the data X_i whose value is less than x. The Kolmogorov-Smirnov test [23] uses a law of large numbers result which implies that:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x) \to \mathbb{P}(X_1 \leq x) = F(x) \tag{1.38}$$

It can be shown that this approximation holds uniformly over all $x \in \mathbb{R}$:

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \to 0 \tag{1.39}$$

To use the Kolmogorov-Smirnov test, we will need another weaker result which we formulate using the central limit theorem:

$$\sqrt{n}(F_n(x) - F(x)) \to^{dist} N\left(0, F(x)(1 - F(x))\right) \tag{1.40}$$

where N is the Normal distribution, $\to^{dist}$ denotes the convergence in distribution and $F(X)(1 - F(x)$ is the variance of $I(X_1 \leq x)$. Using the previous observation we can state the following result:

$$\mathbb{P}(\sqrt(n) \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \leq t) \to H(t) = 1 - 2\sum_{i=1}^{\infty}(-1)^{i-1}e^{-2i^2t} \tag{1.41}$$

where $H(t)$ is the cdf of Kolmogorov-Smirnov distribution.
and by reformulating the hypotheses 1.36 in terms of $F(x)$:

$$H_0 : F = F_0 \qquad H_1 : F \neq F_0 \tag{1.42}$$

where F_0 is the cdf of $\mathbb{P}_0$.

Let us introduce the statistics:

$$D_n = \sqrt{n} \sup_{x\in\mathbb{R}} |F_n(x) - F_0(x)| \tag{1.43}$$

Using Eq. 1.36, if the null hypothesis is true, the D_n distribution can be tabulated as it will depend only on n and it is approximated by H if $n \to \infty$. If the null hypothesis fails, F_n will converge but it will not approximate F_0:

$$\sup_{x} |F_n(x) - F_0(x)| > \delta \tag{1.44}$$

with a small $\delta \in \mathbb{R}$. If we now multiply by $\sqrt{n}$:

$$D_n = \sqrt{n} \sup_{x\in\mathbb{R}} |F_n(x) - F_0(x)| > \sqrt{n}\delta \tag{1.45}$$

if H_0 fails then $D_n > \sqrt{n}\delta \to \infty$ as $n \to \infty$.
It is possible to use the following decision rule:

$$\delta = \begin{cases} H_0 : & D_n \leq c \\ H_1 : & D_n > c \end{cases} \tag{1.46}$$

c is a threshold tabulated under H_0. The level of significance α can be related to c by:

$$\alpha = \mathbb{P}(\delta \neq H_0 | H_0) = \mathbb{P}(D_n \geq c | H_0) \tag{1.47}$$

and if n is large the we can use the H distribution to compute c since:

$$\alpha = \mathbb{P}(D_n \geq c | H_0) \approx 1 - H(c) \tag{1.48}$$

1.1.6 The Peak over Threshold approach:

The Gnedenko-Pickands-Balkema-de Haan (hereinafter GPBH) theorem states that the distribution of large events conditioned to be larger than some threshold may be characterized by using the Generalized Pareto Distribution GPD.
The GPD can be derived from the GEV distribution in Eq. 1.6 by taking the $G(x, \xi)$ of the largest value given by:

$$H(x/\xi, s) = 1 + \ln(G(x/s, \xi) = 1 - (1 + \xi x/s)^{-1/\xi} \tag{1.49}$$

where $s = s(u)$ is a positive function.Nonetheless, the usual way to obtain the GPBH theorem is related to the definition of the excess distribution $F_u(x)$:

$$F_u(x) = P\{X - u < x | X > u\}, \qquad x \geq 0 \tag{1.50}$$

Let $F(x)$ be a distribution function with excess distribution $F_u(x), u > 0$. Then, for $-\infty < \xi < +\infty$, $F(x) \in D(x, \xi)$ maximum domain of attraction of $G(x, \xi)$ if and only if there exists a positive function $s(u)$ such that:

$$\lim_{u \to x_F} \sup_{0 \leq x \leq x_F - u} |\bar{F}_u(x) - \bar{H}(x/\xi, s(u))| = 0 \tag{1.51}$$

$\bar{F}(x)$ indicates the complementary cumulative of the distribution $F(x)$:

$$\bar{F}(x) = 1 - F(x) \tag{1.52}$$

This results is due to Pickands [24] and Balkema de Haan [25]. In the class of distributions for which this result holds may be further subdivided into three groups according to the value of the parameter ξ in the limiting GPD approximation to the excess distribution. For a complete derivation see, for example, Theorem 3.4.13 in [26].

The importance of the GPBH theorem resides in the statement which is not only valid for the largest value of a data set, as in the case of the Extreme Value Theory, by making a full use of all the data presents in the tail.
The Block Maxima approach uses only a part of the parent distribution tail discarding a significant part of the data. This is the reason why the Peak over Threshold approach is usually preferred in the applications as it requires generally smaller datasets [27].

1.1.7 Maximum likelihood estimation for the GPD

Let us denote by N_u the number of observation above a threshold u and as $y_1, ..., y_{N_u}$ the observations such that:

$$u : y_i = x_{j(i)} - u; x_{j(i)} > u \tag{1.53}$$

The GPBH theorem yields an approximation to the tail $\bar{F}(x)$ using the Pareto distribution as estimator:

$$\bar{F}(x + u) \simeq \bar{H}(x/\hat{\xi}, \hat{s}) \times (N_u/N) \tag{1.54}$$

As for the GEV distribution we can infer the estimates of the two parameters $\hat{\xi}$, $\hat{s}$ through the Maximum Likelihood Estimation (MLE). Here the log-likelihood l must equals:

$$l = -N_u \ln s - (1 + 1/\xi) \sum_{1}^{N_u} \ln(1 + \xi y_i/s) \tag{1.55}$$

The q-quantile estimator x_q, which denotes the value of the random variable not overpassed with probability q, can be written as:

$$x_q = u + (\hat{s}/\hat{\xi})(N/N_u(1-q)^{-\hat{\xi}} - 1) \tag{1.56}$$

While the scale parameter $s = s(u)$ depends on the threshold value u, the shape parameter ξ is, in theory, independent on it and uniquely determined by the distribution function $F(x)$ of the dataset. In other words, a very simple way to assess if the threshold value is high enough is to look at the fluctuations of the shape parameter: whenever no appreciable changes of ξ are observed when rising the value of the threshold, the asymptotic regime is reached [28].

1.2 Recurrences and Hitting time statistics

In this section we present the basic theory of Return Time Statistics (also known as Poincaré recurrences) showing the well known results obtained for a wide class of dynamical systems. In the past years both strong numerical evidences and theoretical results indicate that this theory is deeply connected with diffusion processes, correlation and not only for mixing dynamical systems but also for integrable maps [29], [30], [31]. For the understanding of the papers enclosed, these results are useful to establish a connection between the Hitting time statistics and the Extreme Value Theory that is widely used in [2, 3, 4, 5, 6].

1.2.1 Poincaré Recurrence Theorem.

Let us consider a standard probability space $(\Omega, \mathcal{F}, \nu)$:

1. *Ω is a sample space, the set of all possible outcomes.*
2. *$\mathcal{F}$ is a σ-algebra, the collection of all events which characterize groups of outcomes.*
3. *ν is a standard probability measure.*

and let $f : \Omega \to \Omega$ be a measure preserving map:

$$\nu(f^{-1}A) = \nu(A), \qquad \forall A \in \mathcal{F} \tag{1.57}$$

Then, for any $A \in \mathcal{F}$, the set of those points x of A such that $f^n(x) \notin A$ for all $n > 0$ has zero measure, that is almost every point returns infinitely often. i.e

$$\nu\{x \in A : f^n x \in A, \text{ for some } n > 0\} = \nu(A) \tag{1.58}$$

Proof: Let us introduce $A_n = \cup_{k=n}^{\infty} f^{-k}A$ Then, $A \subset A_0$ and $A_i \subset A_j$ when $j \leq i$. Also, $A_i = f^{j-i}A_j$, so that $\nu(A_i) = \nu(A_j)$ for all $i, j \geq 0$, by the f-invariance of ν. Now for any $n > 0$ we have $A - A_n \subset A_0 - A_n$, so that

$$\nu(A - A_n) \leq \nu(A_0 - A_n) = \nu(A_0) - \nu(A_n) = 0.$$

Hence $\nu(A - A_n) = 0$ for all $n > 0$, so that $\nu(A - \cap_{n=1}^{\infty} A_n) = \nu(\cup_{n=1}^{\infty} A - A_n) = 0$. But $A - \cap_{n=1}^{\infty} A_n$ is precisely the set of those $x \in A$ such that for some n and for all $k > n$ we have $f^k(x) \notin A$ [32]. A proof of theorem which follows the original Poincaré formulation can be found in [33]. □

First Visiting Time and Kac's Theorem

In the following we will assume that ν is *ergodic* for f:

$$\nexists A \in \mathcal{F}, 0 < \nu(A) < 1, \text{ such that } \nu(A \Delta f^{-1}A) = 0 \tag{1.59}$$

where $\triangle$ denotes the symmetric difference. This is equivalent to say that ν-almost every point in the whole space visits A; using the ergodicity we can introduce the concept of *visiting time* to a fixed set A of positive measure for ν-almost every point in Ω [34].
Let us fix $A \in \mathcal{F}$ with $\nu(A) > 0$, we define $\tau : \Omega \to \mathbb{N}$ as:

$$\tau = \inf\{n > 0 : f^n x \in A\} \tag{1.60}$$

first visiting time of A (note that $\tau(x) < \infty$ almost everywhere).
The smaller the set, the more time we have to wait in order to observe a return in A. This result is stated formally by the following theorem:

Kac's Theorem.

Let $(\Omega, \mathcal{F}, \nu, f)$ be an ergodic dynamical system and let $A \in \mathcal{F}$ be a set of positive measure, then

$$\mathbb{E}_A(\tau) = \frac{1}{\nu(A)} \int_A \tau(x) d\nu = \frac{1}{\nu(A)} \tag{1.61}$$

In accordance with this theorem one then looks at the return times which are normalized by the measure of the return set.

To prove Kac's Theorem we follow [34], [35] and [36].

First we introduce Rokhlin Tower definition:
In a dynamical system $(\Omega, \mathcal{F}, \nu, f)$ we define Rokhlin Tower a finite family of measurable sets.

$$\Psi = (B, fB, ..., f^{h-1}B) \tag{1.62}$$

where $f^j B$ are two-by-two disjoint sets. The set B represents the tower base, $f^j B$ is a floor and h is the height.

It is possible to partition the set A as follows. Let us define

$$A_i = \{x \in A : \tau(x) = i\} \qquad i \geq 1 \tag{1.63}$$

then:

$$\nu(A) = \nu\left(\bigcup_{i\geq 1} A_i\right). \qquad A = \bigcup_{i=1}^{\infty} A_i \tag{1.64}$$

and we consider the Rokhlin Tower of height k:

$$(A_k, fA_k, ..., f^{k-1}A_k) \tag{1.65}$$

Eventually, we introduce a lemma which is useful to prove Kac's Theorem:

Rokhlin's Lemma. *Let $(\Omega, \mathcal{F}, \nu, f)$ be an aperiodic dynamical system. Then $\forall \epsilon > 0, \forall n \in \mathbb{N}, \exists B$ measurable such that:*

- $f^i B, 0 \leq i \leq n$ are two-by-two disjoint
- $\nu(\bigcup_{i=0}^{n-1} f^i B) > 1 - \epsilon$

We can eventually prove Kac's theorem as follow. ν is ergodic and $\nu(A) > 0$ then:

$$\nu\left(\bigcup_{n\geq 0} f^n A\right) = 1 \tag{1.66}$$

We have seen that A is a disjoint union of A_k and then $\bigcup_{n\geq 0} f^n A$ is composed by the whole $f^j A_k$ in a separate way: on the one hand for the A_k definition, on the other, by the injectivity of T, we know that $fA_j, f^2A_j, ..., f^{j-1}A_j$ cannot join $fA_k, f^2A_k, ..., f^{k-1}A_j$ with $k \neq j$, then:

$$\sum_{k=0}^{+\infty} \sum_{j=0}^{k-1} \nu(f^j A_k) = 1 \tag{1.67}$$

and with the invariance measure we obtain:

$$\sum_{k=0}^{+\infty} k\nu(A_k) = 1 \tag{1.68}$$

and

$$\sum_{k=0}^{+\infty} k\nu(A_k) = \int_A \tau(x) d\nu(x) \tag{1.69}$$

□.

1.2.2 Return Time Statistics

Kac's theorem suggests that for every set B of positive measure, the expectation value of τ is equal to $1/\nu(B)$. For a wide class of dynamical systems an exponential behaviour of first return times is observed when $\nu(B) \to 0$. We are interested in finding out which are the suitable conditions in order to obtain this kind of distribution. This property is once again linked to the existence of mixing properties for the system considered, as stated in [37]. In order to frame rigorously the exponential decay of first return times we have at first to introduce some useful definitions:

Conditional Measure: *Let $(\Omega, \mathcal{F}, \nu)$ be an ergodic dynamical system. We define conditional measure of the set A:*

$$\nu_A(B) = \frac{\nu(A \cap B)}{\nu(A)} \qquad B \subseteq \Omega \tag{1.70}$$

Weak convergence *We say that a sequence of distribution functions*

$$F_n \qquad n = 1, 2, ...$$

weakly converges to a function F (which might not be a distribution itself) if F is non-strictly increasing, right continuous and satisfies:

$$\lim_{n\to\infty} F_n(t) = F(t) \tag{1.71}$$

at every point t of continuity of F .

Return times spectrum: *Let τ_A be the first return time over a set A and $\mathbb{E}(\tau)$ its expectation value, we define Return times Spectrum $\tilde{F}_A(t)$*

$$\tilde{F}_A(t) = \nu_A\left(x \in A : \frac{\tau_A(x)}{\tau_{\mathbb{A}}} \leq t\right) \tag{1.72}$$

Let us consider a neighbourhood sequence $A_K(x)$ of a point $x \in \Omega$ so that $\nu(A_k(x)) \to 0$ as $k \to +\infty$, we define *Limiting Spectrum* or Return Time Statistics $\tilde{F}(t)$:

$$\tilde{F}(t) = \lim_{k\to+\infty} \tilde{F}_{A_k(x)}(t) \tag{1.73}$$

if the limit exists.

1.2.3 Mixing Systems

There exist a great variety of mixing definitions which can be used to characterize this fundamental property of dynamical systems. We start by introducing an intuitive point of view, leaving the formalism for the following section.

Strong mixing system: *Let (Ω, f, ν) be a dynamical system, ν an invariant probability measure $\nu(f^{-1}A) = \nu(A)$. It is said to be strongly mixing if $\forall\varphi, \psi \in L^2(\Omega)$:*

$$\lim_{k\to\infty} \int_\Omega \varphi(f^{-k}(x))\psi(x)d\nu = \int_\Omega \varphi(x)d\nu \int_\Omega \psi(x)d\nu \tag{1.74}$$

For the special case of characteristic functions whose value is 1 at points of A (B) and 0 at points of $\Omega - A$ $(\Omega - B)$, we have that χ_A and χ_B of sets $A, B \in \Omega$:

$$\lim_{k\to\infty} \int_\Omega \chi_A(f^{-k}(x))\chi_B(x)d\nu = \int_\Omega \chi_A(x)d\nu \int_\Omega \chi_B(x)d\nu \tag{1.75}$$

$$\lim_{k\to\infty} \nu(f^{-k}A \cup B) = \nu(A)\nu(B) \tag{1.76}$$

and in the last passage we have used measure invariance.
Following Arnold, we present a simple physical example of mixing: the Cuba libre: suppose that a glass initially contains 20% rum (the set A) and 80% cola (the set B) in separate regions. After stirring the glass, any region of the glass contains approximately 20% rum. Furthermore, the stirred mixture is in a certain sense inseparable: no matter where one looks, or how small a region one looks at, one will find 80% cola and 20% rum [38]. In physics, a dynamical system is said to be mixing if the phase space has a *coarse-graining* structure. Every mixing transformation is ergodic, but there are ergodic transformations which are not mixing. An heuristic argument can be framed as follows:

Let us consider the set $A_{>t}$. It contains all the points which have a recurrence time greater than $t \cdot \mathbb{E}(\tau_A)$. Calling A_c the complementary set with respect to the relative measure, it is evident that $\forall x \in A_{>t}$ holds:

$$x \in A_{>T} \Longleftrightarrow f^k(x) \in A_c, \qquad k = \tau_A, \tau_A + 1, ..., t \cdot \mathbb{E}(\tau_{\mathbb{A}}) \tag{1.77}$$

When $\nu(A) \to 0$ then $\tau_A \gg 1$. If we apply the strong mixing condition (Eq. 1.74) and measure invariance, then:

$$\nu\left(f^{-k}(A_c) \cap f^{-n}(A_c)\right) = \nu(f^{-k}A_c)\nu(A_c) = \nu(A_c)^2 \tag{1.78}$$

which holds for $n \gg 1$. Now we can write:

$$\nu_A(A_{>t}) = \nu\left(f^{-\tau_A}(A_c) \cap f^{-\tau_A-1}(A_c) \cap ... \cap f^{-t\mathbb{E}(\tau_A)}(A_c)\right) \tag{1.79}$$

Since some sets is contained in the others, we obtain:

$$\nu_A(A_{>t}) = \nu\left(A_c\right)^{t\mathbb{E}(\tau_A)-\tau_A} \tag{1.80}$$

It follows that:

$$\nu_A(A_{>t}) = (1-\nu(A))^{t\mathbb{E}(\tau_A)-\tau_A} \tag{1.81}$$

$$\nu_A(A_{>t}) = \exp\left[(t\mathbb{E}(\tau_A) - \tau_A)\log(1-\nu(A))\right] \tag{1.82}$$

Eventually, taking the limit $\nu(A) \to 0$, we obtain:

$$\tilde{F}(t) = \lim_{\nu(A)\to 0} \nu_A\left(A_{>t}\right) = e^{-t} \tag{1.83}$$

1.2.4 Hitting Time Statistics

We have just introduced the Return Time Statistics $\tilde{F}(t)$ but we may also study the Hitting Time Statistics $F(t)$ that differ by $\tilde{F}(t)$ since τ_A is defined on the whole of Ω and not simply restricted to A:

$$F_A(t) = \nu\left(\frac{\tau_A(x)}{\tau_{\mathbb{A}}} \leq t\right) \tag{1.84}$$

$$F(t) = \lim_{k\to+\infty} F_{A_k(x)}(t) \tag{1.85}$$

To state rigorous results about Hitting Time Statistics, we focus our attention to the class of functions $\mathcal{A}$ and $\tilde{\mathcal{A}}$ which have the following properties [39]:

- $\mathcal{A} = \{F : \mathbb{R} \to [0,1], F \equiv 0 \text{ on }]-\infty, 0], F \text{ (non strictly) increasing,}$

  $$\text{continuous, concave on } [0.+\infty[, F(t) \leq t \text{ for } t \geq 0\}$$

- $\tilde{\mathcal{A}} = \{\tilde{F} : \mathbb{R} \to [0,1], F \equiv 0 \text{ on }]-\infty, 0], F \text{ (non strictly) increasing,}$

  $$\text{right-continuous, } \int_0^{+\infty} (1-\tilde{F}(s))ds \leq 1\}$$

Hitting time statistics

Following the work of [40], we introduce their mixing definitions and then proof the exponential spectral decay of recurrence times:
We say that $(\Omega, \mathcal{F}, \nu, f)$ is, for all integers $n \geq 1$ and $l \geq 0$,

- γ-*mixing* (or uniform mixing) if

$$\sup_{B\in\mathcal{F}_{\{0,\dots,n\}},C\in\mathcal{F}_{\{n\geq 0\}}} \left|\nu(B\cap f^{-(n+l+1)}) - \nu(B)\nu(C)\right| = \gamma(l) \qquad (1.86)$$

- α-*mixing* if

$$\sup_{B\in\mathcal{F}_{\{0,\dots,n\}},C\in\mathcal{F}_{\{n\geq 0\}}} \frac{\left|\nu(B\cap f^{-(n+l+1)}) - \nu(B)\nu(C)\right|}{\nu(B)} = \alpha(l) \qquad (1.87)$$

- ϕ-*mixing* if

$$\sup_{B\in\mathcal{F}_{\{0,\dots,n\}},C\in\mathcal{F}_{\{n\geq 0\}}} \frac{\left|\nu(B\cap f^{-(n+l+1)}) - \nu(B)\nu(C)\right|}{\nu(B)\nu(C)} = \phi(l) \qquad (1.88)$$

where $\mathcal{F}_{\{0,\dots,n\}}$ are finite or countable measurable σ-algebra.

Following [40], we state a general implication verified by the preceding types of mixing:

Remark: *ϕ-mixing implies α-mixing which implies uniform mixing:*

$$\gamma(n) \leq \alpha(n) \leq \phi(n) \qquad \forall n \in \mathbb{N}$$

We now state a fundamental theorem which connects the Return Time Statistics with the Hitting Time Statistics, showing that, if one converges the same applies to the other. This theorem prepares to the main results obtained by Freitas et al. [41] reported in Section 1.4.

Theorem (Haydn-Lacroix-Vaienti) *Let (Ω, f, ν) be an ergodic system and consider a sequence $\{A_n \in \Omega : n \geq 1\}$ a sequence of positive measure measurable subsets. Then the sequence of functions $\tilde{F}_{A_n}$ converges weakly if and*

only if the functions F_{A_n} converge weakly. Moreover, if the convergence holds, then:

$$F(t) = \int_0^t (1 - \tilde{F}(s))ds, \qquad t \geq 0 \tag{1.89}$$

where $\tilde{F}$ and F are the corresponding limiting (sub-probability) distributions The only previous result in this direction was obtained by [40] where it is shown that $\tilde{F}(A_n) \to \tilde{F}$ and $\tilde{F}(t) = 1 - e^{-t}$ for $t \geq 0$ if and only if $F(A_n) - \tilde{F}(A_n) \to 0$ in the supremum norm on the real line. The Haydn-Lacroix-Vaienti Theorem shows that the exponential distribution is the only distribution which can be asymptotic to both return and hitting times, as it is clearly the only fixed point of 1.89. Linking this two theorems one can obtain:

$$F(t) = \int_0^t (1 - \tilde{F}(s))ds = \int_0^t (1 - (1 - e^{-s}))ds = 1 - e^{-t} \tag{1.90}$$

Proof: Let us denote:

$$B_k = \{x \in A : \tau_A(x) = k\}$$

$$A_k = \{x \in \Omega : \tau_A(x) = k\}$$

Then, a set of zero measure over Ω can be written as a disjoint union of:

$$\bigcup_{j=0}^{k-1} f^j B_k, \qquad k = 1, 2, ...$$

On the other hand we can write:

$$A_k = \bigcup_{j=0}^{+\infty} f^j B_{k+j}$$

which implies:

$$\nu(A_k) = \sum_{j=k}^{+\infty} \nu(B_k) \tag{1.91}$$

$$F_A(t) = \nu\left(\bigcup_{k=1}^{[t/\nu(A)]} A_k\right) = \sum_{k=1}^{[t/\nu(A)]}$$

F_A is the partition function of a discrete aleatory variable so that it is a simple function : more exactly it is constant over the intervals

$$[k\nu(A), (k+1)\nu(A)[$$

and its value is established by the precedent sum.
It has a jump discontinuity of height $\nu(A_k)$ at the point $k\nu(A)$.
Let us define $\bar{F}_A$ the piecewise linear function which has the same values of F_A for all point $k\nu(A)$, and that is linear over the intervals $[k\nu(A), (k+1)\nu(A)]$. Using Eq. 1.91, all the jump discontinuities of F_A are decreasing which implies that $\bar{F}_A$ is concave. On the other hand it is continue and increasing and its right derivative for a point $t \in [k\nu(a), (k+1)\nu(A)[$ can be written as:

$$\tilde{F'}_A(t) = \frac{\nu(A_k + 1)}{\nu(A)} \tag{1.92}$$

in the same way, the return

$$\tilde{F}_A(t) = \frac{1}{\nu(A)} \sum_{k=1}^{[t/\nu(A)]} \nu(B_k) \tag{1.93}$$

is constant over $[k\nu(a), (k+1)\nu(A)[$ and has a jump of height $\frac{\nu(B_k)}{\nu(A)}$ in $k\nu(A)$.
$\forall t \geq 0$ we can write:

$$\bar{F'}_A(t) = 1 - \tilde{F}_A(t) \tag{1.94}$$

By the definition of $\bar{F}_A$ we get:

$$||F_A - \bar{F}_A||_\infty \leq \nu(A) \tag{1.95}$$

Let $(A_n)_n$ a sequence of measurable sets so that $\nu(A_n) \to 0$ and $\tilde{F}_{A_n} \Rightarrow \tilde{F}$ (in this case $\tilde{F} \in \tilde{\mathcal{A}}$). Since $\tilde{F}$ is increasing and $\in [0.1]$ then $\tilde{F}_{A_n} \to \tilde{F}$ a.e. over

$[0, +\infty[$. Now, by applying the dominated convergence theorem over $[0, t]$ for $t \geq 0$ and by using Eq. 1.94 we have:

$$\bar{F}_{A_n}(t) = \int_0^t (1 - \tilde{F}_{A_n}(s))ds \to \int_0^t (1 - \tilde{F}(s))ds =: F(t) \tag{1.96}$$

Posing $F(t) = 0$ if $t > 0$, then $\tilde{F} \in \tilde{\mathcal{A}}$ and $F \in \mathcal{A}$. On the other hand, by using Eq. 1.95, $F_{A_n}(t) \to F(t), \forall t \in \mathbb{R}$. Eventually, if $\tilde{F}_{A_n} \Rightarrow \tilde{F}$ then $F_{A_n} \Rightarrow F$ previously defined.

1.3 Reversibility Error and Correlations

Important statistical properties of dynamical systems can be inferred also studying the behaviour of the so called reversibility error and of the correlations functions. The former is a very well known tool in the scientific community for its versatility - it can be computed for a general class of observables - and it has also been widely used in the study of critical transition. The latter is widely commented in [1]. In this section we present the theory used to describe the asymptotic decay of correlations which is strictly connected to recurrences statistics and extreme value theory as we show through the papers enclosed.

The decay of correlations plays a very important role in non-equilibrium statistical mechanics. It is essential in the studies of relaxation to equilibrium. Correlations and autocorrelations functions are explicitly involved in the formulas for transport coefficients, such as heat conductivity, electrical resistance, viscosity, and the diffusion coefficient [42].

1.3.1 Reversibility Error

It is possible to introduce the reversibility error after n iteration of a map f as:

$$\Delta_n = \text{dist}(f^{-n}(f^n(\vec{x})), \vec{x}) \tag{1.97}$$

This is clearly a quantity obtained only numerically. In a pure analytical set-up $\Delta_n = 0$ always.
Using this definition [43] have investigated numerically the role of numerical round-off. In a regular region Δ_n is about of the same order of magnitude of the machine precision $10^{-13} - 10^{-16}$ while it is orders of magnitude greater in chaotic regions.
For this reason, by using this indicator it is possible to distinguish chaotic from regular region with a great detail. In [1] we present some examples of map with different kind of regions depicted using reversibility error.

1.3.2 Correlations function

Let $(\Omega, \mathcal{F}, \nu, f)$ be a dynamical system and let us denote by f^n the dynamical evolution on the phase space Ω with ergodic measure ν, then, for any choice of $\varphi, \psi \in L^2(\Omega, \nu)$, we define the correlation function as:

$$C_{\varphi,\psi}(n) = \int_\Omega \varphi(f^n x)\psi(x)d\nu(x) - \int_\Omega \varphi(x)d\nu(x) \cdot \int_\Omega \psi(x)d\nu(x) \qquad (1.98)$$

The key property for analysing the dynamics of a system is the speed at which the correlation function decay.We will distinguish the cases for which the correlations function $C_{\varphi,\psi}(n)$ decays exponentially with n - for suitable φ, ψ -, from other cases we will observe a power law decay:

- *decay exponentially* if $|C_{\varphi,\psi}(n)| < \text{const} \cdot e^{-cn}$ for some $c > 0$.
- *decay polynomially* if $|C_{\varphi,\psi}(n)| < \text{const} \cdot n^{-\alpha}$ for some $\alpha > 0$.

One way to obtain analytical results for simple dynamical systems is to consider the Fourier transform of the correlations function:

$$\hat{C}_{\varphi,\psi}(\omega) = \sum_{n\in\mathbb{Z}} e^{in\omega} C_{\varphi,\psi}(n) \qquad (1.99)$$

By substituting $e^{in\omega}$ with z in the expression above, we obtain an equation in form of a power series in z. We know the first n coefficients up to the

maximum correlation time. In [31] and [44] the authors have shown that the exponential decay is associated to the analyticity of $\hat{C}_{\varphi,\psi}(\omega)$: if it has a pole in $\omega* = \alpha + \mathrm{i}\beta$, its contribution to the correlation function is written as:

$$\exp(-\beta \cdot n + \mathrm{i}\alpha \cdot n)$$

The analyticity of $\hat{C}_{\varphi,\psi}(\omega)$ regards a strip centred around the real axis. The properties of the Fourier transform of the correlations function have been studied in [45] and [46] showing that the behaviour of the system is connected to its analyticity properties.

1.3.3 Correlations and Recurrences

To show the connections between correlations decay and Poincaré Recurrences, we follow Young's method for hyperbolic dynamical systems [47]. We recall that hyperbolic dynamics is characterized by the presence of expanding and contracting directions for the derivative. This stretching and folding typically gives rise to complicated long-term behaviour in these systems (see [48] for a comprehensive description).
We describe the basic idea of Young's method shown in [49], skipping mathematical and technical details which can be found in the original papers.
Let $f : \Omega \to \Omega$ be an hyperbolic map acting on Ω and ν an absolute continue ergodic measure. We are looking for sufficient conditions under which correlations for the map f have a summable decay. Young constructs a set Δ_0 with a hyperbolic structure obtained by intersection of a family of unstable manifolds with a family of stable manifolds similar to an 'horseshoe'. We iterate points $x \in \Delta_0$ using the map f until they make return to Δ_0.
We can now construct a tower, Δ, with base Δ_0, where the induced map f_Δ moves every points one floor up until they hit the ceiling, eventually falling onto the base again. The tower Δ is identified with Ω, and f_Δ with f. Now, we can introduce the first return time defined in Eq. 1.60 which in this case, for $x \in \Delta$, can be written as:

$$\tau_{\Delta_0}(x) = \inf\{k > 0 : f^k_\Delta(x) \in \Delta_0\} \tag{1.100}$$

that is the first return time of the point X to the base of the tower Δ_0. Since the tower has infinitely many levels, τ is unbounded. As in Eqs. 1.72 and 1.73, we can introduce the Hitting Times statistics:

$$F(x \in \Delta : \tau_{\Delta_0}(x) > n).$$

Young proves both that if the Hitting Time Statistics is exponential (that is exponentially small probability of long returns), then correlations function decays exponentially and that correlations decay polynomially for systems with slower mixing rates:

Theorem (Young) *Let $f : \Omega \to \Omega$ be an hyperbolic map acting on Ω and ν an absolute continue ergodic measure, if*

- *The Hitting Time Statistics shows an exponential decay:*

$$F(x \in \Delta : \tau_{\Delta_0}(x) > n) \leq \text{const} \cdot \theta^n \quad \forall n \geq 1 \tag{1.101}$$

  where $\theta < 1$ is a constant, then

$$|C_{\varphi,\psi}(n)| < \text{const} \cdot e^{-cn} \quad \text{for some } n > 0 \tag{1.102}$$

- *The Hitting Time Statistics shows a power law decay:*

$$F(x \in \Delta : \tau_{\Delta_0}(x) > n) \leq \text{const} \cdot n^{-\alpha} \quad \forall n \geq 1 \tag{1.103}$$

  where $\alpha > 0$ is a constant, then

$$|C_{\varphi,\psi}(n)| < \text{const} \cdot n^{-\alpha} \tag{1.104}$$

To verify directly the tail bound in specific systems, we have to iterate the map f and construct a Δ_0, an approach that might be quite difficult.

To overcome this problem, we can use another approach, which allows for avoiding the computation of Young's tower. Imagine that one can localize, on the Ω manifold, a set $M \subset \Omega$ having a strong hyperbolic behaviour. Consider the hitting time statistics $F : M \to M$, then there exists an horseshoe $\Delta_O \subset M$ which return times are exponentially bounded under the map f.

We state now the main theorem for the power law decay of correlations. It will be useful to better understand the results presented in [8], where we use extensively the summable decay of correlations.

Theorem (Markarian): *Let $f : \Omega \to \Omega$ a non-uniformly hyperbolic map. Suppose $M \subset \Omega$ is a subset such that the Hitting time statistic $F : M \to M$ satisfies Eq. 1.101 for $\tau_{\Delta_0}(x)$ to a rectangle $\Delta_0 \subset M$. If the return times $\tau_M(x)$ satisfies the bound*:

$$F(x \in \Omega : \tau_M(x) > n) \leq \text{const} \cdot n^{-\alpha} \qquad \forall n \geq 1 \tag{1.105}$$

then

$$|C_{\varphi,\psi}(n)| < \text{const} \cdot (\ln n)^{\alpha+1} n^{-\alpha} \tag{1.106}$$

Proof: For every $n \geq 1$ and $x \in \Omega$ define:

$$r_M(x) = \{1 \leq i \leq n : f^i(x) \in M\} \tag{1.107}$$

and

$$\begin{aligned} A_n &= \{x \in \Omega : \tau_{\Delta_0}(x) > n\} \\ B_{n,b} &= \{x \in \Omega : r_M(x) > b \ln n\} \end{aligned} \tag{1.108}$$

where b is a constant to determine. Using equation 1.101:

$$\nu(A_n \cap B_{n,b}) \leq \text{const} \cdot n\theta^{b \ln n}$$

Choosing b large enough makes this bound less than const $\cdot\, n^{-\alpha}$.
We have to bound $\nu(A_n \backslash B_{n,b})$ noting that points $x \in A_n \backslash B_{n,b}$ return to M

at most $b \ln n$ times during the first n iterates of f. There are less or almost $b \ln n$ time intervals between successive returns to M and the longest such interval I has length $\geq n/(b \ln n)$. By applying equation to the interval I we obtain:

$$\nu(A_n \backslash B_{n,b}) \leq \text{const} \cdot \frac{n(\ln n)^{\alpha+1}}{n^{\alpha+1}}$$

Note that Markarian includes the extra factor n because the interval I may appear anywhere within the interval $[1, n]$, and the measure ν is invariant. Returning to Young's tower Δ, we obtain:

$$F(x \in \Delta : \tau_{\Delta_0}(x) > n) \leq \text{const} \cdot (\ln n)^{\alpha+1} n^{-\alpha} \qquad n \geq 1 \tag{1.109}$$

and this bound differs from Young's by the extra factor $(\ln n)^{\alpha+1}$. A similar argument can be used to prove a bound for the power law decay.

1.4 Hitting time statistics and Extreme Value Theory

In this section we explore the link between Extreme Value Laws presented in the first section and the Hitting Time Statistics, described in section three, following the paper by Freitas et al. [10]. Their main finding has been showing the deep connection between the two theories for dynamical systems which exhibit an exponential return time statistics.
The results presented in this section are widely used among the enclosed papers. We wish to rephrase them here in a less technical language that can be useful to access the results presented in the subsequent chapters.

Uniform mixing conditions

We have introduced the statistics of maxima $M_n = \max\{X_0, ..., X_{n-1}\}$ for i.i.d variables $X_0, X_1..$ and we have already observed in 1.1.3 that exists a limit distribution under certain conditions.

The same limit laws apply to stationary stochastic processes, under certain conditions on the dependence structure, which allows for the reduction to the independent case. Freitas et al. [10] associate to a given stochastic process $X_0, X_1, ...$ an i.i.d. sequence $Y_0, Y_1, ...$ whose distribution function is the same as that of X_0, and whose partial maximum is defined as:

$$\hat{M}_n = \max\{Y_0, ..., Y_{n-1}\}$$

In the dependent context, the general strategy is to prove that if $X_0, X_1, ..$ satisfies some conditions, then the same limit law for $\hat{M}_n$ applies to M_n with the same normalizing sequences a_n and b_n. See [50] for the complete proof. This is equivalent to showing that the following mixing conditions hold.

Condition $D_2(u_n)$. We say that $D_2(u_n)$ holds for the sequence $X_0, X_1, ...$ if for any integers l, t and n

$$|\nu(\{X_0 > u_n\} \cap \{\max\{X_y, ..., X_{t+l-1}\}\}) - \nu(\{X_0 > u_n\})\nu(\{M_l \leq u_n\})| \leq \gamma(n, t)$$

where $\gamma(n, t)$ is non-increasing in t for each n.

Condition $D'(u_n)$. We say that $D'(u_n)$ holds for the sequence $X_0, X_1, ...$ if

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=1}^{[n/k]} \nu(\{X_0 > u_n\} \cap \{X_j > u_n\}) = 0$$

The main result states that if $D_2(u_n)$ and $D'(u_n)$ hold for the process $X_0, X_1, ...$, then the following limits exists:

$$\lim_{n\to\infty} \nu(\hat{M}_n \leq u_n) = \lim_{n\to\infty} \nu(M_n \leq u_n) \tag{1.110}$$

Before stating the main results, we recall the Lebesgue's differentiation theorem which will have a key role in the subsequent proofs.

Lebesgue's Differentiation Theorem *Let us consider a Lebesgue integrable real or complex-valued function f on $\mathbb{R}^n$, the indefinite integral is a set function which maps a measurable set A to the Lebesgue integral of $f \cdot \chi_A$, where χ_A denotes the characteristic function of the set A. We can write it as:*

$$\int_A f \, \mathrm{d}\nu_{Leb}$$

with ν_{Leb} the n-dimensional Lebesgue measure.
The derivative of this integral at x is defined to be

$$\lim_{B \to x} \frac{1}{|B|} \int_B f \, \mathrm{d}\nu_{Leb},$$

where $|B|$ denotes the volume (i.e., the Lebesgue measure) of a ball B centred at x, and $B \to x$ means that the diameter of B tends to zero. Then this derivative exists and is equal to $f(x)$ at almost every point $x \in \mathbb{R}^n$.

The points x for which this equality holds are called Lebesgue points.

1.4.1 Freitas-Freitas-Todd Theorem

Let $(\Omega, \mathcal{F}, \nu, f)$ be a dynamical system where ν is absolute continuous invariant probability measure (that is $\nu(f^{-1}(A)) = \nu(A)$) for all $A \in \Omega$, a d-dimensional Riemannian Manifold . We introduce:

- $B_\delta(\zeta) = \{x \in \Omega : \mathrm{dist}(x, \zeta) < \delta\}$ is the ball of radius δ centred in ζ
- ν_{Leb} is Lebesgue measure on Ω.
- $\rho = \frac{d\nu}{d\nu_{Leb}}$ is a measure density.
- 'dist' a Riemannian metric on Ω.

Consider the stationary stochastic process $X_0, X_1, ...$ which is given by:

$$X_n(x) = g(\mathrm{dist}(f^n x, \zeta)) \qquad \forall n \in \mathbb{N} \tag{1.111}$$

Define the partial maxima as:

$$M_n = \max\{X_0, ..., X_{n-1}\} \tag{1.112}$$

As in subsection we are interested in knowing if there are normalizing sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ such that:

$$\nu(\{x : a_n(M_n - b_n) \leq y\}) = \nu(\{x : M_n \leq u_n\}) \to G(y) \tag{1.113}$$

where $u_n = u_n(x) = x/a_n + b_n$.

Observable functions g_i and normalizing sequences

Following Freitas et al. [41] we now describe the three types of observable $g_i \quad i = 1, 2, 3.$ which are suitable to obtain a GEV distribution for normalized maxima according the Gnedenko's theorem stated in subsection 1.1.3:

Observation 4.1: *Let ζ be a chosen point in the phase space Ω and consider the function:*

$$g(\text{dist}(x, \zeta)) : \Omega \to \mathbb{R} \tag{1.114}$$

g is such that 0 is a global maximum. $g : V \to W$ is a strictly decreasing bijection and it has one of the following behaviour:

- *Type 1: There exists some positive function $p : W \to \mathbb{R}$ such that for all $y \in R$*

$$\lim_{s\to g_1(0)} \frac{g_1^{-1}(s + yp(s))}{g_1^{-1}(s)} = e^{-y} \tag{1.115}$$

- *Type 2: $g_2(0) = +\infty$ and there exists $\beta > 0$ such that $\forall y > 0$*

$$\lim_{s\to+\infty} \frac{g_2^{-1}(sy)}{g^{-1}(s)} = y^{-\beta} \tag{1.116}$$

- *Type 3: $g_3(0) = D < +\infty$ and there exists $\gamma > 0$ such that for all $y > 0$*

$$\lim_{s\to 0} \frac{g_3^{-1}(D - sy)}{g_3^{-1}(D - s)} = y^{\gamma} \tag{1.117}$$

g_i conditions are just translations in terms of the shape of g^{-1} of the conditions on the tail of F, since the distribution function F is given by:

$$F(u) = \nu(X_0 \leq u)$$

Note that in this case the dynamic of the system is not introduced in F definition.

$$1 - F(u) = 1 - \nu(X_0 \leq u) \tag{1.118}$$

It follows from X_0 definition that:

$$1 - F(u) = 1 - \nu(g(\text{dist}(x, \zeta) \leq u)) \tag{1.119}$$

using Lebesgue Differentiation Theorem:

$$1 - F(u) = 1 - \nu(\text{dist}(x, \zeta) \geq g^{-1}(u)) \tag{1.120}$$

If we observe that:

$$\nu(\text{dist}(x, \zeta) \geq g^{-1}(u)) = 1 - \rho(\zeta)|B_{g^{-1}(u)}(\zeta)|$$

Then we obtain:

$$1 - F(u) \sim \rho(\zeta)|B_{g^{-1}(u)}(\zeta)|$$

$$\text{dist}(f^j x, \zeta) \leq g^{-1}(u) \tag{1.121}$$

It is now possible to apply the corollary in subsection 1.1.3 in order to obtain the normalizing sequences a_n e b_n:

Observation 4.2: *Let $\{\delta_n\}_{n\in\mathbb{N}}$ be such that $\delta_n \to 0$ if $n \to \infty$, for each $\zeta \in \Omega$ and define $k \in]0, \infty[$ such that $|B_{\delta_n}(\zeta)| \sim k \cdot \delta_n^d$, then:*

- *Type g_1:* $y \in \mathbb{R}$ $\quad u_n = g_1((k\rho(\zeta)n)^{-1/d}) + p(g_1((k\rho(\zeta)n)^{-1/d}))\frac{y}{d}$
- *Type g_2:* $y > 0$ $\quad u_n = g_2((k\rho(\zeta)n)^{-1/d})y$
- *Type g_3:* $y < 0$ $\quad u_n = D - (D - g_3((k\rho(\zeta)n)^{-1/d}))(-y)$

We are now ready to state the main result obtained by Freitas et al. which links the extreme value laws presented in Section 1.1 with the hitting time statistics widely presented in Section 1.2.4.

Theorem (Freitas-Freitas-Todd): *Let $(\Omega, \mathcal{F}, \nu, f)$ be a dynamical system where ν is an absolute continue invariant probability measure, and consider $\zeta \in \Omega$ for which Lebesgue's Differentiation Theorem holds.*

- *If we have an exponential Hitting Time Statistics $F(t) = e^{-t}$ (defined in Eq. 1.85) to balls centred on $\zeta \in \Omega$, then we have an Extreme Value Law for M_n which applies to the observables g_i (described in Observation 4.1) achieving a maximum at ζ.*

- *If we have an exponential Hitting Time Statistics to balls at $\zeta \in \Omega$, then we have an Extreme Value Law for M_n which coincides with that of $\hat{M}_n$ (meaning that Eq. 1.110 holds). In particular, this Extreme Value Law must be one of the 3 classical types described in Eq. 1.3 for observables g_1, in Eq. 1.4 for observables g_2 and in Eq. 1.5 for observables g_3 where, in all the cases considered, $\mu = 0$ and $\sigma = 1$ since normalizing sequences a_n and b_n are used.*

Proof: We outline the part of the proof which demonstrate the theorem for the observable class g_1. The complete proof for all the observables g_i can be found in [10]. Let us denote g_i simply as g. The outline is divided into two parts:

- ***Part 1*** For n sufficiently large we have:

$$\{x : M_n(x) \leq u_n\} = \bigcap_{j=0}^{n-1} \{x : g(\text{dist}(f^j x, \zeta)) \leq u_n\} \tag{1.122}$$

We now apply the Lebesgue Differentiation Theorem:

$$\bigcap_{j=0}^{n-1}\{x : g(\text{dist}(f^j x, \zeta)) \leq u_n\} = \bigcap_{j=0}^{n-1}\{x : \text{dist}(f^j x, \zeta) \geq g^{-1}(u_n)\} \tag{1.123}$$

note the change in the inequality sign. The last equation allow us to introduce the hitting time statistics over the ball $B_{g^{-1}(u_n)}$:

$$\bigcap_{j=0}^{n-1}\{x : \text{dist}(f^j x, \zeta) \geq g^{-1}(u_n)\} = \{x : \tau_{B_{g^{-1}(u_n)}}(x) \geq n\} \tag{1.124}$$

From Eq. 1.115 we can write:

$$\begin{aligned} g^{-1}(u_n) &= g^{-1}\left[g_1((k\rho(\zeta)n)^{-1/d}) + p(g_1((k\rho(\zeta)n)^{-1/d}))\frac{y}{d}\right] \\ &\sim g^{-1}\left[g_1((k\rho(\zeta)n)^{-1/d})\right] e^{-y/d} = \left(\frac{e^{-y}}{k\rho(\zeta)n}\right)^{1/d} \end{aligned} \tag{1.125}$$

Thus:

$$g^{-1}(u_n) \sim \left(\frac{e^{-y}}{k\rho(\zeta)n}\right)^{1/d} \tag{1.126}$$

As said before, since Lebesgue's Differentiation Theorem holds for $\zeta \in \Omega$, we have $\frac{\nu(B_\delta(\zeta)}{|B_\delta(\zeta)|} \to \rho(\zeta)$ as $\delta \to 0$. Consequently, since it is obvious that $g^{-1}(u_n) \to 0$ as $n \to \infty$, then:

$$\nu(B_{g^{-1}u_n)}(\zeta)) \sim \rho(\zeta)|B_{g^{-1}(u_n)}(\zeta)| \sim \rho(\zeta)k(g^{-1}(u_n))^d \tag{1.127}$$

and the last passage is obtained by using the relation 1.121.

$$\rho(\zeta)k(g^{-1}(u_n))^d = \rho(\zeta)k\frac{e^{-y}}{k\rho(\zeta)n} = \frac{e^{-y}}{n} \tag{1.128}$$

We can now express n as:

$$n \sim \frac{e^{-y}}{\nu(B_{g^{-1}(u_n)(\zeta)})} \tag{1.129}$$

Now, by substituting Eq. 1.129 in Eq 1.124 and taking the limit, we have:

$$\lim_{n\to\infty} \nu\left(\left\{x : \tau_{B_{g^{-1}(u_n)}}(x) \geq \frac{e^{-y}}{\nu(B_{g^{-1}(u_n)(\zeta)})}\right\}\right) = F(e^{-y}) = e^{-e^{-y}} \tag{1.130}$$

Eventually it is possible to write:

$$\lim_{n\to\infty} \nu(\{x : M_n(x) \leq u_n\}) = e^{-e^{-y}} \tag{1.131}$$

which is Eq. 1.3 with $\mu = 0$ and $\sigma = 0$ obtained using the normalizing sequences as stated by Gnedenko (see Section 1.1.3).

- ***Part 2:*** We use the exponential Hitting Time Statistics hypothesis

$$F(t) = e^{-t} \tag{1.132}$$

to proof the second part. It follows by Eqs. 1.131 and 1.110 that also:

$$\lim_{n\to\infty} \nu(\{x : \hat{M}_n(x) \leq u_n\}) = e^{-e^{-y}} \tag{1.133}$$

which means that the Extreme Value Laws applies to M_n (rather than $\hat{M}_n$). In the case of observable g_1 which we have considered this is true $\forall y \in \mathbb{R}$.

1.5 Methodological notes and informatics tools

The numerical analysis contained in the enclosed paper have required to devise and write several codes and involved the use of hundreds and hundreds CPU hours, producing several GB of data. Some of the results have

been presented by using MATLAB Statistics Toolbox function such as *gevfit* and *gevcdf*. These functions returns maximum likelihood estimates of the parameters for the generalized extreme value (GEV) distribution and 95% confidence intervals for the parameter estimates [51].
The basic code have been devised to follow a well defined algorithmic procedure which consists in iterating the map, dividing the series into n bins each containing m observation, extracting the maxima or the minima in each bin and fitting them to the GEV distribution. Other self written codes have been used to complete the analysis. Further explanations may be found in the individual papers.

Chapter 2

Article: Analysis of Round Off Errors with Reversibility Test as a Dynamical Indicator

International Journal of Bifurcation and Chaos, Vol. 22, No. 9 (2012) 1250215 (14 pages)

DOI: 10.1142/S021812741250215X

ANALYSIS OF ROUND OFF ERRORS WITH REVERSIBILITY TEST AS A DYNAMICAL INDICATOR

DAVIDE FARANDA
Department of Physics, University of Bologna,
Via Irnerio 46, Bologna 40126, Italy
davide.faranda@studio.unibo.it

MARTÍN FEDERICO MESTRE
Grupo de Caos en Sistemas Hamiltonianos,
Facultad de Ciencias Astronómicas y Geofísicas,
Universidad Nacional de La Plata,
Instituto de Astrofísica de La Plata
(CCT La Plata – CONICET, UNLP),
Paseo del Bosque S/N, La Plata 1900, Argentina
mmestre@fcaglp.unlp.edu.ar

GIORGIO TURCHETTI
Department of Physics, University of Bologna,
INFN-Bologna, Via Irnerio 46,
Bologna 40126, Italy
turchett@bo.infn.it

Received November 8, 2010; Revised May 31, 2011

We compare the divergence of orbits and the reversibility error for discrete time dynamical systems. These two quantities are used to explore the behavior of the global error induced by round off in the computation of orbits. The similarity of results found for any system we have analyzed suggests the use of the reversibility error, whose computation is straightforward since it does not require the knowledge of the exact orbit, as a dynamical indicator.

The statistics of fluctuations induced by round off for an ensemble of initial conditions has been compared with the results obtained in the case of random perturbations. Significant differences are observed in the case of regular orbits due to the correlations of round off error, whereas the results obtained for the chaotic case are nearly the same.

Both the reversibility error and the orbit divergence computed for the same number of iterations on the whole phase space provide an insight on the local dynamical properties with a detail comparable with other dynamical indicators based on variational methods such as the finite time maximum Lyapunov characteristic exponent, the mean exponential growth factor of nearby orbits and the smaller alignment index. For 2D symplectic maps, the differentiation between regular and chaotic regions is well full-filled. For 4D symplectic maps, the structure of the resonance web as well as the nearby weakly chaotic regions are accurately described.

Keywords: Discrete time dynamical systems; reversibility error; standard map; LCE; SALI; GALI; MEGNO; mLCE.

1. Introduction

The discrete time dynamical systems are widely used because they present a rich structure (regular, chaotic, intermittent and other types of orbits) in one or two dimensions and because the numerical evaluation of the orbits is straightforward. A large number of dynamical tools known as indicators of stability allow us to improve our understanding of dynamical systems: Lyapunov Characteristic Exponents (LCEs) are known from a long time [Wolf *et al.*, 1985; Rosenstein *et al.*, 1993; Skokos, 2010] as well as Return Times Statistics [Kac, 1934; Gao, 1999; Hu *et al.*, 2004; Buric *et al.*, 2005]. In the recent past, many other indicators have been introduced not only to address the same problem of quantifying the degree of chaoticity of an orbit but also to perform the task fast. The Smaller Alignment Index (SALI), widely described in [Skokos *et al.*, 2002, 2004], allows to discriminate regular from chaotic orbits. Similarly, the Generalized Alignment Index (GALI), introduced in [Skokos *et al.*, 2007], is a family of highly efficient algorithms. Besides, the Mean Exponential Growth factor of Nearby Orbits (MEGNO) discussed in [Cincotta *et al.*, 2003; Goździewski *et al.*, 2001] is a quantity that gives a fast identification of the chaoticity of the orbit while its average slope estimates the maximum LCE (mLCE), see [Maffione *et al.*, 2011] for a test of the MEGNO. Fidelity and correlation decay are also tools that can be successfully used to characterize stability properties as explained in [Vaienti *et al.*, 2007; Turchetti *et al.*, 2010b] as well as Frequency Map Analysis [Laskar, 1999; Robutel & Laskar, 2001].

As it is already known, in any numerical computation of a given trajectory, there is a round off error, and it would be interesting to study its relation with the chaoticity of the orbit and if possible, determine its effect with the help of some dynamical indicator.

The error between an exact and a numerical orbit is due to the finite precision used to represent real numbers and to the arithmetics with round off. This is unavoidable because the length of the binary strings representing real numbers must not change after arithmetic operations. The shadowing lemma is often invoked to state the existence of a true orbit close to a numerical orbit for chaotic systems [Katok & Hasselblatt, 1997; Hammel *et al.*, 1987; Chow & Palmer, 1992, 1991]. However, it does not provide information on error growth for a given numerical orbit and the case of regular numerical orbits is not covered by such a lemma. The global error between the exact and the numerical orbit is unknown because the first one is not computable. Nevertheless, an estimate can be provided by replacing the exact orbit with another one having very high accuracy. If the map is invertible, the reversibility error can be computed without any reference to the exact orbit. Both errors have a similar behavior, namely an average linear growth for regular orbits and an average exponential growth for chaotic orbits and consequently can be used as dynamical indicators. Due to the correlation between the single step errors, there is a substantial difference with respect to the case in which the exact system is perturbed with random uncorrelated noise. This difference has been analyzed in [Turchetti *et al.*, 2010b] by using the fidelity which measures the deviation of the orbits of a given map and its perturbation by integrating over all the initial conditions with the appropriate measure. In the case of regular orbits with random perturbations, the decay of fidelity is exponential whereas with round off errors the decay follows a power law. In the case of chaotic orbits the asymptotic limit is approached super-exponentially for both situations (random uncorrelated perturbations and round off). The symplectic maps of physical interest are generally provided by the composition of a linear map with another one whose generating function is the identity plus a function of position or momentum only. In this case, the inversion is immediately obtained in analytic form. As most of the symplectic maps used in the literature are of this kind, the condition of invertibility of the map is not too restrictive. In addition, the reversibility error can also be applied to time-reversible systems of differential equations.

Here, we present an analysis of the reversibility error and its comparison with the divergence of orbits due to round off and other dynamical indicators of stability such as SALI, MEGNO and a finite time numerical estimation of the mLCE. Moreover, we talk about the similarities and differences between an irreversibility due to the single precision round off and one due to the application of uncorrelated random noise in an orbit iterated with double precision. For two-dimensional area preserving maps the reversibility error for a fixed number of iterations detects the various regions of phase space with different stability properties quite effectively

as well as other dynamical indicators. Besides, the reversibility error allows to study the structure of the resonance web of a four-dimensional symplectic map.

2. Round Off Error Methods

In a computational device a real number x can be represented by a floating-point number x^*, that according to Goldberg [1991] and working with base 2, can be written as

$$x^* = \pm d_0.d_1 \cdots d_{p-1} \times 2^e = \pm \sum_{k=0}^{p-1} d_k 2^{e-k} \qquad (1)$$

where $d_0.d_1 \cdots d_{p-1}$ is called the significand and has p binary digits d_k, whose value is 0 or 1, and where the exponent is an integer that satisfies $e_{\min} \leq e \leq e_{\max}$. The IEEE 754 standard states that for single precision $p = 24$, $e_{\min} = -126$ and $e_{\max} = 127$ while for double precision $p = 53$, $e_{\min} = -1022$ and $e_{\max} = 1023$.

Consequently, a floating point number x^* differs from the real number it represents and the relative error r_p, defined by $x^* = x(1 + r_p)$, satisfies $|r_p| \leq \epsilon \equiv 2^{-p}$, as analyzed by Goldberg [1991] and Knuth [1973]. Therefore, according to IEEE 754 we have that $\epsilon = 2^{-24}$ and $\epsilon = 2^{-53}$ for single and double precision which roughly correspond to seven and 16 decimal digits, respectively. The arithmetic operations such as sums or multiplications imply a round off, which propagates the error affecting each number. Round off algebraic procedures are hardware dependent as detailed in [Knuth, 1973]. Unlike the case of stochastic perturbations, the error strongly depends on x. Suppose we are given a map $M(x)$ then the error with respect to the numerical map $M_*(x)$ after the first step is defined by:

$$\delta_1 \equiv \epsilon\xi_1 = x_1^* - x_1 \equiv M_*(x) - M(x). \qquad (2)$$

Analogously, we define the local error produced in the nth step by $\delta_n \equiv \epsilon\xi_n = M_*(x_{n-1}^*) - M(x_{n-1}^*)$ where $\xi_n = \xi_n(x_{n-1}^*)$.

The global error

$$G_n = M_*^n(x) - M^n(x) \qquad (3)$$

accumulates all the local errors and explicit expressions can be written at first order in ϵ. In the example of a regular map we take the translation on the torus $\mathbb{T}^1$ defined as:

$$M(x) = x + \omega \mod 1, \qquad (4)$$

so $M_*(x) = x^* + \omega^* \mod 1$ and the global error, which includes also the error to the mod 1 operation, becomes:

$$G_n = \epsilon \sum_{k=1}^{n} \xi_k = \epsilon(n\overline{\xi} + w_n), \qquad (5)$$

where $\epsilon\overline{\xi}$ is a time average defined as the limit of G_n/n for $n \to \infty$ and w_n is a bounded fluctuation.

For the chaotic Bernoulli map

$$M(x) = qx \mod 1, \qquad (6)$$

we have that $M_*(x) = qx^* \mod 1$ and that the single step and global error satisfy, respectively, $|\delta_1| \leq C\epsilon q$ and $|G_n| \leq C\epsilon q^n$.

In order to compute the global error we need the knowledge of the exact map, which is usually precluded. A practical way to overcome this difficulty is to replace the exact map with a map computed with an accuracy 2^{-P} where $P \gg p$. For instance, as $p = 24$ corresponds to single precision one might choose $P = 53$ corresponding to double precision. If $p = 53$ then one might choose $P = 100$ and so forth. There are available libraries which allow to compute with any fixed number of bits or significant decimal digits. The computation of the reference orbit is expensive if high precision is used, but there is no other way to evaluate the global error. As a consequence, the "exact" orbit is achievable for a definite number of iterations which depends on P and the nature of the map. Taking into account what we have just mentioned, in the forthcoming numerical experiments, we will use the divergence of orbits, defined by:

$$\Delta_n = M_S^n(x) - M_D^n(x), \qquad (7)$$

where M_S and M_D stand for single and double precision iterations, respectively.

If the map is invertible there is another option to overcome the difficulty of not possessing the true map. We define the reversibility error as

$$R_n = M_*^{-n} \circ M_*^n(x) - x \qquad (8)$$

which is nonzero since the numerical inverse M_*^{-1} of the map is not exactly the inverse of M_* namely $M_*^{-1} \circ M_*(x) \neq x$. Obviously the reversibility error is much easier to compute than the divergence of orbits (if we know explicitly the inverse map) and the information it provides is basically the same as the latter. Both quantities give an average linear growth for a regular map together with an

exponential growth for a chaotic map having positive Lyapunov exponents and strong mixing properties. When computing R_n we will set $M_* = M_S$ in order to compare with Δ_n.

3. Variational Methods

In the forthcoming sections, we will compare the performance of the indicators presented above with three well known and widely accepted dynamical indicators that are based on the behavior of the solution of the variational equations of the system. These are the finite time mLCE, the cumulative moving time average of MEGNO and SALI.

Let us briefly state them for discrete time dynamical systems of the form:

$$\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n), \tag{9}$$

where $\mathbf{x}_n$ is the state vector at time n and $\mathbf{f}$ is a vector valued function.

The concomitant discrete time variational equations, also called tangent map dynamics, associated to a given orbit $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ are the following:

$$\mathbf{v}_{n+1} = \mathbf{Df}(\mathbf{x}_n) \cdot \mathbf{v}_n, \tag{10}$$

where $\mathbf{Df}(\mathbf{x})$ is the Jacobian matrix of the function $\mathbf{f}$ and $\mathbf{v}_n$ is a deviation vector at time n.

Skokos [2010] presented a historical review of the definition of the LCEs and its connection with the divergence of nearby orbits. He also stated the theorems that guarantee the existence of the spectrum of LCEs and, in particular, the current definition of the mLCE in terms of the solution of the variational equations:

$$\text{mLCE} \equiv \lim_{n\to\infty} \frac{1}{n} \ln \frac{\|\mathbf{v}_n\|}{\|\mathbf{v}_0\|}, \tag{11}$$

with $\|\cdot\|$ some norm. For a chaotic orbit the mLCE is positive and this implies an exponential divergence of nearby orbits. On the other hand, for regular orbits mLCE is zero.

In order to have a numerically computable quantity we define the finite time mLCE at time n as

$$\text{mLCE}(n) \equiv \frac{1}{n} \ln \frac{\|\mathbf{v}_n\|}{\|\mathbf{v}_0\|} = \frac{1}{n} \sum_{k=1}^{n} \ln \frac{\|\mathbf{v}_k\|}{\|\mathbf{v}_{k-1}\|}, \tag{12}$$

so that Eq. (11) can be reformulated as

$$\text{mLCE} = \lim_{n\to\infty} \text{mLCE}(n). \tag{13}$$

Cincotta *et al.* [2003] defined a biparametric family of MEGNO indicators:

$$Y_{m,j}(n) = (m+1)n^j \sum_{k=1}^{n} k^m \ln \frac{\|\mathbf{v}_k\|}{\|\mathbf{v}_{k-1}\|}, \tag{14}$$

where m and j are integer numbers. They made experiments with $Y_{2,0}$, $Y_{3,1}$ and $Y_{1,-1}$ and concluded that the last one allows both a fast classification between chaotic and regular orbits, and a clear identification of stable and unstable periodic orbits. Due to this fact, we will use $Y \equiv Y_{1,-1}$ throughout the rest of this article. In order to reduce the fast oscillations that the time evolution of the MEGNO presents, in [Cincotta *et al.*, 2003] they used a time average of this quantity, namely:

$$\overline{Y}(n) = \frac{1}{n} \sum_{k=1}^{n} Y(k). \tag{15}$$

Theoretically, the asymptotic evolution of $\overline{Y}(n)$ for any dynamical regime can be put into a single expression:

$$\overline{Y}(n) \approx \frac{\text{mLCE}}{2} n + c, \tag{16}$$

where $c \approx 0$ and 2 for chaotic and regular motion, respectively.

We also compare our results with the SALI, introduced by Skokos [2001], that measures the degree in which a pair of initially linear independent deviation vectors tend to become aligned. The underlying principle is that, for a chaotic orbit, a deviation vector under the tangent map dynamics changes in order to become aligned with the instantaneous most unstable direction. In other words, for almost every pair of initial deviation vectors ($\mathbf{v}_0$, $\mathbf{u}_0$), the more chaotic the orbit the faster the angle between them will reduce to zero. In the case of regular orbits, the behavior depends strongly on the dimensionality of the map: in maps with dimension ≥ 4 the deviation vectors generally remain unaligned so the SALI tends to a positive nonzero value, while in 2D maps the two deviation vectors tend to align with a time rate that follows a power law. See [Skokos *et al.*, 2004, 2007] for numerical tests of the SALI and its generalization, the GALI family. Denoting the Euclidean norm with $\|\cdot\|$ the SALI is defined as:

$$\text{SALI}(n) = \min\left\{\left\|\frac{\mathbf{v_n}}{\|\mathbf{v_n}\|} + \frac{\mathbf{u_n}}{\|\mathbf{u_n}\|}\right\|, \left\|\frac{\mathbf{v_n}}{\|\mathbf{v_n}\|} - \frac{\mathbf{u_n}}{\|\mathbf{u_n}\|}\right\|\right\}. \tag{17}$$

To control the exponential increase of the norm of the vectors and avoid overflow problem, Skokos *et al.* [2004] have normalized them, at every time step, keeping their norm equal to 1. Defined like this, SALI$(n) \in [0, \sqrt{2}]$ and SALI $= 0$ if and only if the two normalized vectors have the same direction, being equal or opposite.

4. Error Behavior on Simple Maps

From the linear and exponential bounds on the errors it follows that the numerical orbit remains close to the exact one for a number of iterations proportional to $1/\epsilon$ in the regular case and to $\ln(1/\epsilon)$ in the chaotic one.

We consider two types of models where the error grows linearly and exponentially, respectively. The first one is the translation on the torus $\mathbb{T}^1$ defined by Eq. (4).

This is equivalent to the rotations on the unit circle defined by the map

$$R\mathbf{x} = \begin{pmatrix} \cos(2\pi\omega) & -\sin(2\pi\omega) \\ \sin(2\pi\omega) & \cos(2\pi\omega) \end{pmatrix} \begin{pmatrix} \cos(2\pi x) \\ \sin(2\pi x) \end{pmatrix}. \tag{18}$$

The correspondence between the sequences $x_n = M(x_{n-1})$ and $\mathbf{x}_n = R\mathbf{x}_{n-1}$ is evident after writing $\mathbf{x}_n = (\cos(2\pi x_n), \sin(2\pi x_n))$. In spite of the rigorous mathematical equivalence between the translations on the torus and the rotations on the unit circle, there is not such similitude in the numerical evaluation of these maps, as will be explained in the following paragraph.

For map (4) the divergence of orbits and the reversibility error have a linear growth, basically due to the fact that $\omega^* \neq \omega$, as shown theoretically for the global error in Eq. (5). However there are architectures and/or compilers in which the reversibility error may be zero for this simple map that only requires the computation of sums and modulus operations. Without knowing the precise implementation of these peculiar operations, it is impossible to establish *a priori* with which compiler and in which architecture we may obtain this result. We believe this anomaly is due to the peculiar arithmetic operations involved and that it does not occur for a generic map. To support this claim we have checked that the reversibility error never vanishes for the map (18) which involves multiplications and the evaluation of trigonometric functions.

In Fig. 1(a), we compare the divergence and reversibility error for the torus translation with $\omega = \sqrt{2} - 1$ and $x_0 = 0.7$. We can see that both errors approximately satisfy the same expression: in log–log scale it is close to a straight line with unitary slope. This implies that both quantities are linear functions of time. We have checked that the behavior is similar for almost every initial condition and

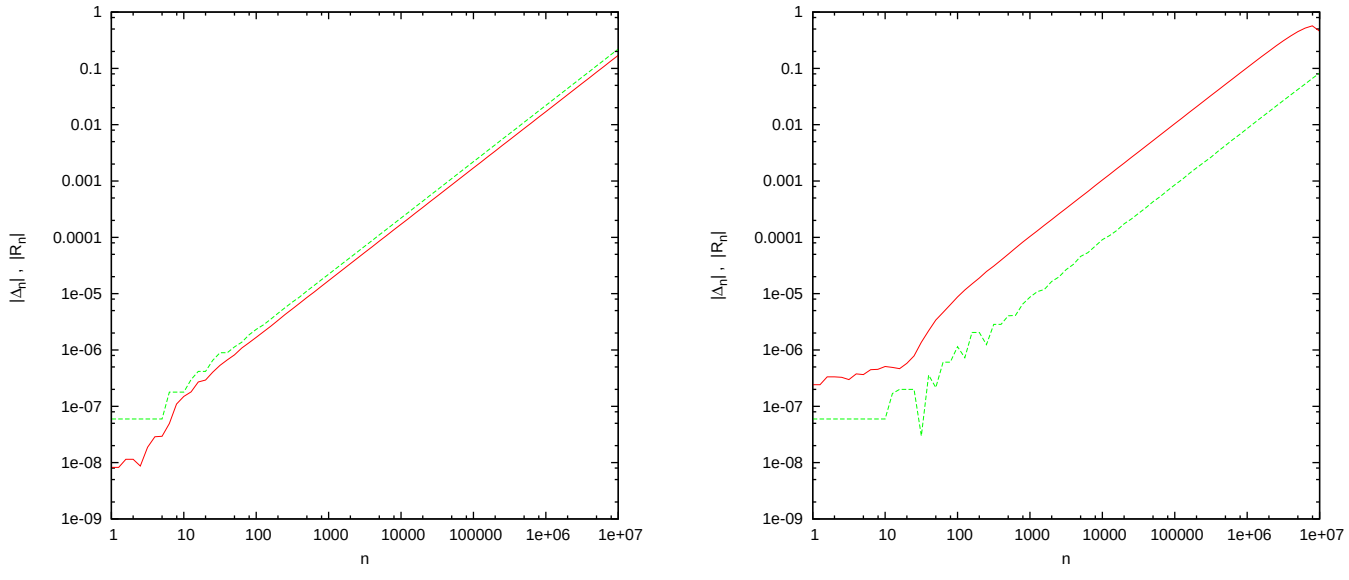


Fig. 1. Comparison between divergence (red) and reversibility (green) for both (a) torus translations and (b) rotations, using a frequency $\omega = \sqrt{2} - 1$ and initial condition $x_0 = 0.7$. In all four cases, the average slope in log–log scale is nearly one.

frequency but in some cases the fluctuations around the average linear growth have a larger amplitude. Figure 1(b) shows that for the map (18) there is an analogous behavior. However, in this case the straight line that corresponds to orbit divergence is a bit shifted upwards with respect to the one of reversibility errors. This implies that the orbit divergence's linear growth has a bigger slope.

The fluctuations w_n have variance $\langle w_n^2 \rangle$ computed with respect to an ensemble of initial conditions x_0 which first grow linearly, but rapidly saturates to a constant value with small oscillations (see Fig. 7 in [Turchetti *et al.*, 2010b]). In this respect, it is quite different with respect to a random perturbation $\epsilon\xi_n$ where ξ_n are independent variables, since in this case we have $w_n = \xi_1 + \cdots + \xi_n$ and the variance is given by $\langle w_n^2 \rangle = n\langle \xi^2 \rangle$ for any value of n. This different behavior is reflected in the decay of fidelity [Vaienti *et al.*, 2007] which follows a power law for round off errors and an exponential law for random perturbations [Turchetti *et al.*, 2010a, 2010b]. This shows that the round off errors decorrelate very slowly unlike the random errors which are uncorrelated.

For chaotic maps, characterized by an exponential increase in the distance between two nearby orbits, the reversibility error due to round off has the same growth. Bounded two-dimensional chaotic maps, like the cat map, exhibit this behavior during a short time before reaching saturation.

5. The Case of the Standard Map

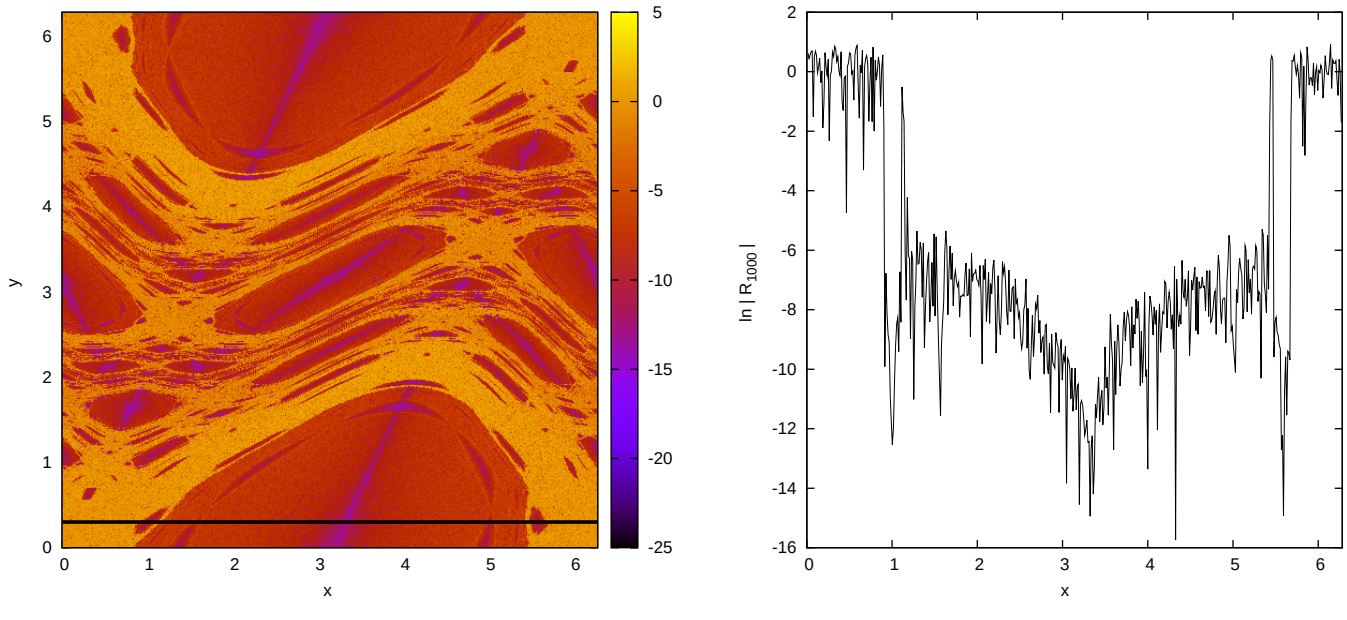
As an example of generic 2D map we have chosen the standard map in a torus:

$$\begin{aligned} y_{n+1} &= y_n + \lambda \sin(x_n) \mod 2\pi; \\ x_{n+1} &= x_n + y_{n+1} \mod 2\pi. \end{aligned} \tag{19}$$

For very low values of λ the divergence of orbits has an average growth linear with n as for the translations on the torus and they depend weakly on the initial condition. As λ is increased, a power law n^α with $\alpha > 1$ is observed and the dependence on the initial conditions becomes appreciable. Reversibility error also shows this behavior for $\lambda \ll 1$.

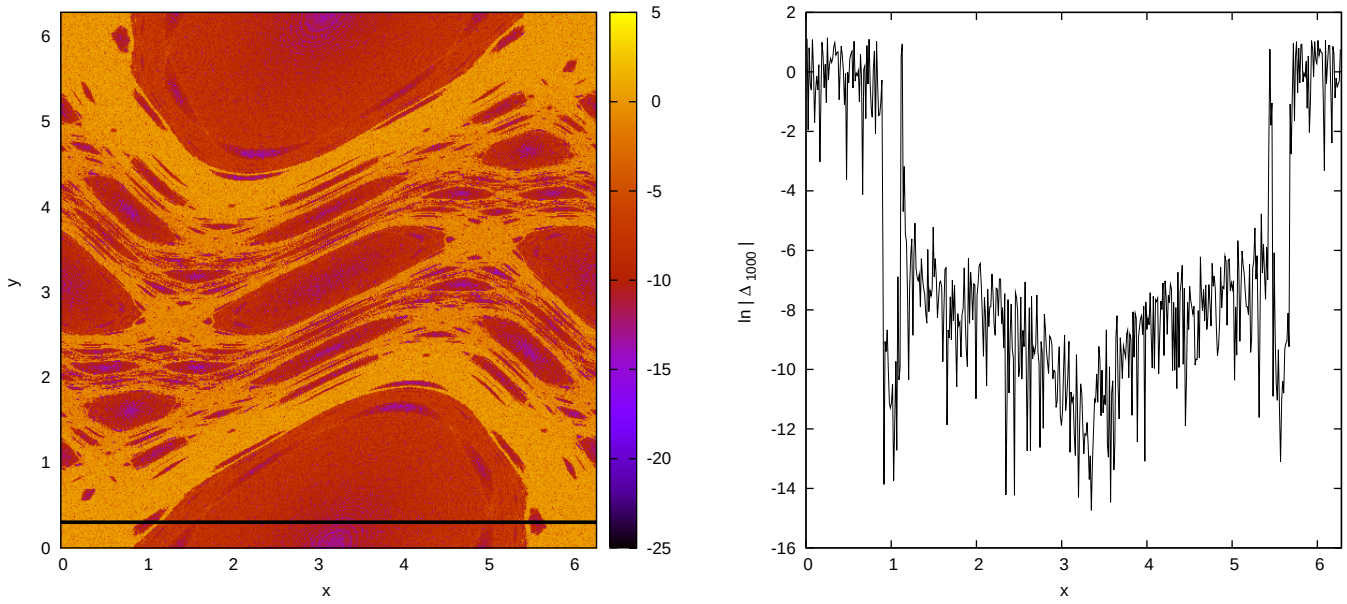
For λ approaching one, we see the coexistence of regular and chaotic regions so that the domain is split into several islands of stable orbits and a chaotic sea. In Figs. 2–6 we show, respectively, the value of the reversibility error (8), the orbit divergence (7), the finite time mLCE (12), the time average of the MEGNO (15) and the SALI (17) obtained iterating Eq. (19).

The standard map with $\lambda = 0.971635$ was iterated $n = 10^3$ times to compute all the dynamical



(a) (b)

Fig. 2. $\ln(R_{10^3})$ displaying the dynamical structure of the phase space of (a) the Standard Map and (b) for torus section $y = 0.3$.

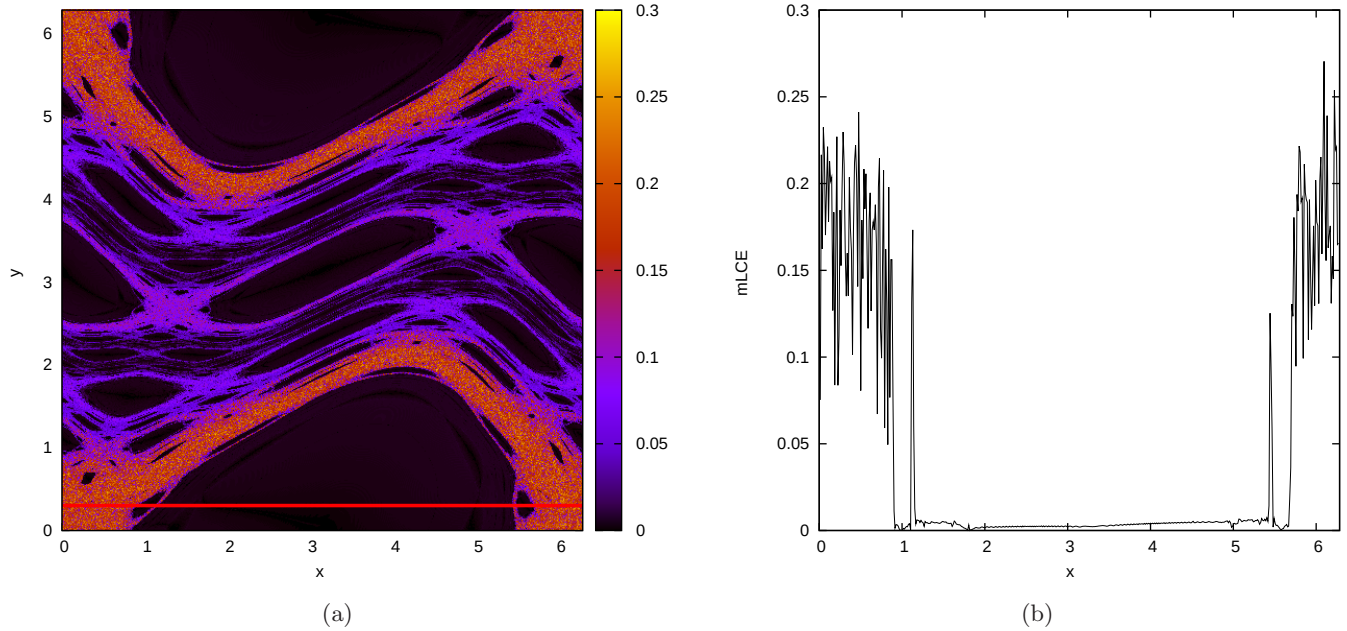


(a) (b)

Fig. 3. $\ln(\Delta_{10^3})$ displaying the dynamical structure of the phase space of (a) the Standard Map and (b) for torus section $y = 0.3$.

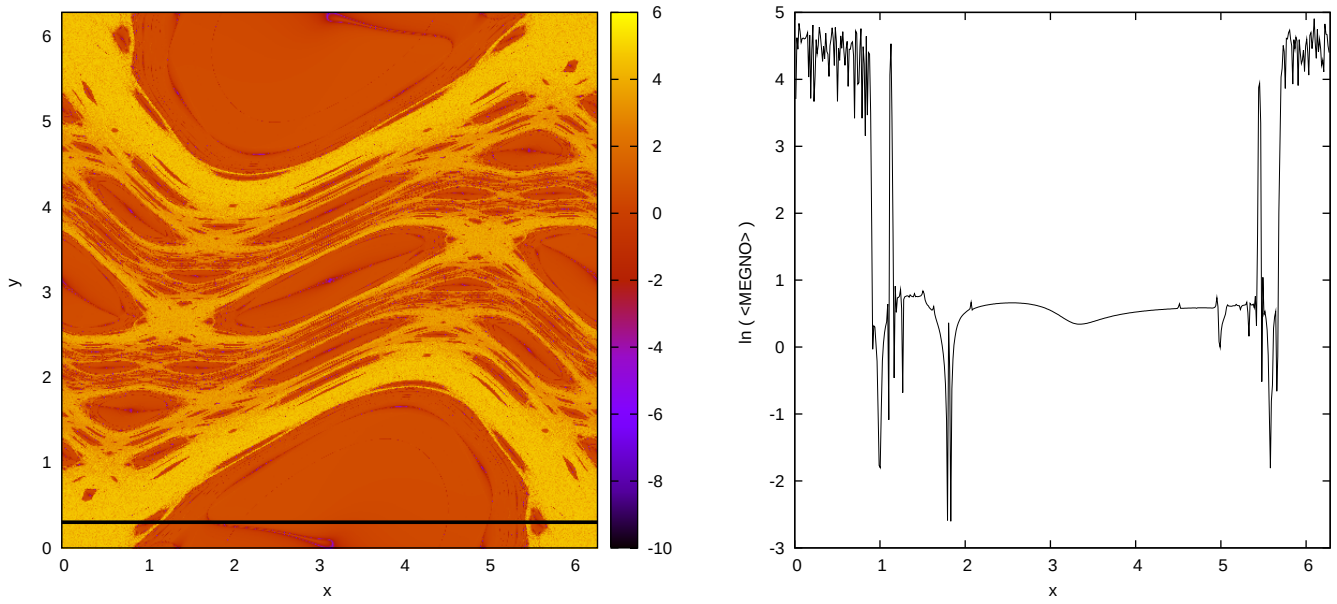
indicators. Even if $n < 10^3$ might be used to describe the chaotic region, the chosen value $n = 10^3$ allows to highlight differences within regular regions where the growth is very slow.

The pictures on the left side (a) show the value of the dynamical indicators in a chromatic scale for a grid of 500×500 initial conditions in the two-dimensional torus whereas the right pictures (b) show the corresponding graph for a fixed value of the action variable ($y = 0.3$) corresponding to the horizontal line in the left figures. Except for mLCE we have used the natural logarithm of the absolute value for all the dynamical indicators.



(a) (b)

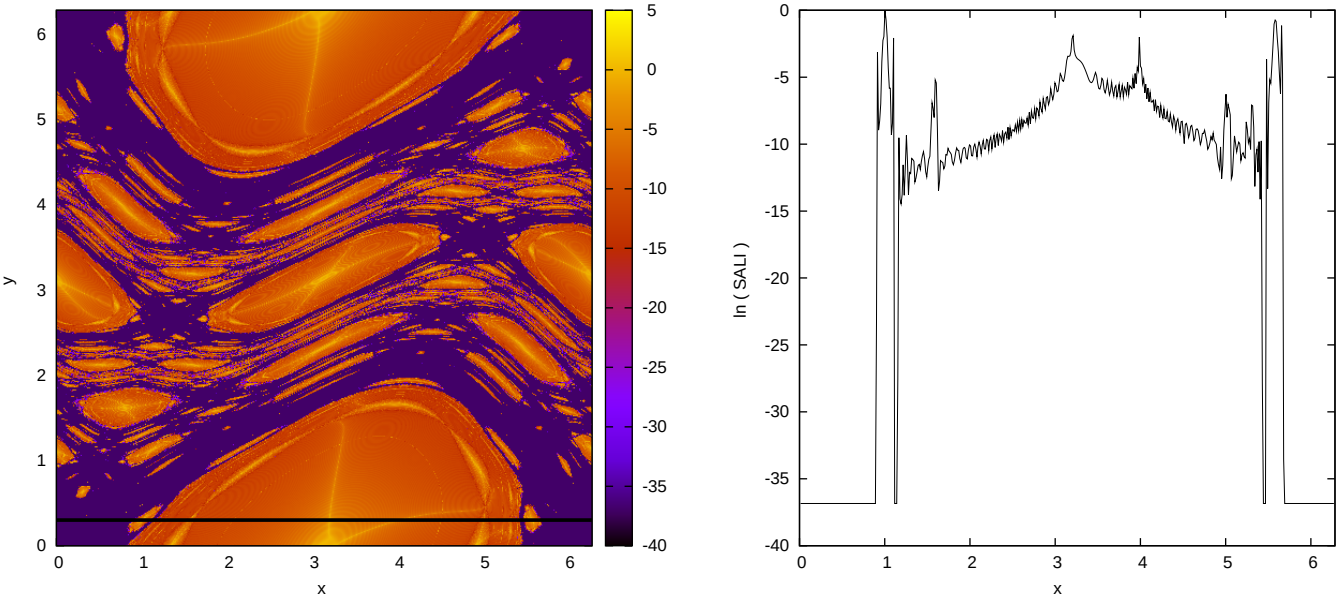
Fig. 4. mLCE(10^3) displaying the dynamical structure of the phase space of (a) the Standard Map and (b) for torus section $y = 0.3$.



(a) (b)

Fig. 5. $\ln(\overline{Y}(10^3))$ displaying the dynamical structure of the phase space of (a) the Standard Map and (b) for torus section $y = 0.3$.

Figures 2 and 3 were obtained by taking into account the error only in the action variable. It is evident that both the reversibility error and the orbits divergence (with respect to the exact one) discriminate regular from chaotic orbits. Minor differences exist within the islands of stability: the orbit divergence approaches the minimum value close to the center, whereas the minimum of the reversibility error occurs on a line crossing the islands possibly due to a mechanism similar to the one reported by Barrio *et al.* [2009] which analyzed spurious errors for variational methods.



(a) (b)

Fig. 6. $\ln(\mathrm{SALI}(10^3))$ displaying the dynamical structure of the phase space of (a) the Standard Map and (b) for torus section $y = 0.3$.

The plots obtained for mLCE and the MEGNO (Figs. 4 and 5, respectively) show no structure within the resonance islands whereas variations and fluctuations appear in the chaotic regions as for the previous indicators.

Figure 6 shows the logarithm of the SALI value. Due to the fact that this indicator converges to zero extremely fast for chaotic orbits, we have used a cut off value at SALI $= 10^{-16}$. The presence of this cut-off erases any structure within the strongly chaotic region.

To summarize, all the indicators discriminate regular and chaotic regions but their sensitivity in these regions is different.

Another aspect to consider when comparing the efficiency of chaos indicators is the computational cost. Each variational method needs to iterate both the map and the tangent map forward for n steps. The tangent map is the computationally most expensive since it needs the evaluation of the Jacobian matrix at every step. When computing SALI, two deviation vectors must be simultaneously computed. We also use two deviation vectors in evaluating MEGNO, selecting at every step the one that stretches more, in order to reduce the probability of having a vector almost orthogonal to the most unstable direction. In the case of the computation of the mLCE, we have used the orthogonalization algorithm developed by Benettin *et al.* [1980]. On the other hand, we notice that the reversibility error method requires only the iteration of the map whereas the orbit divergence method requires the iteration of the single and double precision (or double and higher precision) maps. As a consequence, the computationally most economic method is the one based on the reversibility error which does not require any algorithm except the evaluation of the map itself. Typically, the relevant information can be extracted from a computation in single precision.

From now on, within this section, we will focus on some ensemble quantities in order to compare the effect of round off errors with the effect of random perturbations in the standard map. For the standard map with $\lambda = 10^{-4}$ all the orbits are regular and we follow the evolution of an ensemble of 10 001 initial conditions randomly chosen in $(x, y) \in [1.5,\ 1.5 + 10^{-3}] \times [\pi,\ \pi + 10^{-3}]$. For each iteration we compute the variance of R_n in action $({\sigma_y}^2)$ and angle $({\sigma_x}^2)$ variables [see in Fig. 7(a)] and compare it against the same quantities of a double precision orbit stochastically perturbed with a uniform uncorrelated noise of amplitude 10^{-7} [shown in Fig. 7(b)]. In the latter case, we have found that ${\sigma_y}^2$ and ${\sigma_x}^2$ grow according to a power law with exponents one and three respectively up to numerical uncertainties.

It is interesting to compare the two pictures in Fig. 7 with each other. For n large we see that the behaviors of variances due to round off and random

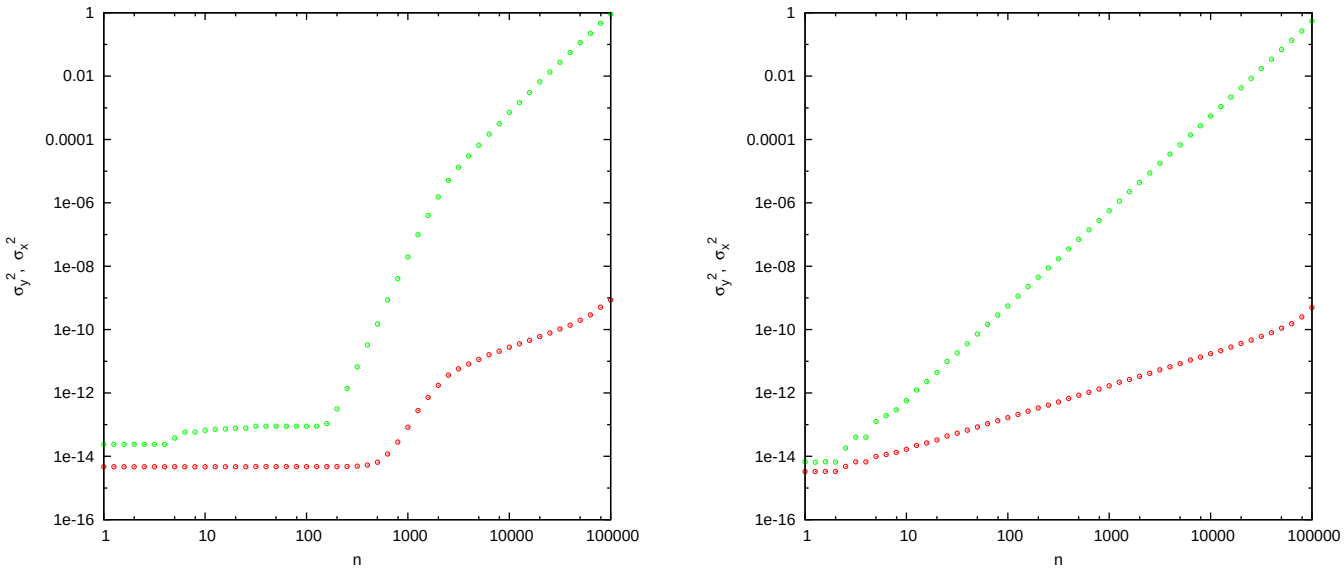
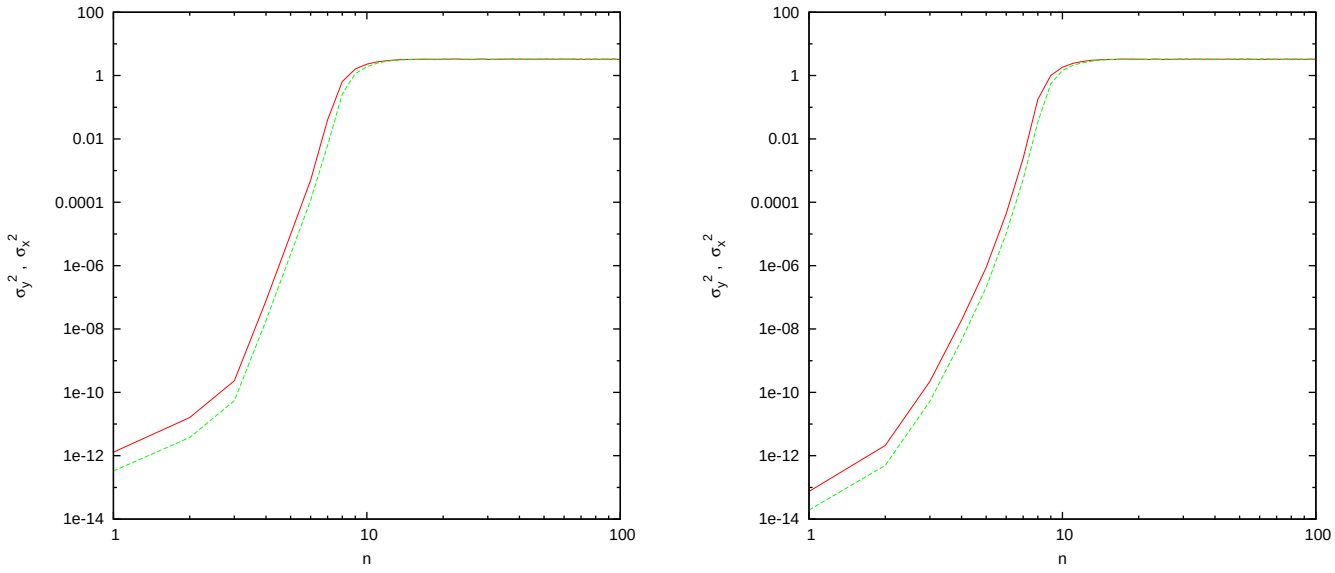


Fig. 7. Evolution of the variance of a distribution of reversibility errors in action (red) and angle (green) variables for the standard map with $\lambda = 10^{-4}$. (a) Round off error, (b) stochastic uncorrelated perturbations of amplitude 10^{-7}.



(a) (b)

Fig. 8. Evolution of the variance of a distribution of reversibility errors in action (red) and angle (green) variables for the standard map with $\lambda = 10$. There is evident equivalence between (a) round off error and (b) stochastic uncorrelated perturbations of amplitude 10^{-7}.

perturbations are similar. The presence of a transient for the reversibility case is very likely due to the initial presence of correlation.

In the fully chaotic regime for the standard map $\lambda \gg 1$ the random perturbation produces very similar results to the round off error as shown in Fig. 8 for $\lambda = 10$ since in the presence of chaotic dynamics the round off error correlations are lost.

The results for the standard map in the small λ regime can be compared to the variances for the skew map:

$$\begin{aligned} y_{n+1} &= y_n \qquad\quad \mod 1; \\ x_{n+1} &= x_n + y_{n+1} \quad \mod 1. \end{aligned} \tag{20}$$

to which the standard map in Eq. (19) reduces for $\lambda = 0$. If the skew map is randomly perturbed:

$$\begin{aligned} y_{n+1} &= y_n + \epsilon\xi_n \qquad\quad \mod 1; \\ x_{n+1} &= x_n + y_{n+1} + \epsilon\chi_n \quad \mod 1, \end{aligned} \tag{21}$$

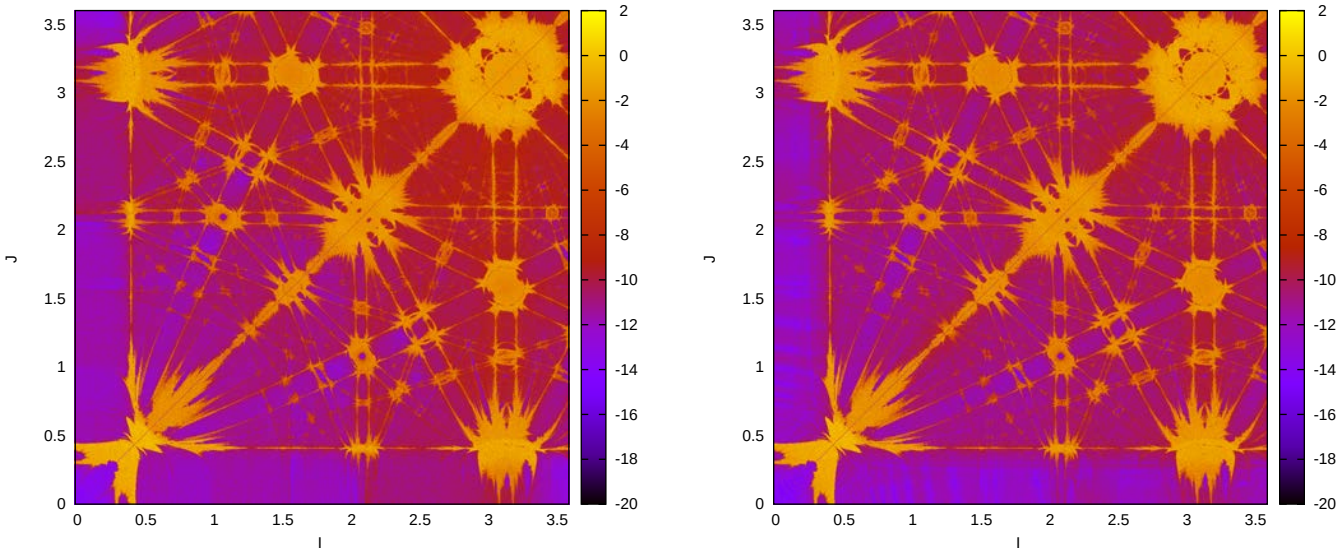
where ξ_n and χ_n are random uncorrelated variables, the growth of the variances σ_y^2 and σ_x^2 follows a linear and cubic law, respectively [Turchetti *et al.*, 2006]. The random perturbation of a standard map with a small value of λ [shown in Fig. 7(b)] shows exactly the same growth.

In the case of the round off error, the behavior of the skew map with respect to the standard map with a small value of λ is quite different. Indeed, the round off error affects the skew map just as the translation on the 1D Torus: it was observed that the global error grows linearly and the variance saturates at a very small value with respect to the size of the torus (see Fig. 7 in [Turchetti *et al.*, 2010a]). For the standard map, the coupling between action and angle, even for very small λ, causes a growth of the variance of the fluctuations due to round off as shown in Fig. 7(a). As a consequence, the effect of round off in a very weakly perturbed standard map is similar to a random perturbation.

6. A 4D Map

In this section, we show how either the reversibility error or the divergence of orbits can be used to analyze the resonance structure of four-dimensional nonintegrable maps. As an example, we consider a symplectic nearly integrable map extensively used in the literature (see [Guzzo & Lega, 2004]). This map is defined as:

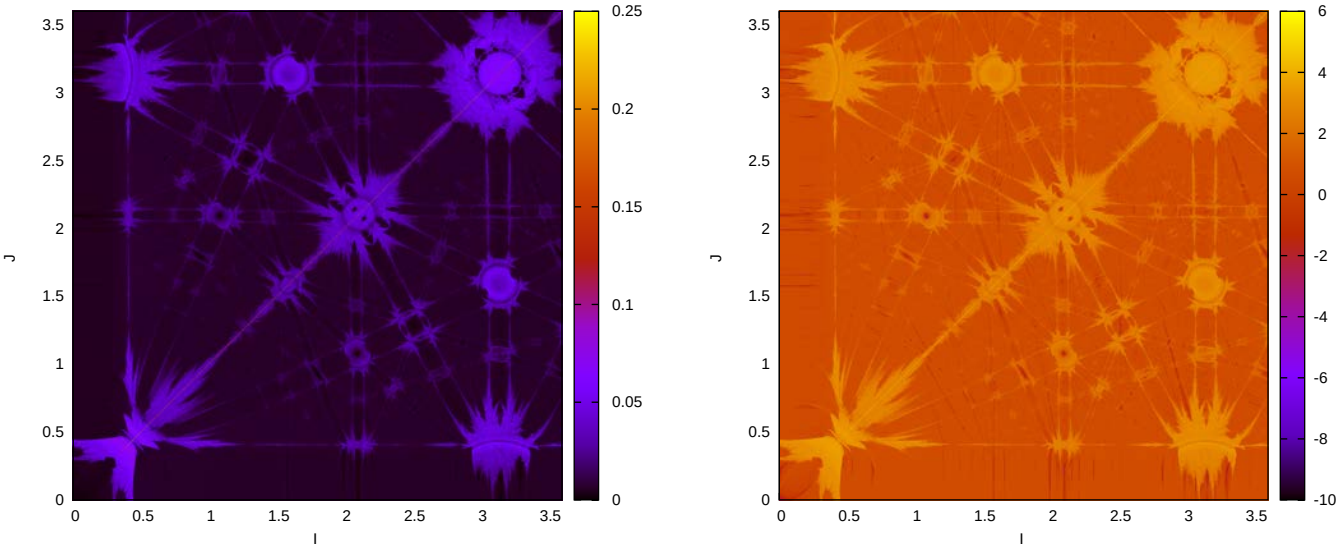
$$\begin{aligned} \theta_{n+1} &= \theta_n + I_n \\ \phi_{n+1} &= \phi_n + J_n \\ I_{n+1} &= I_n - \mu \frac{\partial V(\theta_{n+1}, \phi_{n+1})}{\partial \theta_{n+1}} \\ J_{n+1} &= J_n - \mu \frac{\partial V(\theta_{n+1}, \phi_{n+1})}{\partial \phi_{n+1}}; \end{aligned} \tag{22}$$



(a) (b)

Fig. 9. (a) $\ln(R_{10^3})$ and (b) $\ln(\Delta_{10^3})$ for map (22) using $c = 2$ and $\mu = 0.6$.

where $V \equiv 1/(\cos(\theta) + \cos(\phi) + 2 + c)$, with $c > 0$ and μ is the perturbation parameter. We have taken a grid of 1146×1146 initial conditions with actions $(I, J) \in [0,\ 3.6] \times [0,\ 3.6]$ and a fixed pair of angles, namely $(\theta, \phi) = (0.5,\ 0.5)$ and computed R_n, Δ_n, mLCE(n), $\overline{Y}(n)$ and SALI(n) for $n = 10^3$. Associating these values to each initial condition while using a chromatic scale, we have performed Figs. 9–11, where the parameter values $c = 2$ and $\mu = 0.6$ have been used. As we did in the previous section, in all except for the mLCE we have used the natural logarithm of the absolute value of the concomitant indicator and again we have used a cut off value at SALI $= 10^{-16}$. Figure 9 was done taking into account the Euclidean error in only the action plane. As it happened for the standard map, both reversibility error and orbit divergence present the same order of magnitude. The resonance web appears in every figure and its structure is the same. From a qualitative viewpoint, no substantial differences are found



(a) (b)

Fig. 10. (a) mLCE(10^3) and (b) $\ln(\overline{Y}(10^3))$ for map (22) using $c = 2$ and $\mu = 0.6$.

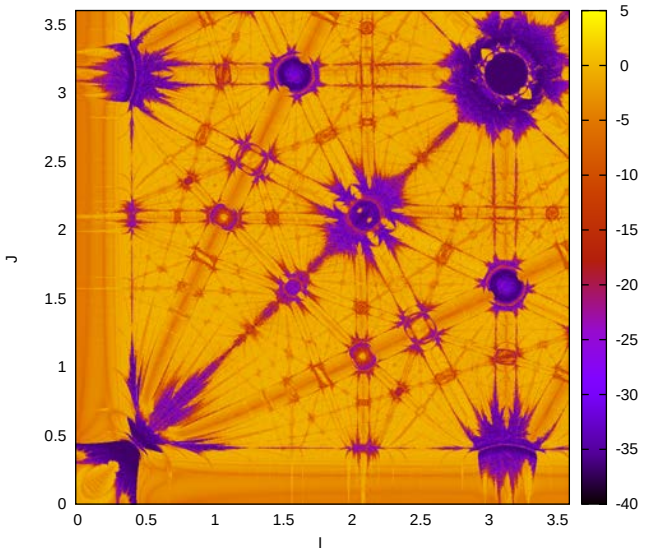


Fig. 11. SALI(10^3) for map (22) using $c = 2$ and $\mu = 0.6$.

and one might conclude that the dynamical information extracted from the reversibility error is the same as for the other four dynamical indicators we have considered. In the case of mLCE, the highest values are located on the diagonal $I = J$ even if they are not well visible in the plot. This is anyway coherent with the fact that on the diagonal we find the highest instability.

The frequency map analysis was not presented for comparison because it is computationally heavier. The error reaches its highest values in small neighborhoods of single resonance lines, because it detects the perturbed separatrices of the pendulum models that one could associate with the resonances, and in relatively large neighborhoods of the intersections of resonances, where the dynamics are widely nonintegrable, as shown by the computation of the interpolating Hamiltonian. These results are certainly not exhaustive but show that the behavior of the reversibility error is strictly related to the divergence of orbits and consequently it is very weak in the integrable regions where it does not have a diffusive character as for a random error.

7. Conclusions

We have examined the orbit divergence and the reversibility error in order to determine the effect of round off error for invertible maps. The knowledge of the exact map is not required to compute the reversibility error and the results obtained are about the same with respect to the case in which we compute the orbit divergence.

By choosing an ensemble of initial data we have examined the statistics of the fluctuations due to round off with respect to the time average of the error. There are different behaviors according to the degree of chaoticity that characterizes the ensemble. For chaotic orbits the variances have an exponential growth similar to the one observed for random perturbations and correspondingly the decay of fidelity is super-exponential in both cases. For regular or quasi-regular orbits differences are observed between the effects generated by round off errors and random perturbations. For a quasi-integrable map the random perturbations produce a growth of variances, linear and cubic, for actions and angles, respectively. On the other hand, the round off errors produce an initial transient that possibly lasts the time that the round off errors of the ensemble orbits need to decorrelate. After this initial transient, the behavior of variances affected by round off and noise are similar to each other. Finally, for a regular map such as the translation of the 1D torus, the round off variance, after a linear growth, saturates before the distribution of errors fills all the torus, unlike in the randomly perturbed case where the growth is always linear until saturation of the full torus. The fidelity has a power law decay for round off whereas it decays exponentially for random perturbations. This is a clear signature of the correlation between the errors due to round off.

To conclude, the reversibility error provides basically the same information as the divergence of orbits and it is easily accessible from a computational view point. This is due to the fact that it does not require the solution of the variational equations and that both the forward and backward orbits can be computed using single precision. When the initial point is varied and the number of iterations is kept fixed the reversibility error due to round off provides an insight into the dynamical structure of the map. For the standard map the reversibility error provides a picture of the dynamical behavior of the map where not only large scale features but also small scale details can be detected. In the case of a 4D symplectic map the reversibility error in action space provides a similar picture where the resonance web and the nearby regions of weakly chaotic motions can be easily highlighted.

Even if no really new information with respect to the standard indicators is provided by the reversibility error, we point out that this type of analysis takes into account not only the dynamics of the map but also the unavoidable effect of finite accuracy due to round off. Moreover, by increasing the accuracy of numerical computations, the time for which the reversibility test is computed can be increased and finer details on the phase space structure can be detected.

Acknowledgments

We are grateful to the anonymous reviewers for their important comments. The stay of M. Mestre at the Physics Department of the University of Bologna was fully supported by a grant from the Erasmus Mundus External Cooperation Window Lot 16 Programme, EADIC, financed by the European Commission. Besides, M. Mestre acknowledges financial support of a grant from the Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET). Also, he is grateful to A. Bazzani for his helpful discussions. D. Faranda acknowledges the financial support of the EU FP7-ERC project NAMASTE Thermodynamics of the Climate System.

References

Barrio, R., Borczyk, W. & Breiter, S. [2009] "Spurious structures in chaos indicators maps," *Chaos Solit. Fract.* **40**, 1697–1714.

Benettin, G., Galgani, L., Giorgilli, A. & Strelcyn, J. M. [1980] "Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory," *Meccanica* **15**, 9–20.

Buric, N., Rampioni, A. & Turchetti, G. [2005] "Statistics of Poincaré recurrences for a class of smooth circle maps," *Chaos Solit. Fract.* **23**, 1829–1840.

Chow, S. N. & Palmer, K. J. [1991] "On the numerical computation of orbits of dynamical systems: The one-dimensional case," *J. Dyn. Diff. Eqs.* **3**, 361–379.

Chow, S. N. & Palmer, K. J. [1992] "On the numerical computation of orbits of dynamical systems: The higher dimensional case," *J. Complexity* **8**, 398–423.

Cincotta, P. M., Giordano, C. M. & Simó, C. [2003] "Phase space structure of multi-dimensional systems by means of the mean exponential growth factor of nearby orbits," *Physica D* **182**, 151–178.

Gao, J. B. [1999] "Recurrence time statistics for chaotic systems and their applications," *Phys. Rev. Lett.* **83**, 3178–3181.

Goldberg, D. [1991] "What every computer scientist should know about floating point arithmetic," *ACM Comput. Surv.* **23**, 5–48.

Goździewski, K., Bois, E., Maciejewski, A. J. & Kiseleva-Eggleton, L. [2001] "Global dynamics of planetary systems with the MEGNO criterion," *Astron. Astrophys.* **378**, 569–586.

Guzzo, M. & Lega, E. [2004] "First numerical evidence of global Arnold diffusion in quasi-integrable systems," *Discr. Cont. Dyn. B* **5**, 687–698.

Hammel, S. M., Yorke, J. A. & Grebogi, C. [1987] "Do numerical orbits of chaotic dynamical processes represent true orbits?" *J. Complexity* **3**, 136–145.

Hu, H., Rampioni, A., Rossi, L., Turchetti, G. & Vaienti, S. [2004] "Statistics of Poincaré recurrences for maps with integrable and ergodic components," *Chaos* **14**, 160.

Kac, M. [1934] "On the notion of recurrence in discrete stochastic processes," *Proc. Natl. Acad. Sci. USA* **20**, 376–379.

Katok, A. & Hasselblatt, B. [1997] *Introduction to the Modern Theory of Dynamical Systems* (Cambridge Univ. Press).

Knuth, D. E. [1973] *The Art of Computer Programming*, Vol. 2.

Laskar, J. [1999] "Introduction to frequency map analysis," *Hamiltonian Systems with Three or More Degrees of Freedom*, pp. 134–150.

Maffione, N. P., Giordano, C. M. & Cincotta, P. M. [2011] "Testing a fast dynamical indicator: The MEGNO," *Int. J. Nonlin. Mech.* **46**, 23–34.

Robutel, P. & Laskar, J. [2001] "Frequency map and global dynamics in the solar system I: Short period dynamics of massless particles," *Icarus* **152**, 4–28.

Rosenstein, M. T., Collins, J. J. & De Luca, C. J. [1993] "A practical method for calculating largest Lyapunov exponents from small data sets," *Physica D* **65**, 117–134.

Skokos, C. [2001] "Alignment indices: A new, simple method for determining the ordered or chaotic nature of orbits," *J. Phys. A-Math. Gen.* **34**, 10029–10043.

Skokos, C., Antonopoulos, C., Bountis, T. & Vrahatis, M. [2002] "Smaller alignment index (SALI): Detecting order and chaos in conservative dynamical systems," *Proc. 4th GRACM Congress on Computational Mechanics* (Citeseer).

Skokos, C., Antonopoulos, C., Bountis, T. C. & Vrahatis, M. [2004] "Detecting order and chaos in Hamiltonian systems by the SALI method," *J. Phys. A-Math. Gen.* **37**, 6269.

Skokos, C., Bountis, T. C. & Antonopoulos, C. [2007] "Geometrical properties of local dynamics in Hamiltonian systems: The Generalized Alignment Index (GALI) method," *Physica D* **231**, 30–54.

Skokos, C. [2010] "The Lyapunov characteristic exponents and their computation," *Lecture Notes in Physics* (Springer-Verlag, Berlin), pp. 63–135.

Turchetti, G., Bassi, G., Bazzani, A., Giorgini, B. & Mais, H. [2006] "Hamiltonian dynamics with a weak noise and the echo effect for the rotator model," *J. Phys. A-Math. Gen.* **39**, 11417–11440.

Turchetti, G., Vaienti, S. & Zanlungo, F. [2010a] "Asymptotic distribution of global errors in the numerical computations of dynamical systems," *Physica A* **389**, 4994–5006.

Turchetti, G., Vaienti, S. & Zanlungo, F. [2010b] "Relaxation to the asymptotic distribution of global errors due to round off," *Europhys. Lett.* **89**, 40006.

Vaienti, S., Liverani, C. & Marie, P. [2007] "Random classical fidelity," *J. Stat. Phys.* **128**, 1079.

Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. [1985] "Determining Lyapunov exponents from a time series," *Physica D* **16**, 285–317.

Chapter 3

Article: Numerical Convergence of the Block-Maxima Approach to the Generalized Extreme Value Distribution

J Stat Phys
DOI 10.1007/s10955-011-0234-7

Numerical Convergence of the Block-Maxima Approach to the Generalized Extreme Value Distribution

Davide Faranda · Valerio Lucarini · Giorgio Turchetti · Sandro Vaienti

Received: 4 March 2011 / Accepted: 23 May 2011

Abstract In this paper we perform an analytical and numerical study of Extreme Value distributions in discrete dynamical systems. In this setting, recent works have shown how to get a statistics of extremes in agreement with the classical Extreme Value Theory. We pursue these investigations by giving analytical expressions of Extreme Value distribution parameters for maps that have an absolutely continuous invariant measure. We compare these analytical results with numerical experiments in which we study the convergence to limiting distributions using the so called block-maxima approach, pointing out in which cases we obtain robust estimation of parameters. In regular maps for which mixing properties do not hold, we show that the fitting procedure to the classical Extreme Value Distribution fails, as expected. However, we obtain an empirical distribution that can be explained starting from a different observable function for which Nicolis et al. (Phys. Rev. Lett. 97(21): 210602, 2006) have found analytical results.

Keywords Extreme values · Dynamical systems · EVT · GEV · Mixing · Logistic map · Chaos

D. Faranda (✉) · V. Lucarini
Department of Mathematics and Statistics, University of Reading, Whiteknights, PO Box 220, Reading RG6 6AX, UK
e-mail: d.faranda@pgr.reading.ac.uk

V. Lucarini
Department of Meteorology, Department of Mathematics, University of Reading, Whiteknights, PO Box 220, Reading RG6 6AX, UK
e-mail: v.lucarini@reading.ac.uk

G. Turchetti
Department of Physics, INFN-Bologna, University of Bologna, Via Irnerio 46, Bologna 40126, Italy
e-mail: turchett@bo.infn.it

S. Vaienti
UMR-6207, Centre de Physique Théorique, CNRS, Universités d'Aix-Marseille I, II, Université du Sud Toulon-Var and FRUMAM (Fédération de Recherche des Unités de Mathématiques de Marseille), CPT, Luminy, Case 907, 13288 Marseille Cedex 09, France
e-mail: vaienti@cpt.univ-mrs.fr

Published online: 27 September 2011

1 Introduction

Extreme Value Theory (EVT) was first developed by Fisher and Tippett [21] and formalized by Gnedenko [29] which showed that the distribution of the block-maxima of a sample of independent identically distributed (i.i.d.) variables converges to a member of the so-called Extreme Value (EV) distribution. It arises from the study of stochastical series that is of great interest in different disciplines: it has been applied to extreme floods [30, 50, 56], amounts of large insurance losses [6, 14]; extreme earthquakes [8, 13, 55]; meteorological and climate events [1, 19, 20, 48, 54, 60]. All these events have a relevant impact on socioeconomic activities and it is crucial to find a way to understand and, if possible, forecast them [34, 39].

The attention of the scientific community to the problem of modeling extreme values is growing. An extensive account of recent results and relevant applications is given in Ghil et al. [27]. Such an interest is mainly due to the fact that this theory is also important in defining risk factor in a wide class of applications such as the modeling of financial risk after the significant instabilities in financial markets worldwide [18, 28, 45], the analysis of seismic and hydrological risk [8, 47]. Even if the probability of extreme events decreases with their magnitude, the damage that they may bring increases rapidly with the magnitude as does the cost of protection against them Nicolis et al. [49].

From a theoretical point of view, extreme values represent extreme fluctuations of a system. Very recently, many authors have shown clearly how the statistics of global observables in correlated systems can be related to EV statistics [5, 15]. Clusel and Bertin [9] have shown how to connect fluctuations of global additive quantities, like total energy or magnetization, by statistics of sums of random variables in such a way that it is possible to identify a class of random variables whose sum follows an extreme value distributions.

The so called block-maxima approach is widely used in EVT since it represents a very natural way to look at extremes. It consists of dividing the data series of some observable into bins of equal length and selecting the maximum (or the minimum) value in each of them [11]. When dealing with climatological or financial data, since we usually have limited dataset, the main problem in applying EVT is related to the choice of a sufficiently large statistics of extremes provided that each bin contains a suitable number of observations. Therefore a smart balance between number of maxima and observations per bin is needed [19, 40–42].

Recently a number of alternative approaches have been studied. One consists in looking at exceedance over high thresholds rather than maxima over fixed time periods. While the idea of looking at extreme value problems from this point of view is very old, the development of a modern theory has started with Todorovic and Zelenhasic [57] that have proposed the so called Peaks Over Threshold approach. At the same time there was a mathematical development of procedures based on a certain number of extreme order statistics [36, 52] and the Generalized Pareto distribution for excesses over thresholds [16, 17, 53].

Since dynamical systems theory can be used to understand features of physical systems like climate and forecast financial behaviors, many authors have studied how to extend EVT to these field. When dealing with dynamical systems we have to know what kind of properties (i.e. stability, degree of mixing, correlations decay) are related to Gnedenko's hypotheses and also which observables we must consider in order to obtain an EV distribution. Furthermore, even if the convergence is achieved, we should evaluate how fast it is depending on all parameters and properties used. Empirical studies show that in some cases a dynamical observable obeys to the extreme value statistics even if the convergence is highly dependent on the kind of observable we choose [58–60]. For example, Balakrishnan et al. [3] and more recently Nicolis et al. [49] and Haiman [33] have shown that for regular orbits of dynamical systems we don't expect to find convergence to EV distribution.

The first rigorous mathematical approach to extreme value theory in dynamical systems goes back to the pioneer paper by P. Collet in 2001 [12]. Collet got the Gumbel Extreme Value Law (see below) for certain one-dimensional non-uniformly hyperbolic maps which admit an absolutely continuous invariant measure and exhibit exponential decay of correlations. Collet's approach used Young towers [61, 62] and his suggestion was successively applied to other systems. Before quoting them, we would like to point out that Collet was able to establish a few conditions (usually called D and D') and which have been introduced by Leadbetter [43] with the aim to associate to the stationary stochastic process given by the dynamical system, a new stationary independent sequence which enjoyed one of the classical three extreme value laws, and this law could be pulled back to the original dynamical sequence. Conditions D and D' require a sort of independence of the stochastic dynamical sequence in terms of uniform mixing condition on the distribution functions. Condition D was successively improved by Freitas and Freitas [22], in the sense that they introduced a new condition, called D_2, which is weaker than D and that could be checked directly by estimating the rate of decay of correlations for Hölder observables.[1] We notice that conditions D_2 and D' allow immediately to get Extreme Value Laws for absolutely continuous invariant measures for uniformly one-dimensional expanding dynamical systems: this is the case for instance of the 1-D maps with constant density studied in Sect. 3 below. Another interesting issue of Collet's paper was the choice of the observables g's whose values along the orbit of the dynamical systems constitute the sequence of events upon which we successively search for the partial maximum. Collet considered a function $g(\mathrm{dist}(x,\zeta))$ of the distance with respect to a given point ζ, with the aim that g achieves a global maximum at almost all points ζ in the phase space; for example $g(x) = -\log x$. Using a different g, Freitas and Freitas [23] were able to get the Weibull law for the family of quadratic maps with the Benedicks-Carleson parameters and for ζ taken as the critical point or the critical value, so improving the previous results by Collet who did not keep such values in his set of full measure.

The latter paper [23] strongly relies on condition D_2; this condition has also been invoked to establish the extreme value laws on towers which model dynamical systems with stable foliations (hyperbolic billiards, Lozi maps, Hénon diffeomorphisms, Lorenz maps and flows). This is the content of the paper by Gupta, Holland and Nicol [32]. We point out that the observable g was taken in one of three different classes g_1, g_2, g_3, see Sect. 2 below, each one being again a function of the distance with respect to a given point ζ. The choice of these particular forms for the g's is just to fit with the necessary and sufficient condition on the tail of the distribution function $F(u)$, see next section, in order to exist a non-degenerate limit distribution for the partial maxima [24, 37]. The paper Gupta et al. [32] also covers the easier case of uniformly hyperbolic diffeomorphisms, for instance the Arnold Cat map which we studied in Sect. 3.2.

Another major step in this field was achieved by establishing a connection between the extreme value laws and the statistics of first return and hitting times, see the papers by

[1] We briefly state here the two conditions, we defer to the next section for more details about the quantities introduced. If $X_n, n \geq 0$ is a stochastic process, we define $M_{j,l} \equiv \{X_j, X_{j+1}, \ldots, X_{j+l}\}$ and we put $M_{0,m} = M_m$. Moreover we set a_m and b_m two normalizing sequences and $u_m = x/a_m + b_m$, where x is a real number, cf. next section for the meaning of these variables. The condition $D_2(u_m)$ holds for the sequence X_m if for any integer l, t, m we have $|\nu(X_0 > u_m, M_{t,l} \leq u_m) - \nu(X_0 > u_m)\nu(M_{t,l} \leq u_m)| \leq \gamma(m,t)$, where $\gamma(m,t)$ is non-increasing in t for each m and $m\gamma(m,t_m) \to 0$ as $m \to \infty$ for some sequence $t_m = o(m)$, $t_m \to \infty$.

We say condition $D'(u_m)$ holds for the sequence X_m if $\lim_{k\to\infty} \limsup_m m \sum_{j=1}^{[m/k]} \nu(X_0 > u_m, X_j > u_m) = 0$. Whenever the process is given by the iteration of a dynamical systems, the previous two conditions could also be formulated in terms of decay of correlation integrals, see Freitas and Freitas [22], Gupta [31].

Freitas, Freitas and Todd [24, 26]. They showed in particular that for dynamical systems preserving an absolutely continuous invariant measure or a singular continuous invariant measure ν, the existence of an exponential hitting time statistics on balls around ν almost any point ζ implies the existence of extreme value laws for one of the observables of type g_i, $i = 1, 2, 3$ described above. The converse is also true, namely if we have an extreme value law which applies to the observables of type g_i, $i = 1, 2, 3$ achieving a maximum at ζ, then we have exponential hitting time statistics to balls with center ζ. Recently these results have been generalized to local returns around balls centered at periodic points [25]. We would like to point out that the equivalence between extreme values laws and the hitting time statistics allowed to prove the former for broad classes of systems for which the statistics of recurrence were known, for instance for expanding maps in higher dimension.

In this work we consider a few aspects of the extreme value theory applied to dynamical systems throughout both analytical results and numerical experiments. In particular we analyze the convergence to EV limiting distributions pointing out how robust are parameters estimations. Furthermore, we check the consistency of block-maxima approach highlighting deviations from theoretical expected behavior depending on the number of maxima and number of block-observation. To perform our analysis we use low dimensional maps with different properties: mixing maps in which we expect to find convergence to EV distributions and regular maps where the convergence is not ensured.

The work is organized as follows: in Sect. 2 we briefly recall methods and results of EVT for independent and identical distributed (i.i.d.) variables and dynamical systems. In Sect. 3 we explicitly compute theoretical expected distributions parameter in respect to the observable functions of type g_i, $i = 1, 2, 3$ for map that have constant density measure. Numerical experiments on low dimensional maps are presented. In Sect. 4 we show that it is possible to derive an asymptotic expression of normalizing sequences when the density measure is not constant. As an example we derive the explicit expressions for the Logistic map. Eventually, in Sect. 5 we repeat the experiment for regular maps showing that extreme values laws do not follow from numerical experiments.

2 Background on EVT

Gnedenko [29] studied the convergence of maxima of i.i.d. variables

$$X_0, X_1, \ldots, X_{m-1}$$

with cumulative distribution (cdf) $F(x)$ of the form:

$$F(x) = P\{a_m(M_m - b_m) \leq x\}$$

Where a_m and b_m are normalizing sequences and $M_m = \max\{X_0, X_1, \ldots, X_{m-1}\}$. It may be rewritten as $F(u_m) = P\{M_m \leq u_m\}$ where $u_m = x/a_m + b_m$. Such types of normalizing sequences converge to one of the three type of Extreme Value (EV) distribution if necessary and sufficient conditions on parent distribution of X_i variables are satisfied [43]. EV distributions include the following three families:

- Gumbel distribution (type 1):

$$F(x) = \exp\{-e^{-x}\} \quad x \in \mathbb{R} \tag{1}$$

- Fréchet distribution (type 2):

$$\begin{cases} F(x) = 0 & x \leq 0 \\ F(x) = \exp\{-x^{1/\xi}\} & x > 0 \end{cases} \tag{2}$$

- Weibull distribution (type 3):

$$\begin{cases} F(x) = \exp\{-(-x)^{1/\xi}\} & x < 0 \\ F(x) = 0 & x \geq 0 \end{cases} \tag{3}$$

Let us define the right endpoint x_F of a distribution function $F(x)$ as:

$$x_F = \sup\{x : F(x) < 1\} \tag{4}$$

then, it is possible to compute normalizing sequences a_m and b_m using the following corollary of Gnedenko's theorem:

Corollary (Gnedenko) *The normalizing sequences a_m and b_m in the convergence of normalized maxima $P\{a_m(M_m - b_m) \leq x\} \to F(x)$ may be taken (in order of increasing complexity) as*:

- *Type* 1: $a_m = [G(\gamma_m)]^{-1}$, $b_m = \gamma_m$;
- *Type* 2: $a_m = \gamma_m^{-1}$, $b_m = 0$ *or* $b_m = c \cdot m^{-\xi}$;
- *Type* 3: $a_m = (x_F - \gamma_m)^{-1}$, $b_m = x_F$;

where

$$\gamma_m = F^{-1}(1 - 1/m) = \inf\{x;\, F(x) \geq 1 - 1/m\} \tag{5}$$

$$G(t) = \int_t^{x_F} \frac{1 - F(u)}{1 - F(t)}\, du \quad t < x_F \tag{6}$$

and $c \in \mathbb{R}$ is a constant. It is important to remark that the choice of normalizing sequences is not unique [43]. *For example for b_m of type* 2 *distribution it is possible to choose either $b_m = 0$ or $b_m = c \cdot m^{-\xi}$. In particular, we will use the last one since it is a more general choice that ensure the convergence for a much broader class of initial distributions* [4].

Instead of Gnedenko's approach it is possible to fit unnormalized data directly to a single family of generalized distribution called GEV distribution with cdf:

$$F_G(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{7}$$

which holds for $1 + \xi(x - \mu)/\sigma > 0$, using $\mu \in \mathbb{R}$ (location parameter) and $\sigma > 0$ (scale parameter) as scaling constants in place of b_m, and a_m [51]. $\xi \in \mathbb{R}$ is the shape parameter also called the tail index: when $\xi \to 0$, the distribution corresponds to a Gumbel type. When the index is negative, it corresponds to a Weibull; when the index is positive, it corresponds to a Fréchet.

In order to adapt the extreme value theory to dynamical systems, we will consider the stationary stochastic process $X_0, X_1, \ldots$ given by:

$$X_m(x) = g(\text{dist}(f^m(x), \zeta)) \quad \forall m \in \mathbb{N} \tag{8}$$

where 'dist' is a Riemannian metric on Ω, ζ is a given point and g is an observable function, and whose partial maximum is defined as:

$$M_m = \max\{X_0, \ldots, X_{m-1}\} \tag{9}$$

The probability measure will be here an invariant measure ν for the dynamical system. As we anticipated in the Introduction, we will use three types of observables g_i, $i = 1, 2, 3$, suitable to obtain one of the three types of EV distribution for normalized maxima:

$$g_1(x) = -\log(\text{dist}(x, \zeta)) \tag{10}$$

$$g_2(x) = \text{dist}(x, \zeta)^{-1/\alpha} \tag{11}$$

$$g_3(x) = C - \text{dist}(x, \zeta)^{1/\alpha} \tag{12}$$

where C is a constant and $\alpha > 0 \in \mathbb{R}$.

These three type of functions are representative of broader classes which are defined, for instance, throughout (1.11) to (1.13) in Freitas et al. [24]; we now explain the reasons and the meaning of these choices. First of all these functions have in common the following properties: (i) they are defined on the positive semi-axis $[0, \infty]$ with values into $\mathbb{R} \cup \{+\infty\}$; (ii) 0 is a global maximum, possibly equal to $+\infty$; (iii) they are a strictly decreasing bijection in a neighborhood V of 0 with image W. Then we consider three types of behavior which generalize the previous specific choices:

Type 1: there is a strictly positive function $p : W \to \mathbb{R}$ such that $\forall y \in \mathbb{R}$ we have

$$\lim_{s \to g_1(0)} \frac{g_1^{-1}(s + yp(s))}{g_1^{-1}(s)} = e^{-y}$$

Type 2: $g_2(0) = +\infty$ and there exists $\beta > 0$ such that $\forall y > 0$ we have

$$\lim_{s \to \infty} \frac{g_2^{-1}(sy)}{g_2^{-1}(s)} = y^{-\beta}$$

Type 3: $g_3(0) = D < +\infty$ and there exists $\gamma > 0$ such that $\forall y > 0$ we have

$$\lim_{s \to 0} \frac{g_3^{-1}(D - sy)}{g_3^{-1}(D - s)} = y^{\gamma}$$

The Gnedenko corollary says that the different kinds of extreme value laws are determined by the distribution of $F(u) = \nu(X_0 \leq u)$ and by the right endpoint of F, x_F. We will see in the next section that the local invertibility of g_i, $i = 1, 2, 3$ in the neighborhood of 0 together with the Lebesgue's differentiation theorem (which basically says that whenever the measure ν is absolutely continuous with respect to Lebesgue with density ρ, the measure of a ball $B_\delta(x_0)$ of radius δ centered around almost any point x_0 scales like $\delta\rho(x_0)$), allow us to compute the tail of F, in fact we have

$$1 - F(u) \sim \rho(\zeta)|B_{g^{-1}(u)}(\zeta)|,$$

where g is any of the three types of functions introduced in (10) to (12) and $|A|$ denotes the diameter of the set A. As we said above the tail of F determines the three limit laws for partial maximum of i.i.d. sequences. In particular Th. 1.6.2 in Leadbetter et al. [43] specifies what kind of conditions the distribution function F must verify to get one specific law: the above type 1, 2, 3 assumptions are just the translation in terms of the shape of g_i of the conditions on the tail of F.

3 Distribution of Extremes in Mixing Maps with Constant Density Measure

Our goal is to use a block-maxima approach and fit our unnormalized data to a GEV distribution; for that it will be necessary to find a linkage among a_m, b_m, μ and σ. At this regard we will use Gnedenko's corollary to compute normalizing sequences showing that they correspond to the parameter we obtain fitting directly data to GEV distribution.

We derive the correct expression for mixing maps with constant density measure and the asymptotic behavior for logistic map that is a case of non-constant density measure.

3.1 Asymptotic Sequences

In this section we will consider the case of uniformly hyperbolic maps which preserve the Lebesgue measure (the density $\rho = 1$) and satisfy the conditions D_2 and D', sufficient to get extreme valuers distributions. For the second map, the algebraic automorphisms of the torus better known as the Arnold cat map, the existence of extreme value laws follows from the theory developed in Gupta et al. [32]. Starting from the definitions provided by Gnedenko we derive as a novel result the exact expression for the normalizing sequences a_m and b_m.

Case 1 $(g_1(x) = -\log(\mathrm{dist}(x, \zeta)))$ By (8) and (9) we know that:

$$\begin{aligned} 1 - F(u) &= 1 - \nu(g(\mathrm{dist}(x, \zeta)) \leq u) \\ &= 1 - \nu(-\log(\mathrm{dist}(x, \zeta)) \leq u) \\ &= 1 - \nu(\mathrm{dist}(x, \zeta) \geq e^{-u}) \end{aligned} \tag{13}$$

and the last line is justified by using Lebesgue's Differentiation Theorem. Then, for maps with constant density measure, we can write:

$$1 - F(u) \simeq \nu(B_{e^{-u}}(\zeta)) = \Omega_d e^{-ud} \tag{14}$$

where d is the dimension of the space and Ω_d is a constant. To use Gnedenko corollary it is necessary to calculate u_F

$$u_F = \sup\{u; F(u) < 1\}$$

in this case $u_F = +\infty$.

Using Gnedenko (6) we can calculate $G(t)$ as follows:

$$G(t) = \int_t^\infty \frac{1 - F(u)}{1 - F(t)}\,\mathrm{d}u = \int_t^\infty \frac{e^{-ud}}{e^{-td}}\,\mathrm{d}u = \frac{1}{d}\int_{td}^\infty \frac{e^{-v}}{e^{-td}}\,\mathrm{d}v = \frac{1}{d} \tag{15}$$

According to the Leadbetter et al. [43] proof of Gnedenko theorem we can study both a_m and b_m or γ_m convergence as:

$$\begin{aligned} \lim_{m\to\infty} m(1 - F\{\gamma_m + xG(\gamma_m)\}) &= e^{-x} \\ \lim_{m\to\infty} m\Omega_d e^{-d(\gamma_m + xG(\gamma_m))} &= e^{-x} \end{aligned} \tag{16}$$

then we can use the connection between γ_m and normalizing sequences to find a_m and b_m.

By (5) or using relation (16):

$$\gamma_m \simeq \frac{\ln(m\Omega_d)}{d}$$

so that:

$$a_m = d \qquad b_m = \frac{1}{d}\ln(m) + \frac{\ln(\Omega_d)}{d}$$

Case 2 $(g_2(x) = \mathrm{dist}(x,\zeta)^{-1/\alpha})$ We can proceed as for g_1:

$$\begin{aligned} 1 - F(u) &= 1 - \nu(\mathrm{dist}(x,\zeta)^{-1/\alpha} \le u) \\ &= 1 - \nu(\mathrm{dist}(x,\zeta) \ge u^{-\alpha}) \\ &= \nu(B_{u^{-\alpha}}(\zeta)) = \Omega_d u^{-\alpha d} \end{aligned} \tag{17}$$

in this case $u_F = +\infty$.

$$\gamma_m = F^{-1}(1 - 1/m) = (m\Omega_d)^{1/(\alpha d)} \tag{18}$$

and, as discussed in Sect. 2, using Beirlant [4] choice of normalizing sequences we expect:

$$b_m = c \cdot m^{-\xi}$$

where $c \in \mathbb{R}$ is a constant.

Case 3 $(g_3(x) = C - \mathrm{dist}(x,\zeta)^{1/\alpha})$ Eventually we compute a_m and b_m for the g_3 observable class:

$$\begin{aligned} 1 - F(u) &= 1 - \nu(C - \mathrm{dist}(x,\zeta)^{1/\alpha} \le u) \\ &= 1 - \nu(\mathrm{dist}(x,\zeta) \ge (C-u)^{\alpha}) \\ &= \nu(B_{(C-u)^{\alpha}}(\zeta)) = \Omega_d (C-u)^{\alpha d} \end{aligned} \tag{19}$$

in this case $u_F = C$.

$$\gamma_m = F^{-1}(1 - 1/m) = C - (m\Omega_d)^{-1/(\alpha d)} \tag{20}$$

For type 3 distribution:

$$a_m = (u_F - \gamma_m)^{-1} \qquad b_m = u_F \tag{21}$$

3.2 Numerical Experiments

Since we want to show that unnormalized data may be fitted by using the GEV distribution $F_G(x;\mu,\sigma,\xi)$ we expect to find the following equivalence:

$$a_m = 1/\sigma \qquad b_m = \mu$$

where, clearly, $\mu = \mu(m)$ and $\sigma = \sigma(m)$. This fact can be seen as a linear change of variable: the variable $y = a_m(x - b_m)$ has a GEV distribution $F_G(y;\mu = 0,\sigma = 1,\xi)$ (that is an EV one parameter distribution with a_m and b_m normalizing sequences) while x is GEV distributed $F_G(x;\mu = b_m,\sigma = 1/a_m,\xi)$.

As we said above we now apply the previous considerations to two maps which enjoy extreme values laws and have constant density: we summarize below the theoretical results we obtained for all three type of observables. We have obtained the results in terms of m but, since we fix $k = n \cdot m$, the previous results can be translated in terms of n as follows:

For g_1 type observable:

$$\sigma = \frac{1}{d} \qquad \mu \propto \frac{1}{d}\ln\left(\frac{k}{n}\right) \tag{22}$$

For g_2 type observable:

$$\sigma \propto n^{-1/(\alpha d)} \qquad \mu \propto n^{-1/(\alpha d)} \tag{23}$$

For g_3 type observable:

$$\sigma \propto n^{1/(\alpha d)} \qquad \mu = C \tag{24}$$

Following Freitas et al. [24] we obtain the expression for the shape parameters: $\xi = 0$ for g_1 type, $\xi = 1/(\alpha d)$ for g_2 type and $\xi = -1/(\alpha d)$ for g_3 type.

In order to provide a numerical test of our results we consider a one-dimensional and a two dimensional map. The one dimensional map used is a Bernoulli Shift map:

$$x_{t+1} = q x_t \mod 1 \quad q > 1 \in \mathbb{N} \tag{25}$$

with $q = 3$.

The considered two dimensional map is the famous Arnold's cat map defined on the 2-torus by:

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} \mod 1 \tag{26}$$

A wide description of properties of these maps can be found in Arnold and Avez [2] and Hasselblatt and Katok [35].

We proceed as follows. For each map we run a long simulation up to k iterations starting from a given initial condition ζ. Note that the results—as we tested—do not depend on the choice of ζ. From the trajectory we compute the sequence of observables g_1, g_2, g_3

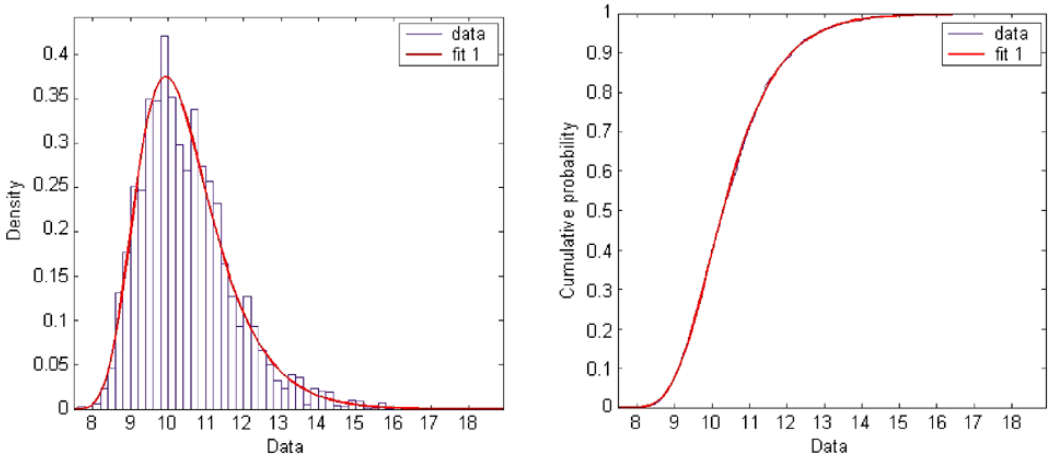


Fig. 1 *Left*: g_1 observable empirical histogram and fitted GEV pdf. *Right*: g_1 observable empirical cdf and fitted GEV cdf. Logistic map, $n = 10^4$, $m = 10^4$

as follows dividing it into n bins each containing $m = k/n$ observations. Then, we test the degree of agreement between the empirical distribution of the maxima and the GEV distribution according to the theoretical values presented above. A priori, it is reasonable to assume GEV as a suitable family of statistical models. For some selected values of n, the maxima are normalized and fitted to GEV distributions $F_G(x; \mu, \sigma, \xi)$ using a maximum likelihood method which selects values of the model parameters that produce the distribution most likely to have resulted in the observed data.

All the numerical analysis contained in this work has been performed using MATLAB Statistics Toolbox functions such as *gevfit* and *gevcdf*. These functions return maximum likelihood estimates of the parameters for the generalized extreme value (GEV) distribution giving 95% confidence intervals for estimates [46].

As in every fitting procedure, it is necessary to test the a posteriori goodness of fit. We anticipate that in every case considered, fitted distributions passed, with maximum confidence interval, the Kolmogorov-Smirnov test described in Lilliefors [44]. For illustration purposes, we present in Fig. 1 an empirical pdf and cdf with the corresponding fits.

Once k is set to a given value (in our case $k = 10^7$), the numerical simulations allow us to explore two limiting cases of great interest in applications where the statistical inference is intrinsically problematic:

1. n is small (m is large), so that we extract only few maxima, each corresponding to a *very* extreme event.
2. m is small (n is large), so that we extract many maxima but most of those will not be as extreme as in case (1).

In case (1), we have only few data—of high quality—to fit our statistical models whereas in case (2) we have many data but the sampling may be spoiled by the inclusion of data not giving a good representation of extreme events. We have in general that in order to obtain a reliable fit for a distribution with p parameters we need 10^p independent data [19] so that we expect that fit procedure gives reliable results for $n > 10^3$. As the value of m determines to which extend the extracted bin maximum is representative of an extreme, below a certain value $m_{\min}$ our selection procedure will be unavoidably misleading. We have no obvious theoretical argument to define the value of $m_{\min}$. We expect to obtain good fits throughout

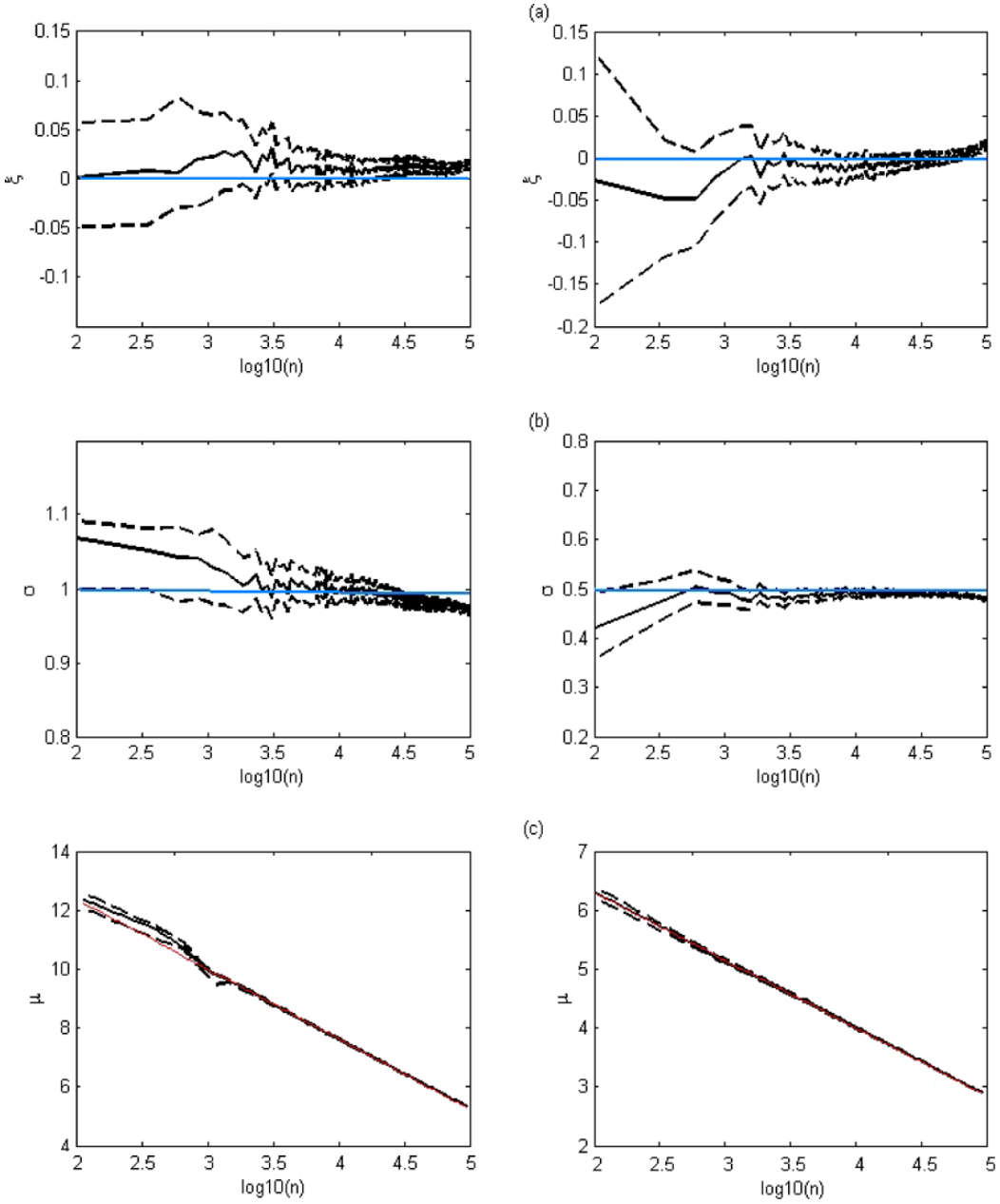


Fig. 2 (Color online) g_1 observable, $\zeta \simeq 0.51$. (**a**) ξ VS $\log_{10}(n)$; (**b**) σ VS $\log_{10}(n)$; (**c**) μ VS $\log_{10}(n)$. *Left*: Bernoulli Shift map. *Right*: Arnold Cat Map. *Dotted lines* represent computed confidence interval, *red lines* represent a linear fit, *blue lines* are theoretical values

the parametric region where the constraints on n, m are satisfied. Therefore, our flexibility in choosing satisfying pairs (n, m) increases with larger values of k.

For a g_1 type observable function the behavior against n of the three parameters is presented in Fig. 2. According to (22) we expect to find $\xi = 0$. For relatively small values of n the sample is too small to ensure a good convergence to analytical ξ and confidence intervals are wide. On the other hand we see deviations from expected value as $m < 10^3$ that is when $n > 10^4$. For the scale parameter a similar behavior is achieved and deviations from expected theoretical values $\sigma = 1/2$ for Arnold Cat Map and $\sigma = 1$ for Bernoulli Shift are

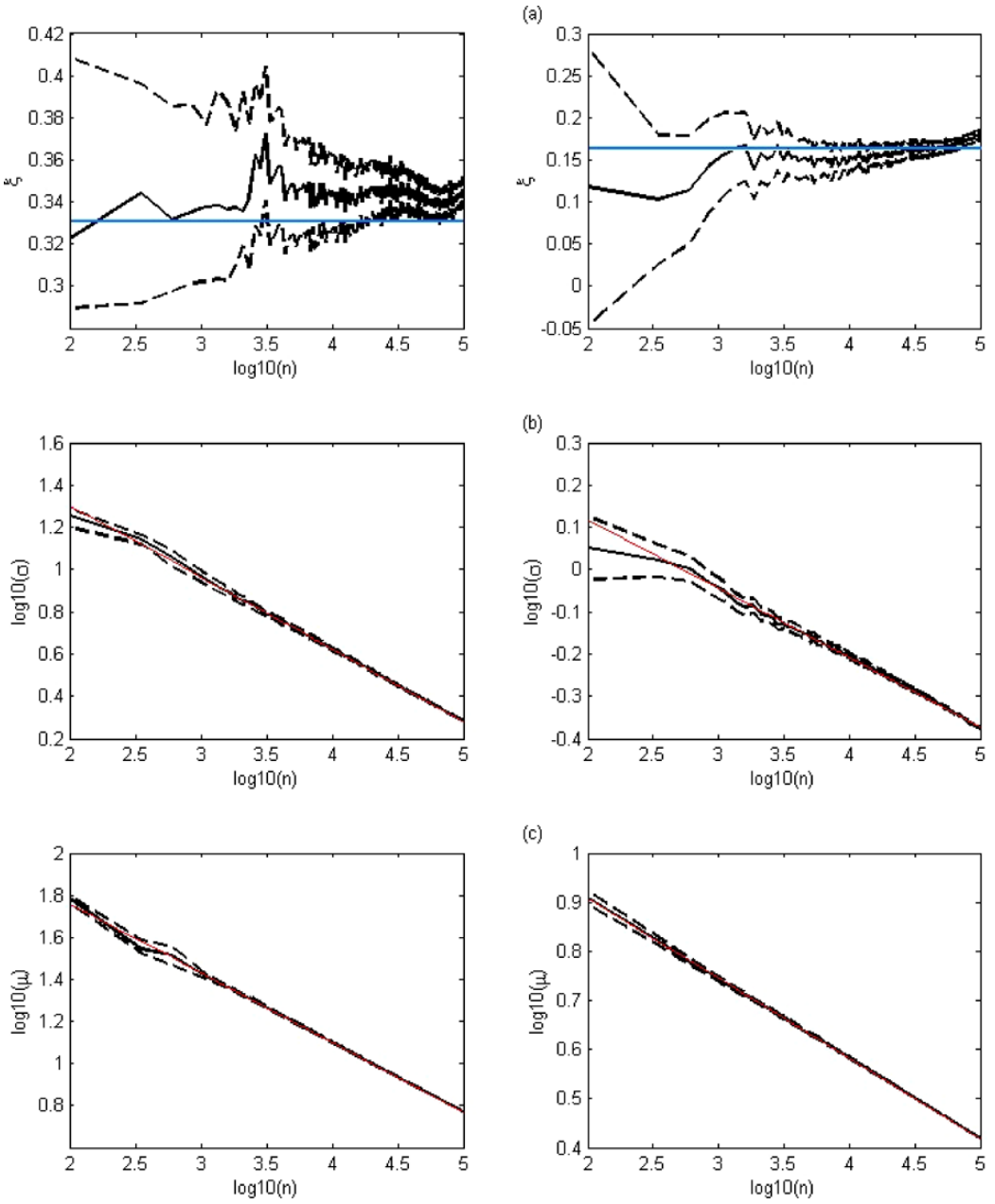


Fig. 3 (Color online) g_2 observable, $\zeta \simeq 0.51$. (**a**) ξ VS $\log_{10}(n)$; (**b**) $\log_{10}(\sigma)$ VS $\log_{10}(n)$; (**c**) $\log_{10}(\mu)$ VS $\log_{10}(n)$. *Left*: Bernoulli Shift map. *Right*: Arnold Cat Map. *Dotted lines* represent computed confidence interval, *red lines* represent a linear fit, *blue lines* are theoretical values

found when $n < 10^3$ or $m < 10^3$. Location parameter μ shows a logarithm decay with n as expected from (22). A linear fit of μ in respect to $\log(n)$ is shown with a red line in Fig. 2. The linear fit computed angular coefficients K^* of (22) well approximate $1/d$: for Bernoulli Shift map we obtain $|K^*| = 1.001 \pm 0.001$ while for Arnold Cat map $|K^*| = 0.489 \pm 0.001$. We find that ξ values have best matching with theoretical ones with reliable confidence interval when both $n > 10^3$ and $m > 10^3$. These results are confirmed even for g_2 type and g_3 type observable functions as shown in Figs. 3(a) and 4(a) respectively. We present the fit results for $\alpha = 3$ but we have done tests for different α and for fixed n and different α.

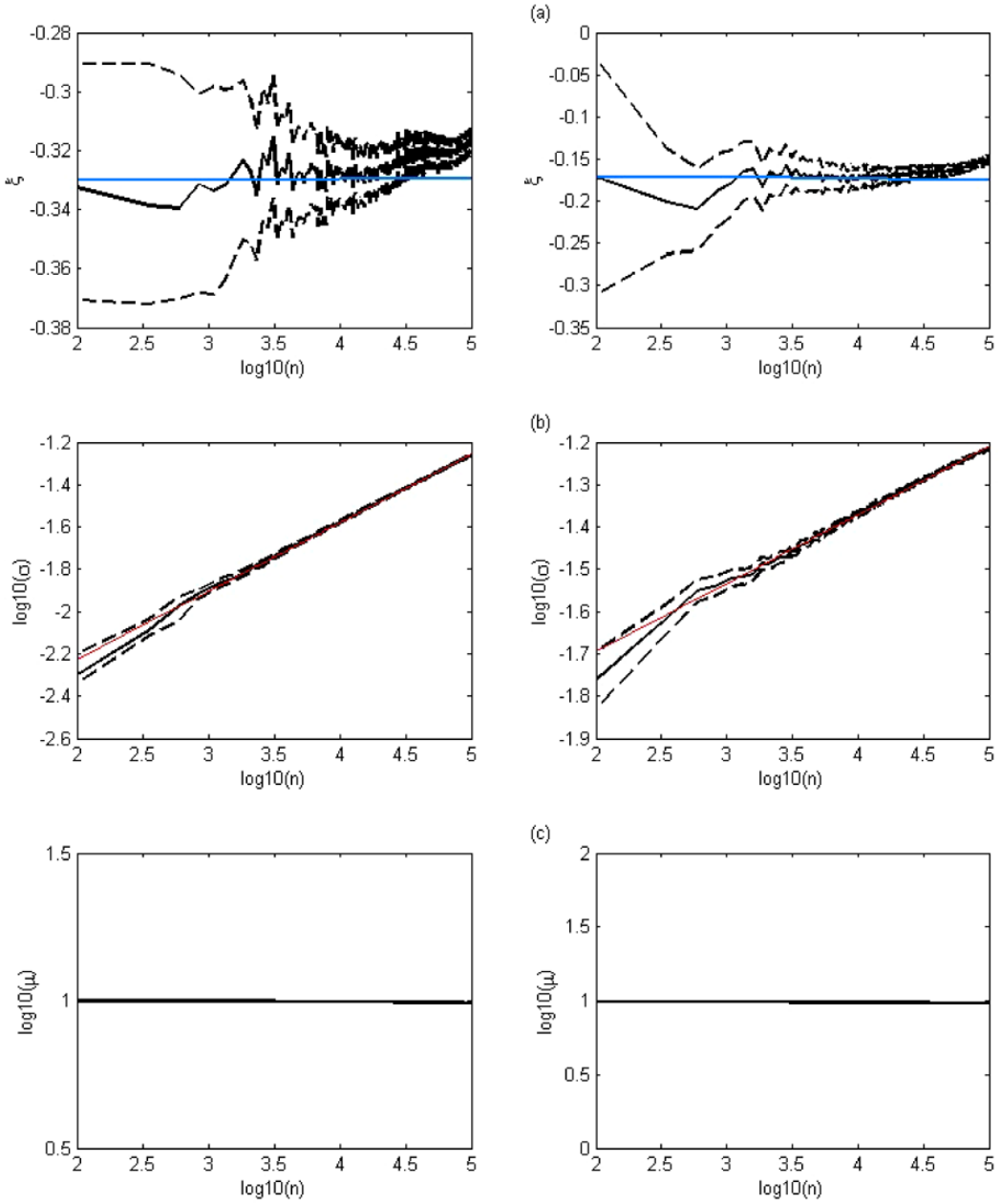


Fig. 4 (Color online) g_3 observable, $\zeta \simeq 0.51$. (**a**) ξ VS $\log_{10}(n)$; (**b**) $\log_{10}(\sigma)$ VS $\log_{10}(n)$; (**c**) $\log_{10}(\mu)$ VS $\log_{10}(n)$. *Left*: Bernoulli Shift map. *Right*: Arnold Cat Map. *Dotted lines* represent computed confidence interval, *red lines* represent a linear fit, *blue lines* are theoretical values

For g_2 observable function we can also check that μ and σ parameters follow a power law as described in (23). In the log-log plot in Fig. 3(b), 3(c), we can see a very clear linear behavior. For the Bernoulli Shift map we obtain $|K^*| = 0.330 \pm 0.001$ for μ series, $|K^*| = 0.341 \pm 0.001$ for σ in good agreement with theoretical value of $1/3$. For Arnold Cat map we expect to find $K^* = 1/6$, from the experimental data we obtain $|K^*| = 0.163 \pm 0.001$ for μ and $|K^*| = 0.164 \pm 0.001$ for σ.

Eventually, computing g_3 as observable function we expect to find a constant value for μ while σ has to grow with a power law in respect to n as expected from (24). As in g_2 case we expect $|K^*| = 1/(\alpha d)$ and numerical results shown in Fig. 4(b), 4(c) are consistent with the

theoretical one since $|K^*| = 0.323 \pm 0.006$ for Bernoulli shift map and $|K^*| = 0.162 \pm 0.006$ for Arnold Cat map.

In all cases considered the analytical behavior described in (23) and (24) is achieved and the fit quality improves if $n > 10^3$ and $m > 10^3$. The g_3 type observable constant has been chosen $C = 10$. The nature of these lower bound is quite different.

4 Distributions of Extremes in Mixing Map with Non-constant Density Measure

4.1 Asymptotic Sequences

The main problem when dealing with maps that have absolutely continuous but non-constant density measure $\rho(\zeta)$ is in the computation of the integral:

$$\nu(B_\delta(\zeta)) = \int_{B_\delta(\zeta)} \rho(x)\, dx \tag{27}$$

where $B_\delta(\zeta)$ is the d-dimensional ball of radius δ centered in ζ.

We have to know the value of this integral in order to evaluate $F(u)$ and, therefore, the sequences a_m and b_m.

As shown in the previous section δ is linked to the observable type: in all cases, since we substitute $u = 1 - 1/m$, $\delta \to 0$ means that we are interested in $m \to \infty$.

In this limit, a first order approximation of the previous integral is:

$$\nu(B_\delta(\zeta)) \simeq \rho(\zeta)\delta^d + \mathcal{O}(\delta^{d+1}) \tag{28}$$

that is valid if we are not in a neighborhood of a singular point of $\rho(\zeta)$.

As an example we compute the asymptotic sequences for a logistic map:

$$x_{t+1} = r x_t (1 - x_t) \tag{29}$$

with $r = 4$. This map satisfies hypothesis described in the analysis performed for Benedicks-Carleson maps in Moreira Freitas and Freitas [23].

For this map the density of the absolutely continuous invariant measure is explicit and reads:

$$\rho(\zeta) = \frac{1}{\pi\sqrt{\zeta(1-\zeta)}} \quad \zeta \in (0, 1) \tag{30}$$

So that:

$$\int_{B_\delta(\zeta)} \rho(\zeta)\, d\zeta = \frac{2}{\pi}\left[\arcsin(\sqrt{\zeta + \delta}) - \arcsin(\sqrt{\zeta - \delta})\right] \tag{31}$$

where $\zeta + \delta < 1$ and $\zeta - \delta > 0$. Since Extreme Value Theory effectively works only if n, m are large enough, the results in (31) can be replaced by a series expansion for $\delta \to 0$:

$$\frac{2}{\pi}\left[\arcsin(\sqrt{\zeta + \delta}) - \arcsin(\sqrt{\zeta - \delta})\right] = \frac{1}{\pi}\frac{2\delta}{\sqrt{\zeta(1-\zeta)}}[1 + \delta^2 P(\zeta) + \cdots] \tag{32}$$

up to order δ^3, where:

$$P(\zeta) = \frac{1}{8\zeta^2} - \frac{2}{\zeta(1-\zeta)} + \frac{2}{\zeta^2(1-\zeta)} + \frac{6}{\zeta^2(1-\zeta)^2} \tag{33}$$

Using the last two equations we are able to compute asymptotic normalizing sequences a_m and b_m for all g_i observables.

Case 1 $(g_1(x) = -\log(\text{dist}(x, \zeta)))$ For g_1 observable functions we set $\delta = e^{-ud}$. In case of logistic map $d = 1$. First we have to compute $G(t)$ using (15) and the expansion in equation:

$$G(t) = \frac{\int_t^\infty du(e^{-u} + e^{-3u}P(\zeta))}{e^{-t} + e^{-3t}P(\zeta)} \simeq 1 - \frac{2}{3}e^{-2t}P(\zeta) \tag{34}$$

We can compute γ_m, if $m \gg 1$, as follows:

$$F(\gamma_m) \simeq 1 - \frac{1}{m} \tag{35}$$

At the first order in (32) we get

$$\frac{1}{m} \simeq \frac{1}{\pi}\frac{2e^{-\gamma_m}}{\sqrt{\zeta(1-\zeta)}} \tag{36}$$

so that:

$$\gamma_m \simeq \ln(m) + \ln\left(\frac{2}{\pi\sqrt{\zeta(1-\zeta)}}\right) \tag{37}$$

Therefore, the sequences a_m and b_m if $m \gg 1$ are:

$$a_m \simeq [G(\gamma_m)]^{-1} \simeq 1 + \frac{2}{3}\frac{\pi^2}{4m^2}\zeta(\zeta-1)P(\zeta) \tag{38}$$

$$b_m \simeq \gamma_m \simeq \ln(m) + \ln(2\rho(\zeta)) \tag{39}$$

Case 2 $(g_2(x) = \text{dist}(x, \zeta)^{-1/\alpha})$ We can proceed as for g_1 setting $\delta = (\alpha u)^{-\alpha}$, computing γ_m we get at the first order in (32):

$$\frac{1}{m} \simeq \frac{1}{\pi}\frac{2\gamma_m^{-\alpha}}{\sqrt{\zeta(1-\zeta)}} = 2\rho(\zeta)(\alpha\gamma_m)^{-\alpha} \tag{40}$$

$$\gamma_m = \frac{1}{\alpha}\left(\frac{1}{2m\rho(\zeta)}\right)^{-1/\alpha} \tag{41}$$

We can respectively compute a_m and b_m as:

$$a_m = \gamma_m^{-1} \qquad b_m = (2m\rho(\zeta))^{-\xi} \tag{42}$$

Case 3 $(g_3(x) = C - \mathrm{dist}(x, \zeta)^{1/\alpha})$ As in the previous cases, we compute γ_m up to the first order setting $\delta = [\alpha(C - \gamma_m)]^\alpha$:

$$\frac{1}{m} \simeq \frac{1}{\pi} \frac{2[\alpha(C - \gamma_m)]^\alpha}{\sqrt{\zeta(1-\zeta)}} = 2\rho(\zeta)[\alpha(C - \gamma_m)]^\alpha \tag{43}$$

$$\gamma_m = C - \frac{1}{\alpha}\left(\frac{1}{2m\rho(\zeta)}\right)^{1/\alpha} \tag{44}$$

For type 3 distribution:

$$a_m = (u_F - \gamma_m)^{-1} \qquad b_m = u_F; \tag{45}$$

where $u_f = C$.

4.2 Numerical Experiment on the Logistic Map

Following the same procedure detailed in Sect. 3.2, we want to show the equivalence between EV computed normalizing sequences a_m and b_m and the parameters of a GEV distribution obtained directly fitting the data even in case of logistic map that has not constant density measure. Using (38)–(39) for g_1, we obtain the following theoretical expression:

$$\sigma(m, \zeta) \simeq 1 + \frac{2}{3}\frac{\pi^2}{4m^2}\zeta(\zeta - 1)P(\zeta) \qquad \mu(m, \zeta) \simeq \ln(m) + \ln(2\rho(\zeta)) \tag{46}$$

From (39), for g_2 observable type, we write:

$$\sigma(m, \zeta) \simeq \frac{1}{\alpha}(2m\rho(\zeta))^{\frac{1}{\alpha}} \qquad \mu(m, \zeta) \simeq (2m\rho(\zeta))^{\frac{1}{\alpha}} \tag{47}$$

and in g_3 case using (45), we expect to find:

$$\sigma(m, \zeta) \simeq \frac{1}{\alpha}(2m\rho(\zeta))^{-1/\alpha} \qquad \mu(m, \zeta) \simeq C = u_F \tag{48}$$

Values of ξ are independent on density and, as stated in Freitas' $\xi = 0$ for g_1 type, $\xi = 1/(\alpha d)$ for g_2 type and $\xi = -1/(\alpha d)$ for g_3 type.

In Figs. 5, 6, 7 we presents a numerical test of the asymptotic behavior described in (46)–(48) on logistic map for $d = 1$, $a = 3$, $C = u_F = 10$, $\zeta = 0.3$ against the variable n. As shown in previous section, block maxima approach works well with maps with constant density measure when n and m are at least 10^3: In fact, regarding ξ parameter. Significant deviations from the theoretical value are achieved when $n < 1000$ or $m < 1000$ even in the case of the Logistic Map.

Regarding μ and σ, for g_1 observable a linear fit of μ in respect to $\log(n)$ give us $|K^*| = 0.999 \pm 0.002$, while σ shows the same behavior of ξ since the best agreement with theoretical value $\sigma = 1$ is achieved when $n, m > 10^3$. In the log-log plots of Fig. 6(b), 6(c) for g_2 observable, we can observe again the expected linear behavior for μ and σ with $|K^*|$ corresponding to $1/(\alpha d)$. From numerical fit we obtain $|K^*| = 0.3334 \pm 0.0007$ for μ series and $|K^*| = 0.337 \pm 0.002$ for σ in good agreement with theoretical value of $1/3$. By applying a linear fit to the log-log plot in Fig. 7(b), the angular coefficient corresponding to σ series is $|K| = 0.323 \pm 0.003$ again consistent with the theory.

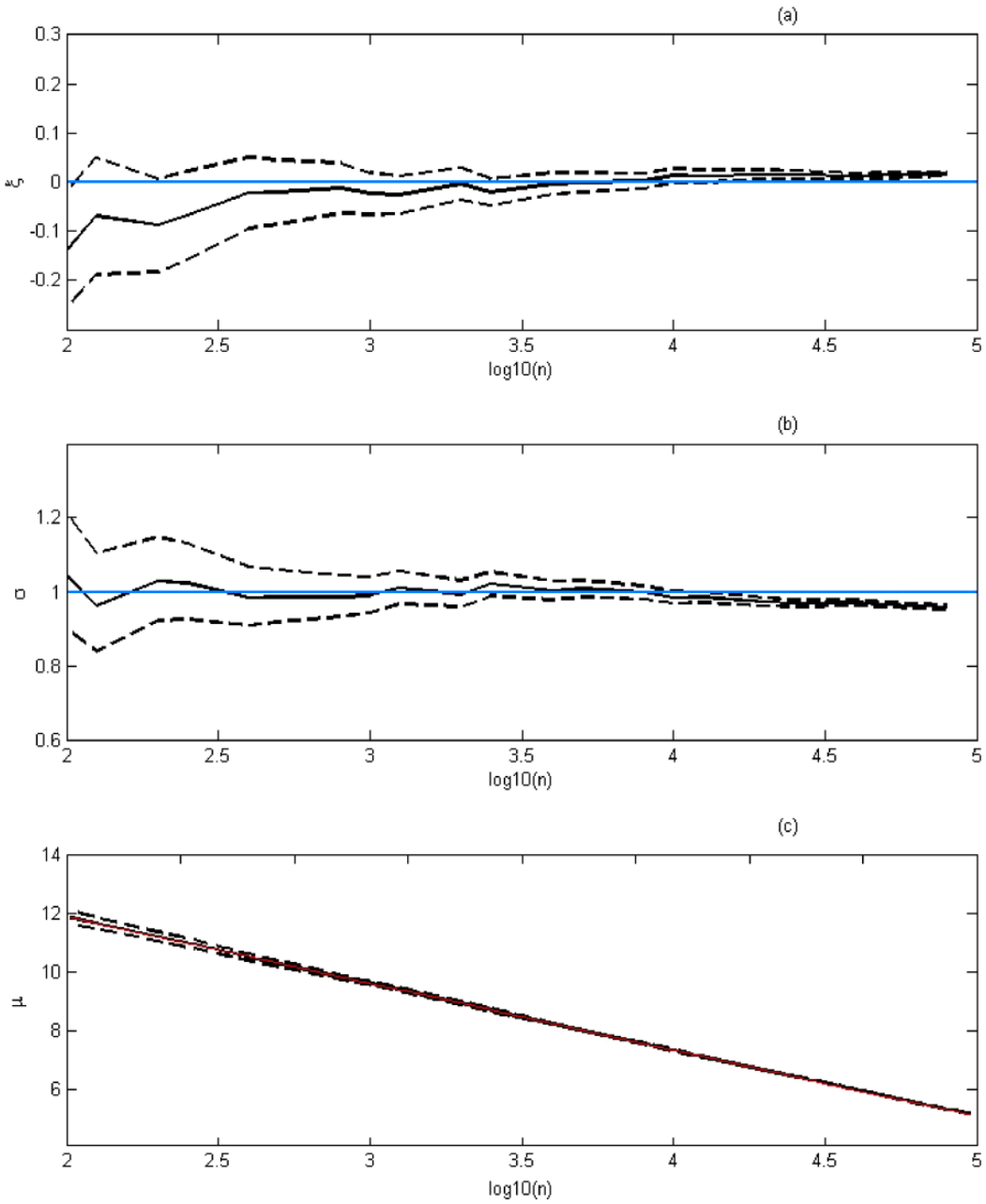


Fig. 5 (Color online) g_1 observable, $\zeta = 0.31$. (**a**) ξ VS $\log_{10}(n)$; (**b**) σ VS $\log_{10}(n)$; (**c**) μ VS $\log_{10}(n)$. Logistic map. *Dotted lines* represent computed confidence interval, *red lines* represent a linear fit, *blue lines* are theoretical values

For a logistic map we can also check the GEV behavior in respect to initial conditions. If we fix $n^* = m^* = 10^3$ and fit our data to GEV distribution for 10^3 different $\zeta \in (0, 1)$ an asymptotic behavior is reached as shown from the previous analysis. For g_1 observable function we have observed that the first order approximation works well for all three parameters. Deviation from this behavior are achieved for $\zeta \to 1$ and $\zeta \to 0$ as the measure become singular when we move to these points and we should take in account other terms of the series expansion. Numerically, we found that deviations from first order approximation are meaningful only if $\zeta < 10^{-3}$ and $\zeta > 1 - 10^{-3}$. Averaging over ζ both ξ and σ we

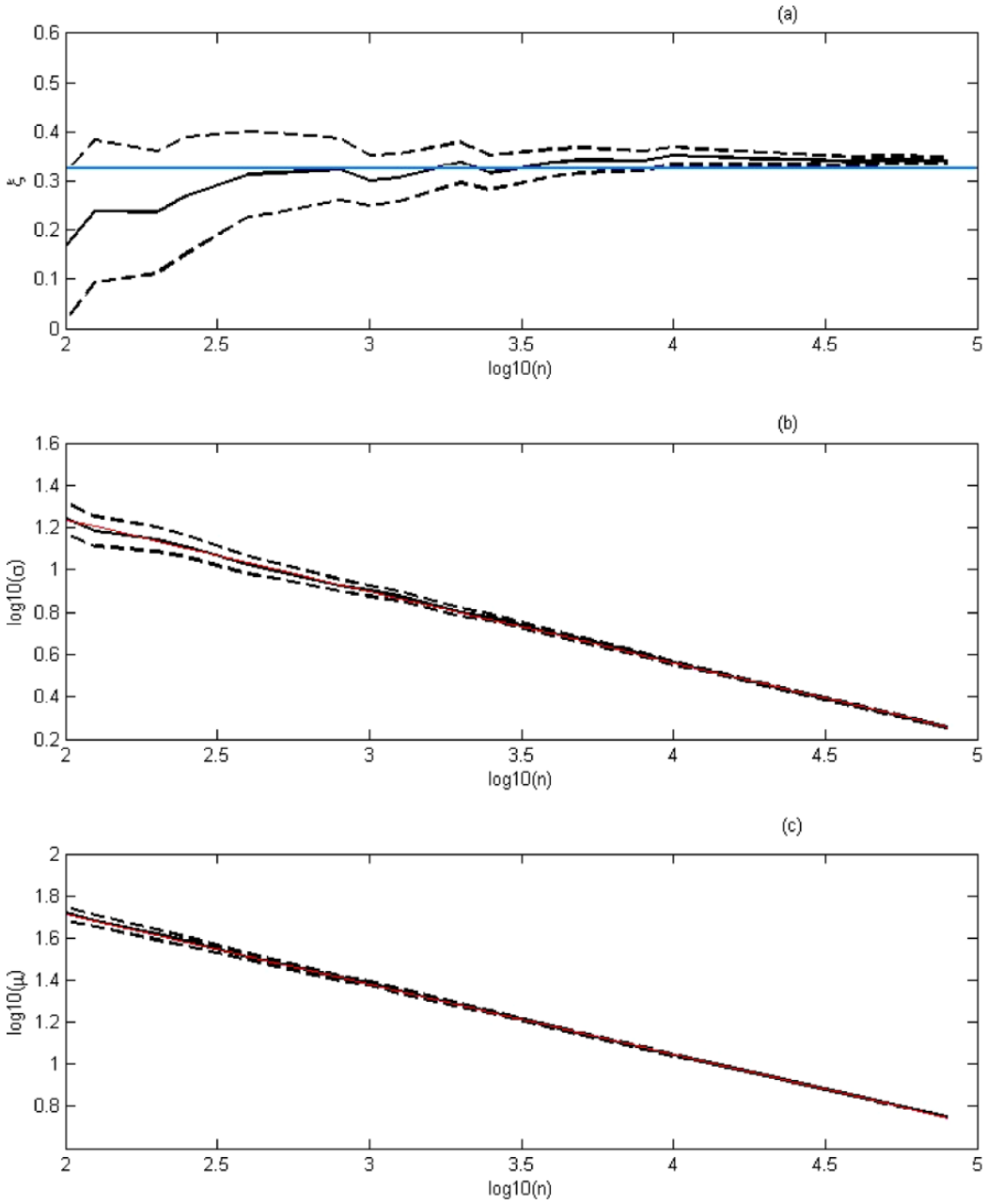


Fig. 6 (Color online) g_2 observable, $\zeta = 0.3$. (**a**) ξ VS $\log_{10}(n)$; (**b**) $\log_{10}(\sigma)$ VS $\log_{10}(n)$; (**c**) $\log_{10}(\mu)$ VS $\log_{10}(n)$. Logistic map. *Dotted lines* represent computed confidence interval, *red lines* represent a linear fit, *blue lines* are theoretical values

obtain $\langle \xi \rangle = 1.000 \pm 0.009$ and $\langle \sigma \rangle = 1.00 \pm 0.03$ where the uncertainties are computed with respect to the estimator. Since we expect $\xi = 0$ and $\sigma = 1$ at zero order approximation, numerical results are consistent with the theoretical ones; furthermore, experimental data are normally distributed around theoretical values.

Asymptotic expansion also works well for g_2 observables: we obtain $\langle \xi \rangle = 0.334 \pm 0.001$ in excellent agreement with theoretical value $\xi = 1/3$. Eventually, in g_3, averaging ξ over different initial conditions we get $\langle \xi \rangle = -0.334 \pm 0.002$ that is again consistent to theoretical value $-1/3$.

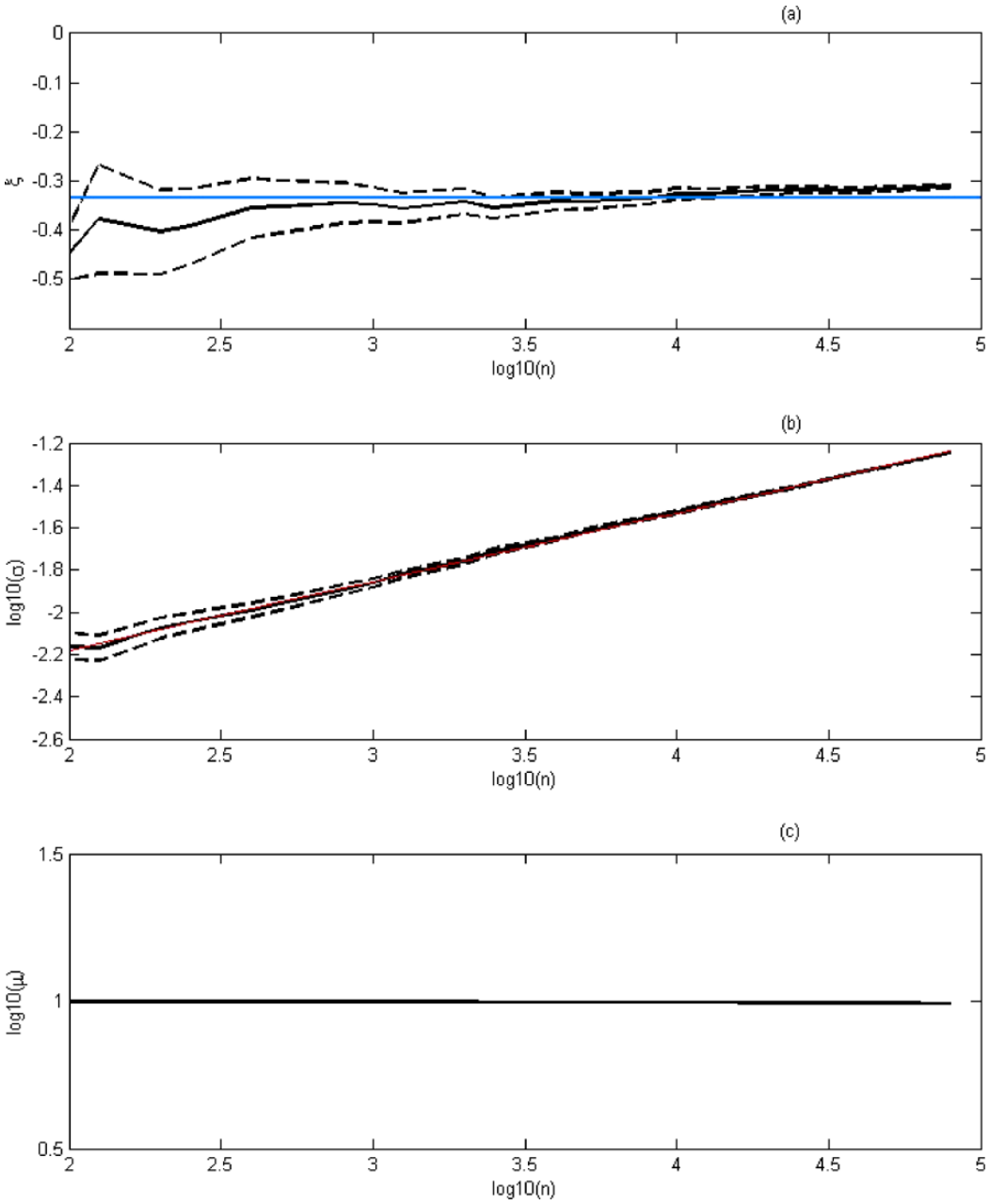


Fig. 7 (Color online) g_3 observable, $\zeta = 0.3$. (**a**) ξ VS $\log_{10}(n)$; (**b**) $\log_{10}(\sigma)$ VS $\log_{10}(n)$; (**c**) $\log_{10}(\mu)$ VS $\log_{10}(n)$. Logistic map. *Dotted lines* represent computed confidence interval, *red lines* represent a linear fit, *blue lines* are theoretical values

5 The Case of Regular Maps

Freitas and Freitas [22] have posed the problem of dependent extreme values in dynamical systems that show uniform quasi periodic motion. Here we try to investigate this problem numerically. We have used a one-dimensional and a bi-dimensional discrete map. The first one is the irrational translation on the torus defined by:

$$x_{t+1} = x_t + \beta \mod 1 \quad \beta \in [0, 1] \setminus \mathbb{Q}. \tag{49}$$

And for the bidimensional case, we use the so called standard map:

$$y_{t+1} = y_t + \frac{\lambda}{2\pi}\sin(2\pi x_t) \mod 1 \qquad x_{t+1} = x_t + y_{t+1} \mod 1 \tag{50}$$

with $\lambda = 10^{-4}$. For this value of λ, the standard map exhibits a regular behavior and it is not mixing, as well as torus translations. This means that these maps fail in satisfying hypothesis D_2 and D' and moreover they do not enjoy as well an exponential hitting time statistics. About this latter statistics, it is however known that it exists for torus translation and it is given by a particular piecewise linear function or a uniform distribution depending on which sequence of sets A_k is considered [10]. In a similar way, a non-exponential Hitting Time Statistics (HTS) is achieved for standard map when $\lambda \ll 1$ as well as for a skew map, that is a standard map with $\lambda = 0$ [7]. Therefore we expect not to obtain a GEV distribution of any type using g_i observables.

We have pointed out that the observable functions choice is crucial in order to observe some kind of distribution of extreme values when we are dealing with dynamical systems instead of stochastic series. Nicolis et al. [49] have shown how it is possible to obtain an analytical EV distribution which does not belong to GEV family choosing a simple observable: they considered the series of distances between the iterated trajectory and the initial condition. Using the same notation of Sect. 2 we can write:

$$Y_m(x = f^t\zeta) = \mathrm{dist}(f^t\zeta, \zeta) \qquad \hat{M}_m = \min\{Y_0, \ldots, Y_{m-1}\}$$

For this observable they have shown that the cumulative distribution $F(x) = P\{a_m(\hat{M}_m - b_m) \leq x\}$ of a uniform quasi periodic motions is not smooth but piecewise linear (Nicolis et al. [49], Fig. 3). Furthermore slop changes of $F(x)$ can be explained by constructing the intersections between different iterates of (49). $F(x)$ must correspond to a density distribution continuous obtained as a composition of box functions: each box must be related to a change in the slope of $F(x)$.

The numerical results we report below confirm that for the maps (49) and (50) the distributions of maxima for various observables cannot be fitted with a GEV since they are multi modal. We recall that the return times into a sphere of vanishing radius do not have a spectrum, if the orbits have the same frequency, whereas a spectrum appears if the frequency varies continuously with the action, as in the standard map for λ close to zero [38]. Since the EV statistics refers to a single orbit, no change due to the local mixing, which insures the existence of a return times spectrum [38], can be observed. Considering that the GEV exists when the system is mixing and does not when it is integrable, one might use the quality of fit to GEV as a dynamical indicator, for systems which exhibit regions with different dynamical properties, ranging from integrable to mixing as it occurs for the standard map when λ is order 1. Indeed we expect that in the neighborhood of a low order resonance, where the homoclinic tangle of intersecting separatrices appears, a GEV fit is possible. Preliminary computations carried out for the standard map and for a model with parametric resonance confirm this claim, that will be carefully tested in the near future.

Using the theoretical framework provided in Nicolis et al. [49] we check numerically the behavior of maps described in (49)–(50) analyzing EV distributions for g_i observable functions. Proceeding as in Sect. 3 for mixing maps, we try to perform a fit to GEV distribution starting with different initial conditions ζ, a set of different α values and (n, m) combinations. In all cases analyzed the Kolmogorov-Smirnov test fails and this means that GEV distribution is not useful to describe the behavior of this kind of statistics. This result is in

agreement with Freitas et al. [26] but we may find out which kind of empirical distribution is obtained.

Looking in details at M_m histograms that correspond to empirical density distributions, they appear always to be multi modal and each mode have a well defined shape: for g_1 type observable function modes are exponential while, for g_2 and g_3, their shape depends on α value of observable function. Furthermore, the number of modes and their positions are highly dependent on both n, m and initial conditions.

Using Nicolis et al. [49] results it is possible to understand why we obtain this kind of histograms: since density distribution of $\hat{M}_m$ is a composition of box functions, when we apply g_i observables we modulate it changing the shape of the boxes. Therefore, we obtain a multi modal distribution modified according to the observable functions g_i.

An example is shown in Fig. 8 for standard map: the left figures correspond to the histogram of the minimum distance obtained without computing g_i observable and reproduce a composition of box functions. The figures in the right show how this distribution is modified by applying g_1 observable to the series of minimum distances. We can see two exponential modes, while the third is hidden in the linear scale but can be highlighted using a log-scale. The upper figures are drawn using $n = 3300$, $m = 3300$, the lower with $n = 10000$, $m = 1000$.

6 Concluding Remarks

EVT was developed to study a wide class of problems of great interest in different disciplines: the need of modeling events that occur with very small probability comes from the fact that they can affect in a strong way several socioeconomic activities: floods, insurance losses, earthquakes, catastrophes. A very extensive account of EVT applications has been recently given in Ghil et al. [27]. EVT was applied on limited data series using the block-maxima approach facing the problem of having a good statistics of extreme values retaining a sufficient number of observation in each bin. Often, since no theoretical a priori values of GEV parameters are available for this kind of applications, we may obtain a biased fit to GEV distribution even if tests of statistical significance succeed. The recent development of an extreme value theory in dynamical systems give us the theoretical framework to test the consistency of block-maxima approach when analytical results for distribution parameters are available. This theory relies on the global properties of the dynamical systems considered (such as the degree of mixing or the decay rate of the Hitting Time Statistics) but also on the observable functions we chose.

Our main finding is that a block-maxima approach for GEV distribution is totally equivalent to fit an EV distribution after normalizing sequences are computed. To prove this we have derived analytical expressions for a_m and b_m normalizing sequences, showing that μ and σ of fitted GEV distribution can replace them. This approach works for maps that have an absolutely continuous invariant measure and retain some mixing properties that can be directly related to the exponential decay of HTS. Since GEV approach does not require the a priori knowledge of the measure density that is instead require by the EV approach, it is possible to use it in many numerical applications.

Furthermore, if we compare analytical and numerical results we can study what is the minimum number of maxima and how big the set of observations in which the maximum is taken has to be. To accomplish this goal we have analyzed maps with constant density measure finding that a good agreement between numerical and analytical value is achieved when both the number of maxima n and the observations per bin m are at least 10^3. We remark that the fits have passed Kolmogorov-Smirnov test with maximum confidence interval

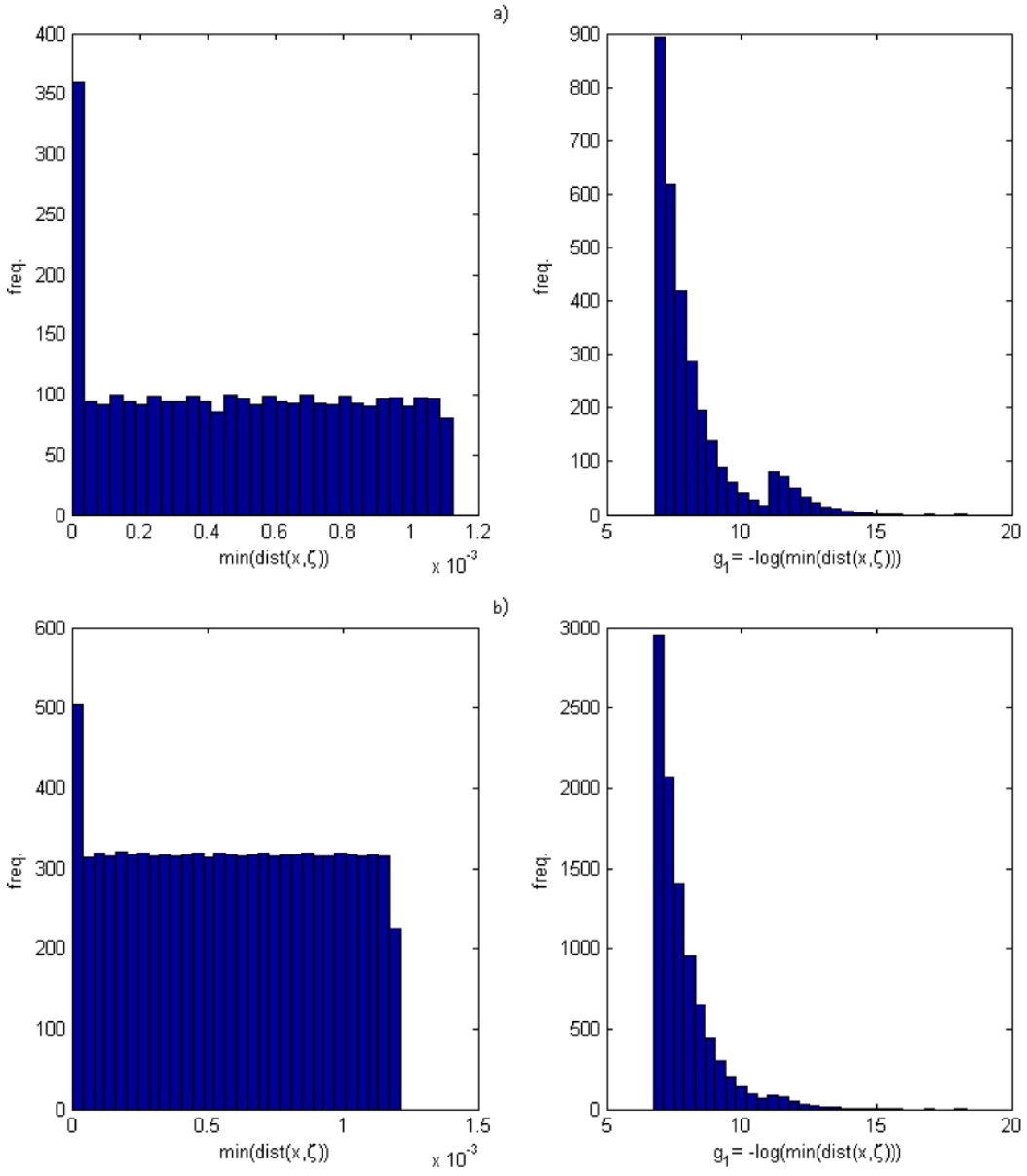


Fig. 8 Histogram of maxima for g_1 type observable function, standard map, $x_0 = y_0 = \sqrt{2} - 1$. *Left*: series of $\min(\text{dist}(f^t\zeta, \zeta))$. *Right*: series of $g_1 = -\log(\min(\text{dist}(f^t\zeta, \zeta)))$. (**a**) $n = 3300$, $m = 3300$. (**b**) $n = 10000$, $m = 1000$

even if $n < 10^3$ or $< m < 10^3$ so that parametric or non-parametric tests are not the only thing to take in account when dealing with extreme value distributions: if maxima are not proper extreme values (which means m is not large enough) the fit is good but parameters are different from expected values. The lower bound of n can be explained using the argument that a fit to a 3-parameters distribution needs at least 10^3 independent data to give reliable informations.

Therefore, we checked that in case of non-constant absolutely continuous density measure the asymptotic expressions used to compute μ and σ works when we consider n and m of order 10^3. For logistic map the numerical values of parameters we obtain averaging over different initial conditions are totally in agreement with the theoretical ones. In regu-

lar maps, as expected, the fit to a GEV distribution is unreliable. We obtain a multi modal distribution, that, for the analyzed maps, is the result of a composition of modes in which the shape depends on observable types. This behavior can be explained pointing out that this kind of systems have not an exponential HTS decay and therefore have no EV law for observables of type g_i.

To conclude, we claim that we have provided a reliable way to investigate properties of extreme values in mixing dynamical systems which may satisfy mixing conditions (like D_2 and D'), finding an equivalence among a_m, b_m, μ and σ behavior for absolutely continuous measures. In our future work we intend to address the case of singular measure. Recently the theorem was generalized to the case of non-smooth observations and therefore it holds also with non-absolutely continuous invariant probability measure [26]. In this case we expect the same for all the procedure described here. Understanding the extreme values behavior for singular measures will be crucial to apply proficiently this analysis to operative geophysical models since in these case we are always dealing with singular measures. In this way we will provide a complete tool to study extreme events in complex dynamical systems used in geophysical or financial applications.

Acknowledgements S.V. was supported by the CNRS-PEPS Project *Mathematical Methods of Climate Models*, and he thanks the GDRE Grefi-Mefi for having supported exchanges with Italy. V.L. and D.F. acknowledge the financial support of the EU FP7-ERC project NAMASTE: Thermodynamics of the Climate System.

References

1. Altmann, E.G., Hallerberg, S., Kantz, H.: Reactions to extreme events: moving threshold model. Phys. A, Stat. Mech. Appl. **364**, 435–444 (2006)
2. Arnold, V.I., Avez, A.: Ergodic Problems of Classical Mechanics. Benjamin, New York (1968)
3. Balakrishnan, V., Nicolis, C., Nicolis, G.: Extreme value distributions in chaotic dynamics. J. Stat. Phys. **80**(1), 307–336 (1995). ISSN 0022-4715
4. Beirlant, J.: Statistics of Extremes: Theory and Applications. Wiley, New York (2004). ISBN 0471976474
5. Bertin, E.: Global fluctuations and Gumbel statistics. Phys. Rev. Lett. **95**(17), 170601 (2005). ISSN 1079-7114
6. Brodin, E., Kluppelberg, C.: Extreme value theory in finance. In: Encyclopedia of Quantitative Risk Analysis and Assesment (2008). doi:10.1002/9780470061596.risk0431. ISBN:0-470-03549-8, 978-0-470-03549-8
7. Buric, N., Rampioni, A., Turchetti, G.: Statistics of Poincaré recurrences for a class of smooth circle maps. Chaos Solitons Fractals **23**(5), 1829–1840 (2005)
8. Burton, P.W.: Seismic risk in southern Europe through to India examined using Gumbel's third distribution of extreme values. Geophys. J. R. Astron. Soc. **59**(2), 249–280 (1979). ISSN 1365-246X
9. Clusel, M., Bertin, E.: Global fluctuations in physical systems: a subtle interplay between sum and extreme value statistics. Int. J. Mod. Phys. B **22**(20), 3311–3368 (2008). ISSN 0217-9792
10. Coelho, Z., De Faria, E.: Limit laws of entrance times for homeomorphisms of the circle. Isr. J. Math. **93**(1), 93–112 (1996). ISSN 0021-2172
11. Coles, S., Heffernan, J., Tawn, J.: Dependence measures for extreme value analyses. Extremes **2**(4), 339–365 (1999). ISSN 1386-1999
12. Collet, P.: Statistics of closest return for some non-uniformly hyperbolic systems. Ergod. Theory Dyn. Syst. **21**(02), 401–420 (2001)
13. Cornell, C.A.: Engineering seismic risk analysis. Bull. Seismol. Soc. Am. **58**(5), 1583 (1968). ISSN 0037-1106
14. Cruz, M.G.: Modeling, Measuring and Hedging Operational Risk. Wiley, New York (2002). ISBN 0471515604
15. Dahlstedt, K., Jensen, H.J.: Universal fluctuations and extreme-value statistics. J. Phys. A, Math. Gen. **34**, 11193 (2001)

16. Davison, A.C.: Modelling excesses over high thresholds, with an application. In: Statistical Extremes and Applications, pp. 461–482. Reidel, Dordrecht (1984)
17. Davison, A.C., Smith, R.L.: Models for exceedances over high thresholds. J. R. Stat. Soc., Ser. B, Methodol. **52**(3), 393–442 (1990). ISSN 0035-9246
18. Embrechts, P., Resnick, S.I., Samorodnitsky, G.: Extreme value theory as a risk management tool. N. Am. Actuar. J. **3**, 30–41 (1999). ISSN 1092-0277
19. Felici, M., Lucarini, V., Speranza, A., Vitolo, R.: Extreme value statistics of the total energy in an intermediate complexity model of the mid-latitude atmospheric jet. Part I: Stationary case. J. Atmos. Sci. **64**, 2137–2158 (2007)
20. Felici, M., Lucarini, V., Speranza, A., Vitolo, R.: Extreme value statistics of the total energy in an intermediate complexity model of the mid-latitude atmospheric jet. Part II: Trend detection and assessment. J. Atmos. Sci. **64**, 2159–2175 (2007)
21. Fisher, R.A., Tippett, L.H.C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Philos. Soc. **24**, 180 (1928)
22. Freitas, A.C.M., Freitas, J.M.: On the link between dependence and independence in extreme value theory for dynamical systems. Stat. Probab. Lett. **78**(9), 1088–1093 (2008). ISSN 0167-7152
23. Freitas, A.C.M., Freitas, J.M.: Extreme values for Benedicks-Carleson quadratic maps. Ergod. Theory Dyn. Syst. **28**(04), 1117–1133 (2008). ISSN 0143-3857
24. Freitas, A.C.M., Freitas, J.M., Todd, M.: Hitting time statistics and extreme value theory. Probab. Theory Relat. Fields **147**(3–4), 675–710 (2010)
25. Freitas, A.C.M., Freitas, J.M., Todd, M.: Extremal index, hitting time statistics and periodicity. Arxiv preprint, arXiv:1008.1350 (2010)
26. Freitas, A.C.M., Freitas, J.M., Todd, M., Gardas, B., Drichel, D., Flohr, M., Thompson, R.T., Cummer, S.A., Frauendiener, J., Doliwa, A., et al.: Extreme value laws in dynamical systems for non-smooth observations. Arxiv preprint, arXiv:1006.3276 (2010)
27. Ghil, M., et al.: Extreme events: dynamics, statistics and prediction. Nonlinear Process. Geophys. **18**, 295–350 (2011)
28. Gilli, M., Këllezi, E.: An application of extreme value theory for measuring financial risk. Comput. Econ. **27**(2), 207–228 (2006). ISSN 0927-7099
29. Gnedenko, B.: Sur la distribution limite du terme maximum d'une série aléatoire. Ann. Math. **44**(3), 423–453 (1943)
30. Gumbel, E.J.: The return period of flood flows. Ann. Math. Stat. **12**(2), 163–190 (1941). ISSN 0003-4851
31. Gupta, C.: Extreme-value distributions for some classes of non-uniformly partially hyperbolic dynamical systems. Ergod. Theory Dyn. Syst. **30**(03), 757–771 (2010)
32. Gupta, C., Holland, M., Nicol, M.: Extreme value theory for dispersing billiards and a class of hyperbolic maps with singularities. Preprint (2009)
33. Haiman, G.: Extreme values of the tent map process. Stat. Probab. Lett. **65**(4), 451–456 (2003). ISSN 0167-7152
34. Hallerberg, S., Kantz, H.: Influence of the event magnitude on the predictability of an extreme event. Phys. Rev. E **77**(1), 11108 (2008). ISSN 1550-2376
35. Hasselblatt, B., Katok, A.B.: A First Course in Dynamics: With a Panorama of Recent Developments. Cambridge University Press, Cambridge (2003)
36. Hill, B.M.: A simple general approach to inference about the tail of a distribution. Ann. Stat. **3**(5), 1163–1174 (1975). ISSN 0090-5364
37. Holland, M., Nicol, M., Török, A.: Extreme value distributions for non-uniformly hyperbolic dynamical systems. Preprint (2008)
38. Hu, H., Rampioni, A., Rossi, L., Turchetti, G., Vaienti, S.: Statistics of Poincaré recurrences for maps with integrable and ergodic components. Chaos, Interdiscip. J. Nonlinear Sci. **14**, 160 (2004)
39. Kantz, H., Altmann, E., Hallerberg, S., Holstein, D., Riegert, A.: Dynamical interpretation of extreme events: predictability and predictions. In: Extreme Events in Nature and Society, pp. 69–93. Springer, Berlin (2006)
40. Katz, R.W.: Extreme value theory for precipitation: sensitivity analysis for climate change. Adv. Water Resour. **23**(2), 133–139 (1999). ISSN 0309-1708
41. Katz, R.W., Brown, B.G.: Extreme events in a changing climate: variability is more important than averages. Clim. Change **21**(3), 289–302 (1992). ISSN 0165-0009
42. Katz, R.W., Brush, G.S., Parlange, M.B.: Statistics of extremes: modeling ecological disturbances. Ecology **86**(5), 1124–1134 (2005). ISSN 0012-9658
43. Leadbetter, M.R., Lindgren, G., Rootzen, H.: Extremes and Related Properties of Random Sequences and Processes. Springer, New York (1983)

44. Lilliefors, H.W.: On the Kolmogorov-Smirnov test for normality with mean and variance unknown. J. Am. Stat. Assoc. **62**(318), 399–402 (1967). ISSN 0162-1459
45. Longin, F.M.: From value at risk to stress testing: the extreme value approach. J. Bank. Finance **24**(7), 1097–1130 (2000). ISSN 0378-4266
46. Martinez, W.L., Martinez, A.R.: Computational Statistics Handbook with MATLAB. CRC Press, Boca Raton (2002)
47. Martins, E.S., Stedinger, J.R.: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resour. Res. **36**(3), 737–744 (2000). ISSN 0043-1397
48. Nicholls, N.: CLIVAR and IPCC interests in extreme events. In: Workshop Proceedings on Indices and Indicators for Climate Extremes, Asheville, NC. Sponsors, CLIVAR, GCOS and WMO (1997)
49. Nicolis, C., Balakrishnan, V., Nicolis, G.: Extreme events in deterministic dynamical systems. Phys. Rev. Lett. **97**(21), 210602 (2006). ISSN 1079-7114
50. Friederichs, P., Hense, A.: Statistical downscaling of extreme precipitation events using censored quantile regression. Mon. Weather Rev. **135**(6), 2365–2378 (2007). ISSN 0027-0644
51. Pickands III, J.: Moment convergence of sample extremes. Ann. Math. Stat. **39**(3), 881–889 (1968)
52. Pickands III, J.: Statistical inference using extreme order statistics. Ann. Stat. **3**(1), 119–131 (1975). ISSN 0090-5364
53. Smith, R.L.: Threshold methods for sample extremes. Stat. Extremes Appl. **621**, 638 (1984)
54. Smith, R.L.: Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. Stat. Sci. **4**(4), 367–377 (1989). ISSN 0883-4237
55. Sornette, D., Knopoff, L., Kagan, Y.Y., Vanneste, C.: Rank-ordering statistics of extreme events: application to the distribution of large earthquakes. J. Geophys. Res. **101**(B6), 13883 (1996). ISSN 0148-0227
56. Sveinsson, O.G.B., Boes, D.C.: Regional frequency analysis of extreme precipitation in Northeastern Colorado and Fort Collins flood of 1997. J. Hydrol. Eng. **7**, 49 (2002)
57. Todorovic, P., Zelenhasic, E.: A stochastic model for flood analysis. Water Resour. Res. **6**(6), 1641–1648 (1970). ISSN 0043-1397
58. Vannitsem, S.: Statistical properties of the temperature maxima in an intermediate order Quasi-Geostrophic model. Tellus A **59**(1), 80–95 (2007). ISSN 1600-0870
59. Vitolo, R., Holland, M.P., Ferro, C.A.T.: Robust extremes in chaotic deterministic systems. Chaos, Interdiscip. J. Nonlinear Sci. **19**, 043127 (2009)
60. Vitolo, R., Ruti, P.M., Dell'Aquila, A., Felici, M., Lucarini, V., Speranza, A.: Accessing extremes of mid-latitudinal wave activity: methodology and application. Tellus A **61**(1), 35–49 (2009). ISSN 1600-0870
61. Young, L.S.: Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. **147**(3), 585–650 (1998)
62. Young, L.S.: Recurrence times and rates of mixing. Isr. J. Math. **110**(1), 153–188 (1999)

Chapter 4

Article: Generalized Extreme Value Distribution as dynamical indicators of stability

International Journal of Bifurcation and Chaos

Generalized Extreme Value distribution parameters as dynamical indicators of Stability

Faranda, Davide
Klimacampus, Universität Hamburg.
Grindelberg 5, 20144, Hamburg, Germany.
davide.faranda@zmaw.de
Department of Mathematics, University of Reading.
Whiteknights, PO Box 220, Reading RG6 6AX, UK.
d.faranda@pgr.reading.ac.uk

Lucarini, Valerio
Klimacampus, Universität Hamburg.
Grindelberg 5, 20144, Hamburg, Germany.
valerio.lucarini@zmaw.de
Department of Mathematics, University of Reading.
Whiteknights, PO Box 220, Reading RG6 6AX, UK.
v.lucarini@reading.ac.uk

Turchetti, Giorgio
Department of Physics, University of Bologna. INFN-Bologna.
Via Irnerio 46, Bologna, 40126, Italy.
turchett@bo.infn.it

Vaienti, Sandro
UMR-6207, Centre de Physique Théorique, CNRS, Universités d'Aix-Marseille I,II.
Université du Sud Toulon-Var and FRUMAM.
(Fédération de Recherche des Unités de Mathématiques de Marseille);
CPT, Luminy, Case 907, 13288 Marseille Cedex 09, France.
vaienti@cpt.univ-mrs.fr

We introduce a new dynamical indicator of stability based on the Extreme Value statistics showing that it provides an insight on the local stability properties of dynamical systems. The indicator performs faster than others based on the iteration of the tangent map since it requires only the evolution of the original systems and, in the chaotic regions, gives further information about the local information dimension of the attractor. A numerical validation of the method is presented through the analysis of the motions in the Standard map.

1. Introduction

The analysis of stability for discrete and continuous time dynamical systems is of fundamental importance to get insights in the dynamical structure of a system. The distinction among regular and chaotic orbits can be easily made in a dissipative case whereas for conservative systems it is usually an hard task especially

when we are dealing with many degrees of freedom or when the phase space is spitted into chaotic regions and regular islands. A large number of tools known as indicators of stability have been developed to accomplish this task: Lyapunov Characteristic Exponents (LCEs) [Wolf *et al.*, 1985], [Rosenstein *et al.*, 1993], [Skokos, 2010] and the indicators related to the Return Time Statistics [Kac, 1934], [Gao, 1999], [Hu *et al.*, 2004], [Buric *et al.*, 2005] have been used from a long time as efficient indicators. Nevertheless, in the recent past, the need for computing stability properties with faster algorithms and for systems with many degrees of freedom resulted in a renewed interest in the technique and different dynamical indicators have been introduced. The Smaller Alignment Index (SALI) described in Skokos *et al.* [2002] and Skokos *et al.* [2004], the Generalized Alignment Index (GALI), introduced in Skokos *et al.* [2007] and the Mean Exponential Growth factor of Nearby Orbits (MEGNO) discussed in Cincotta *et al.* [2003], Goździewski *et al.* [2001] are suitable to analyse the properties of a single orbit. They are based on the divergence of nearby trajectories and require in principle the knowledge of the exact dynamics. Another class of indicators is based on the round off error properties and has been discussed in Faranda *et al.* [2011c]: the divergence between two trajectories starting from the same initial condition but computed with different numeric precision can be used to illustrate the dynamical structure. The so called Reversibility Error that measures the distance between a certain initial condition and the end point of a trajectory iterated forward and backward for the same number of time steps give basically the same informations.
Other tools such as Fidelity and Correlations decay can be successfully used to characterize stability properties of ensemble of orbits as explained in Liverani *et al.* [2007] and Turchetti *et al.* [2010]. Eventually, Frequency Map Analysis has been used to provide informations about the resonance structure of a system [Laskar, 1999], [Robutel & Laskar, 2001].

Although these indicators have been developed to accomplish the same task, each of them presents specific features and limitations and often it is necessary to combine a certain number of indicators to get quantitative information about the dynamics. Another aspect to consider when comparing the efficiency of chaos indicators is the computational cost: each variational method (SALI, MEGNO, mLCE) needs to iterate both the map and the tangent map forward during n steps. The latter map, although linear, is the computationally most expensive due to the fact that it needs the evaluation of the Jacobian matrix at every step. On the other hand the Round off error methods are the computationally less expensive as they require only the iteration of the dynamics. Fidelity and Correlations decay are usually computed using Monte Carlo simulations and therefore they are inaccessible for systems with many degrees of freedom [Turchetti *et al.*, 2010]. Therefore, it is clear that there is still the need to introduce versatile indicators that distinguish regular from chaotic behaviors and possibly give further informations on the dynamics. The purpose of this paper is to use the series of extrema of some specific observable computed using the orbits of dynamical systems as a new indicator of stability.

Extreme Value Theory was originally introduced by Fisher & Tippett [1928], Gnedenko [1943] to study the maxima of a series of independent and identical distributed variables: under very general hypothesis a limiting distribution called Generalised Extreme Value (GEV) distribution exists for the series of extremes. An extensive account of recent results and relevant applications is given in Ghil et al. [2011]. In the recent past this theory has been adapted to study the output of dynamical systems. As we will explain in detail in the next section it is not trivial to observe the asymptotic GEV distribution in dynamical systems: some sort of independence of maxima must be recovered by requiring certain mixing conditions on the orbits. Furthermore, we need to introduce some peculiar observables that satisfy the condition proposed by Gnedenko on the parent distribution of data: they are related to the closest return of a trajectory in a ball centered around the starting point and therefore allow a very detailed description of the dynamics in the neighborhood of the initial condition.
If all these requirements are satisfied it is possible to observe an Extreme Value (EV) statistics that converges to the GEV distribution family. The parameters of the distribution are intimately related with some relevant dynamical quantities such as the local dimension of the attractor [Freitas *et al.*, 2009], [Faranda *et al.*, 2011a]. We will use these features to show the reliability of GEV distribution parameters to discriminate regular from chaotic behaviors pointing out the further information that is possible to extract

regarding the dynamics.
The paper is organised as follows: in section 2 we explain how to introduce an EV statistics in dynamical systems pointing out the difference between regular and chaotic orbits. In section 3 we describe the numerical algorithm and procedure used to compute the parameters of the GEV distribution. Eventually, in section 4 we perform some tests on the Standard Map to validate numerically the use of GEV parameters as dynamical indicators.

2. Extreme Statistics as dynamical indicator

2.1. *Extreme Value Theory in dynamical systems*

Gnedenko [1943] studied the convergence of maxima of i.i.d. variables

$$X_0, X_1, ..X_{m-1}$$

with cumulative distribution function (cdf) $F(x)$ of the form:

$$F(x) = P\{a_m(M_m - b_m) \leq x\}$$

Where a_m and b_m are normalizing sequences and $M_m = \max\{X_0, X_1, ..., X_{m-1}\}$. It may be rewritten as $F(u_m) = P\{M_m \leq u_m\}$ where $u_m = x/a_m + b_m$. Under general hypothesis on the nature of the parent distribution of data, Gnedenko [1943] showed that the cdf of maxima $F(x)$, up to an affine change of variable, converges to one of the following three limit laws $E_i(x), i = 1, 2, 3$:

- Type 1 (*Gumbel*).

$$E_1(x) = \exp(-e^{-x}),\ -\infty < x < \infty \tag{1}$$

- Type 2 (*Fréchet*).

$$E_2(x) = \begin{cases} 0,\ x \leq 0 \\ \exp(-x^{-\xi}),\ \text{for some } \xi > 0,\ x > 0 \end{cases} \tag{2}$$

- Type 3 (*Weibull*).

$$E_3(x) = \begin{cases} \exp(-(-x)^{\xi}),\ \text{for some } \xi > 0,\ x \leq 0 \\ 1,\ x > 0 \end{cases} \tag{3}$$

Let us define the right endpoint x_F of $F(x)$ as:

$$x_F = \sup\{x : F(x) < 1\}, \tag{4}$$

then, it is possible to compute normalizing sequences a_m and b_m using the following corollary of Gnedenko's theorem :
Corollary (Gnedenko): *The normalizing sequences a_m and b_m in the convergence of normalized maxima $P\{a_m(M_m - b_m) \leq x\} \to F(x)$ may be taken (in order of increasing complexity) as:*

- *Type 1:* $a_m = [G(\gamma_m)]^{-1}, \quad b_m = \gamma_m;$
- *Type 2:* $a_m = \gamma_m^{-1}, \quad b_m = 0;$
- *Type 3:* $a_m = (x_F - \gamma_m)^{-1}, \quad b_m = x_F;$

where

$$\gamma_m = F^{-1}(1 - 1/m) = \inf\{x; F(x) \geq 1 - 1/m\}; \tag{5}$$

$$G(t) = \int_t^{x_F} \frac{1 - F(u)}{1 - F(t)} du, \quad t < x_F; \tag{6}$$

In Faranda *et al.* [2011b] we have shown that this approach is equivalent to fit unnormalized data directly to a single family of generalized distribution called GEV distribution with cdf:

$$F_G(x;\mu,\sigma,\xi') = \exp\left\{-\left[1+\xi'\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi'}\right\}; \tag{7}$$

which holds for $1+\xi'(x-\mu)/\sigma > 0$, using $\mu \in \mathbb{R}$ (location parameter) and $\sigma > 0$ (scale parameter) as scaling constants in place of b_m, and a_m [Pickands III, 1968], in particular, in Faranda *et al.* [2011b] we have shown that the following relations hold:

$$\mu = b_m \qquad \sigma = \frac{1}{a_m}.$$

$\xi' \in \mathbb{R}$ is the shape parameter also called the tail index: when $\xi' \to 0$, the distribution corresponds to a Gumbel type (Type 1 distribution). When the index is positive, it corresponds to a Fréchet (Type 2 distribution); when the index is negative, it corresponds to a Weibull (Type 3 distribution).

In the last decade many works focused on the possibility of treating time series of observables of deterministic dynamical system using EVT. For example, Balakrishnan *et al.* [1995] and more recently Nicolis *et al.* [2006] and Haiman [2003] have shown that for regular orbits of dynamical systems we do not expect to find convergence to EV distribution.
The first rigorous mathematical approach to extreme value theory in dynamical systems goes back to the pioneer paper by Collet [2001]. Important contributions have successively been given by Freitas & Freitas [2008], Freitas *et al.* [2009], Freitas *et al.* [2010] and by Gupta *et al.* [2009]. The goal of all these investigations was to associate to the stationary stochastic process given by the dynamical system, a new stationary independent sequence: when the latter sequence satisfies one of the classical three extreme value laws it is possible to show that the same result also holds for the original dynamical sequence.

Let us consider a dynamical system $(\Omega, \mathcal{B}, \nu, f)$, where Ω is the invariant set in some manifold, usually $\mathbb{R}^d$, $\mathcal{B}$ is the Borel σ-algebra, $f : \Omega \to \Omega$ is a measurable map and ν a probability f-invariant Borel measure. In order to adapt the extreme value theory to dynamical systems, we will consider the stationary stochastic process $X_0, X_1, ...$ given by:

$$X_m(x) = g(\mathrm{dist}(f^m(x), \zeta)) \qquad \forall m \in \mathbb{N}, \tag{8}$$

where 'dist' is a distance on the ambient space Ω, ζ is a given point and g is an observable function, and whose partial maximum is defined as:

$$M_m = \max\{X_0, ..., X_{m-1}\}. \tag{9}$$

The probability measure will be here an invariant measure ν for the dynamical system. We will also suppose that our systems verify the condition D_2 and D' which will allow us to use the EVT for i.i.d. sequences. Hereinafter we will use three types of observables $g_i, i = 1, 2, 3$ that are suitable to obtain one of the three types of EV distribution for normalized maxima:

$$g_1(x) = -\log(\mathrm{dist}(x, \zeta)), \tag{10}$$

$$g_2(x) = \mathrm{dist}(x, \zeta)^{-1/\alpha}, \tag{11}$$

$$g_3(x) = C - \mathrm{dist}(x, \zeta)^{1/\alpha}, \tag{12}$$

where C is a constant and $\alpha > 0 \in \mathbb{R}$ [Collet, 2001], [Freitas *et al.*, 2009].

Using these observables we can obtain convergence to the Type 1,2,3 distribution if one can prove two sufficient conditions called D_2 and D' which we briefly explain here: these conditions basically require a sort of independence of the stochastic dynamical sequence in terms of uniform mixing condition on the distribution functions. In particular condition D_2, introduced in its actual form by Freitas & Freitas [2008], could be checked directly by estimating the rate of decay of correlations for Hölder observables.
If $X_m, m \geq 0$ is our stochastic process, we can define $M_{j,l} \equiv \max\{X_j, X_{j+1}, \cdots, X_{j+l}\}$ and we put $M_{0,m} = M_m$.
The condition $D_2(u_m)$ holds for the sequence X_m if for any integer l, t, m we have

$$|\nu(X_0 > u_m, M_{t,l} \leq u_m) - \nu(X_0 > u_m)\nu(M_{t,l} \leq u_m)| \leq \gamma(m,t),$$

where $\gamma(m,t)$ is non-increasing in t for each m and $m\gamma(m,t_m) \to 0$ as $m \to \infty$ for some sequence $t_m = o(m)$, $t_m \to \infty$.
We say condition $D'(u_m)$ holds for the sequence X_m if

$$\lim_{l\to\infty} \limsup_m m \sum_{j=1}^{\lfloor m/l \rfloor} \nu(X_0 > u_m, X_j > u_m) = 0.$$

Here $\lfloor m/l \rfloor$ indicates the integer part of m/l.
Instead of checking the previous conditions, we can use another results that established a connection between the extreme value laws and the statistics of first return and hitting times, see the papers by Freitas *et al.* [2009] and Freitas *et al.* [2011]. We will use the this result instead of checking directly the mixing conditions D' and D_2. To introduce the Hitting Time Statistics (hereinafter HTS) we need first to define the recurrence time τ_A in a measurable set $A \in \Omega$, as

$$\tau_A(x) = \inf_{t\geq 1} \left\{ x \in A : f^t(x) \in A \right\},$$

and the average recurrence time $< \tau_A >$ as

$$< \tau_A >= \int \tau_A(x) \mathrm{d}\mu_A(x) \qquad \mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)},$$

Following Hirata *et al.* [1999] and Buric *et al.* [2003], we define the HTS as the following limit (whenever it exists):

$$H(t) = \lim_{\mu(A)\to 0} \mu_A(A_{>t}) \qquad A_{>t} \equiv \left\{ x \in A : \frac{\tau_A(x)}{< \tau_A >} > t \right\}. \tag{13}$$

In particular, Freitas *et al.* [2009] and Freitas *et al.* [2011] showed that for dynamical systems preserving an absolutely continuous invariant measure or a singular continuous invariant measure ν, the existence of an exponential hitting time statistics on balls around almost any point ζ implies the existence of extreme value laws for one of the observables of type $g_i, i = 1, 2, 3$ described above. The converse is also true, namely if we have an extreme value law which applies to the observables of type $g_i, i = 1, 2, 3$ achieving a maximum at ζ, then we have exponential hitting time statistics to balls with center ζ. Recently these results have been generalized to local returns around balls centered at periodic points [Freitas *et al.*, 2010].

2.2. *Extreme Value Statistics in mixing and regular maps*

In Faranda *et al.* [2011b] and Faranda *et al.* [2011a] we have analised both from an analytical and numerical point of view the Extreme Value distribution in a wide class of low dimensional maps showing that, when the conditions D' and D_2 are verified, the block maxima approach can be used to study extrema. It consists of dividing the data series of length k of some observable into n bins each containing the same number m of observations, and selecting the maximum (or the minimum) value in each of them [Coles *et al.*, 1999].

Using g_i observable functions we have shown that a first order approximation of the GEV parameters in mixing maps can be written in terms of m (or equivalently n) and the local dimension of the attractor d:

For g_1 type observable:

$$\sigma = \frac{1}{d} \qquad \mu \sim \frac{1}{d}\ln(k/n) \qquad \xi' = 0 \tag{14}$$

For g_2 type observable:

$$\sigma \sim n^{-1/(\alpha d)} \qquad \mu \sim n^{-1/(\alpha d)} \qquad \xi' = \frac{1}{\alpha d} \tag{15}$$

For g_3 type observable:

$$\sigma \sim n^{1/(\alpha d)} \qquad \mu = C \qquad \xi' = -\frac{1}{\alpha d} \tag{16}$$

while the higher order terms contain explicit dependence on the density measure.

For regular maps independently on the observable chosen, for periodic or quasi-periodic orbits we do not observe convergence to the GEV distribution. The HTS is not exponential and therefore, according to Freitas *et al.* [2009], we do not expect convergence to the GEV distribution. In Faranda *et al.* [2011b] we have analysed what happens when the observables g_i are used to study empirical maxima distribution in regular maps: Nicolis *et al.* [2006] have shown how it is possible to obtain an analytical EV distribution which does not belong to GEV family choosing a simple observable: they considered the series of distances between the iterated trajectory and the initial condition:

$$Y_m(x = f^t\zeta) = \mathrm{dist}(f^t\zeta, \zeta) \qquad \hat{M}_m = \min\{Y_0, ...Y_{m-1}\}$$

For this observable they have shown that the cumulative density function $F(x) = P\{a_m(\hat{M}_m - b_m) \leq x\}$ of a uniform quasiperiodic motion is not smooth but piecewise linear (Nicolis *et al.* [2006], Figure 3). $F(x)$ must correspond to a probability density function (pdf) obtained as a composition of Heaviside step functions: each box must be related to a change in the slope of $F(x)$. Even if this result is not proven rigorously for the asymptotic regime, this is exactly what we observed numerically in Faranda *et al.* [2011b].
By applying the observable g_i we just remodulate this piecewise linear $F(x)$ but it is clear that we do not obtain any kind of convergence to the GEV distribution. In terms of density functions, we can observe multimodal distributions and the number of modes and their positions are highly dependent on both n, m and initial conditions Faranda *et al.* [2011b].
In the case of pure periodic motion this kind of extreme value distribution must asymptotically be a Dirac delta as we pick up always the same Y_i.

3. The numerical algorithm

As we have already said to observe a GEV distribution of maxima orbits must satisfy D' and D_2 conditions or an exponential decay of the HTS whereas for periodic or quasi-periodic motions we have different distributions but never a GEV. This gives us a way to discriminate the kind of motion simply looking at the extreme value statistics. From a practical point of view we can introduce a simple algorithm to perform this task:

(1) Compute the orbit of the dynamical system for k iterations.
(2) Compute the series $X_m(x) = g(\mathrm{dist}(f^m(x), \zeta))$ where ζ is the initial condition that must be chosen on the attractor [Faranda *et al.*, 2011a].
(3) Divide the series in n bins each containing m data.

(4) Take the maximum in each bin and test if this empirical distribution gives the parameters expected by the theory.

In the next section we will describe how to use it operationally in a meaningful example of a dynamical systems that present coexistence of regular and chaotic motions and therefore allow to test the validity of our indicator: the Standard map. Before presenting the results, we need to clarify the numerical inference procedure that we use to obtain the parameters of the GEV distribution. In Faranda *et al.* [2011b] we have used a Maximum Likelihood Estimation (MLE) procedure working both on pdf and cdf, since our distributions were absolutely continuous and the minimization procedure was well defined. In a general case, when we are dealing also with singular measures, we may have a cdf which is not anymore absolutely continuous and consequently the fitting procedure via MLE could give wrong results [Faranda *et al.*, 2011a]. To avoid these problems we have used an L-moments estimation as detailed in Hosking [1990]. This procedure is completely discrete and can be used both for absolutely continuous or singular continuous cdf. The L-moments are summary statistics for probability distributions and data samples. They are analogous to ordinary moments which mean that they provide measures of location, dispersion, skewness, kurtosis, but are computed from linear combinations of the data values, arranged in increasing order (hence the prefix L). Asymptotic approximations to sampling distributions are better for L-moments than for ordinary moments [Hosking [1990], Figure 4]. The relationship between the moments and the parameters of the GEV distribution is described in Hosking [1990], while the 95% confidence intervals has been derived using a bootstrap procedure.
We have explained in Faranda *et al.* [2011a] the issues of applying directly a Kolmogorov Smirnov or a Chi-square test to see if the data really belongs to the GEV distribution when the measure is singular: even if both tests fail the GEV model is reasonable as it is the closest continuous representation of the empirical discrete distribution we get.
Therefore, instead of using as a dynamical indicator the goodness of the fit to the continuous GEV model, we will check the deviations of the parameters with respect to the theoretically expected values for chaotic orbits.

Our results will be studied against the Divergence of two nearby trajectories and the Reversibility error. In Faranda *et al.* [2011c] we have shown that these indicators give insights into the structure of a dynamical system as well as others such as SALI, MEGNO and mLCE with which these indicators have been compared.
We briefly recall here the definitions and we refer to Faranda *et al.* [2011c] for further clarifications.

The arithmetic operations such as sums or multiplications imply a round off, which propagates the error affecting each number. Round off algebraic procedures are hardware dependent as detailed in Knuth [1973]. Unlike the case of stochastic perturbations, the error strongly depends on x. Suppose we are given a map $f^t(x)$ then we will indicate with $f^t_*(x)$ the correspondent numerical map both evaluated at the t-th iteration. The Divergence of orbits is defined as:

$$\Delta_t = \mathrm{dist}(f^t_S(x), f^t_D(x)), \tag{17}$$

where f^t_S and f^t_D stand for single and double precision iterations respectively, and 'dist' is a suitable metric.

If the map is invertible we can also define the Reversibility error as

$$R_t = \mathrm{dist}(f_*^{-t} \circ f^t_*(x), x) \tag{18}$$

which is non zero since the numerical inverse f_*^{-1} of the map is not exactly the inverse of f_* namely $f_*^{-1} \circ f_*(x) \neq x$. Obviously the reversibility error is much easier to compute than the divergence of orbits (if we know explicitly the inverse map) and the information it provides is basically the same as the latter. Both quantities give an average linear growth for a regular map together with an exponential growth for a chaotic map having positive Lyapounov exponents and strong mixing properties. When computing R_t we will set $f_* = f_S$ in order to compare with Δ_t.

4. A Case-study: the Standard Map

The Standard map (also known as the Chirikov-Taylor map or the Chirikov standard map) is an area-preserving chaotic map defined on the bidimensional torus. It can be thought as a stick that is free of the gravitational force, which can rotate frictionless in a plane around an axis located in one of its tips, and which is periodically kicked on the other tip. This mechanical system is usually called a kicked rotator. It is defined by:

$$\begin{cases} y_{t+1} = y_t - \frac{K}{2\pi}\sin(2\pi x_t) & \mod 1 \\ x_{t+1} = x_t + y_{t+1} & \mod 1 \end{cases} \tag{19}$$

Standard map is one of the most widely-studied examples of dynamical chaos in physics. It can be regular or chaotic, depending on the strength of the impulses: stronger kicks lead to chaotic behaviors. The variables y and x respectively represent the angular position and angular momentum of the stick at the t-th kick.
For $K << 1$ the motion follows quasi periodic orbits for all initial conditions, whereas if $K >> 1$ the motion turns to be chaotic and irregular. An interesting behavior is achieved when $K \sim 1$: in this case we have coexistence of regular and chaotic motions depending on the initial conditions chosen.
Buric *et al.* [2003] studied the HTS of the standard map for different values of the parameter K showing that if $K >> 1$, then $H(t) = e^{-t}$, whereas a power law decay is obtained if $K << 1$. In the intermediate regime $K \sim 1$, the HTS exhibits a superposition of exponential and power law decay. Buric *et al.* [2003] proved analytically the existence of the HTS for the Standard Map and checked it numerically for different value of the parameter K. Following this paper in figure 1 we report the HTS for certain parameter values that we use in the numerical simulations: as we have already said Freitas *et al.* [2009] proved that a GEV distribution as Extreme Value statistics can only be obtained when the HTS is exponential, therefore we will refer to these results to check the consistency between the results provided by HTS and the GEV based indicators. To evaluate the HTS, we use the set $A = [0, \epsilon] \times [0, \epsilon]$ with $\epsilon = 0.01$ but the behavior is qualitatively the same when we change the location of the set A.
First of all, for the Extreme value analysis we have taken an ensemble of 500 initial conditions centered around x_0, y_0 in a small subset of the bidimensional torus. After iterating the map for $k = 10^6$ iterations, we have selected $n = 10^3$ maxima each in $m = 10^3$ observations. These values are in agreement with what emerged from the numerical study presented in Faranda *et al.* [2011b] and guarantee a reliable statistics when chaotic orbits are considered. Once computed the parameters of GEV distribution for each realization, we have averaged them over the different initial conditions checking their convergence towards the expected theoretical values when K is varied from $K = 10^{-4}$ up to $K = 10^2$. As indicators, looking at the equations 14-16, we have selected the shape parameters for the three type observables $\xi'(g_1), \xi'(g_2), \xi'(g_3)$ and the scale parameter for the type 1 observable $\sigma(g_1)$. From all the parameters set, these are the ones that have a dependence on the local dimension of the attractor d but do not have a dependence of m, therefore, once we are in the asymptotic regime, the results are independent on the number of observations in each bin. Results are presented in figure 2. In this example we set $\alpha = 3$ for the observables g_2 and g_3 . For each parameter, the averaged value is represented with a solid line whereas the dotted lines represent one standard deviations of the ensemble. It is clear that for $K > 1$ the parameters converge towards the theoretical values whereas for regular motions, the computed parameters are not representative of a GEV distribution and exhibit a spread that is more than five times bigger with respect to the chaotic counterpart. The results are similar if we change the initial conditions and the value of α.

Let us now fix the value of $K = 6.5$. We want to show that we can depict the structure of the Standard map with the indicators presented above starting from 500×500 initial conditions uniformly distributed on the bidimensional torus. The number of iterations $k = 10^6$, $n = m = 10^3$ and $\alpha = 3$ are fixed as before. Results are shown in figure 3 where we compare the four parameters of GEV distribution (top and middle panels) with the Reversibility Error and Divergence of orbits in logarithm scale (lower panel). The number of iterations for the round off indicators is $t = 100$. It is evident that the structure of the

Standard Map is well highlighted by all the indicators based on GEV distribution. For g_1 the empirical values agree with the theoretical ones $\xi' = 0$ and $\sigma = 1/2$ expected in the chaotic regions, whereas in the small regular islands we observe significant deviation from the expected values. Similar results hold for g_2 and g_3 for which the expected theoretical values are $\xi' = 1/6$ and $\xi' = -1/6$ respectively. The Round off based indicators highlight the same structure: in the chaotic sea the exponential growth of these quantities lead to a saturation at a value comparable with the size of the torus, whereas in the regular region their values remain order of magnitude lower. For this value of the parameter the HTS - shown in figure 1 - shows no significant deviations from the exponential behavior: this means that the chaotic part strongly dominates the dynamics with the exceptions of the small stable islands highlighted in figure 3. All the indicators are comparable in terms of speed of convergence towards the theoretically expected values.

Eventually, we repeat the same analysis changing only the parameter K and fixing it to the value $K = 4.5$. In this case the HTS is a superposition of an exponential statistics and a power law: this is due to the fact that the structure of the orbits in the standard map is now quite different with the presence of major regions of stable orbits surrounded by a layer of variable thickness where Cantori are present. Orbits starting in the chaotic sea intersects this region with evident effects not only on the HTS behavior but also on the statistic of extremes: even if the HTS statistics for $K = 4.5$ is still dominated by an exponential component, its decay is clearly slower if compared with the $K = 6.5$, $K = 10$ cases, all shown in figure 1. This corresponds also to a slower convergence of the partial maxima series to the GEV distribution: choosing again $n = m = 10^3$, the convergence of the parameters to the theoretical ones is slower with respect to the case analysed before for $K = 6.5$. Nonetheless, although slower, the HTS decay is still exponential so a better convergence must be observed if we increase the observations in each bin to $m = 10^4$ as the selection of the extreme values is less affected by the regular motions: this is exactly what is shown in figure 4: in the top panel $\xi'(g_1)$ is compared for the case $m = 10^3$ (Left-hand side) and $m = 10^4$ (Right-hand side): clearly the convergence to the expected theoretical values $\xi' = 0$ improves when the number of observations in each bin increases. Similar results hold for the others GEV parameters: as a further example $\xi'(g_2)$ is presented in the middle panel of figure 4. The structure of the map is again well highlighted by GEV indicators if we compare them with the logarithm of R_t, Δ_t shown in the lower panel for $t = 100$. For this value of t the effect of the superimposition of regular and chaotic it is not appreciable for both the Reversibility error and the Divergence of the orbits as in the chaotic sea they already saturates at a value comparable with the size of the torus.

5. Conclusions

Nowadays there exists a large family of dynamical indicators of stability that can be used to obtain informations about the stability of orbits in dynamical systems. Nonetheless, the numerical algorithm to compute many of them can be computationally expensive as it usually require the solution of variational equations or a Monte Carlo simulation. Moreover each indicator is specialized in detecting differences either in the chaotic regions or in the regular ones. In this framework we have introduced the parameters of the GEV distribution for some special class of observables to show that they effectively work as other dynamical indicators. As shown in Freitas *et al.* [2009], Faranda *et al.* [2011b] and Faranda *et al.* [2011a] a GEV like distribution can be observed only if the dynamical system satisfies certain mixing conditions or exhibit an exponential decay of the Hitting Time Statistics.
We selected the parameters that do not depend explicitly on the number of observations in each bin m, so that our results only depend on the local dimension of the attractor d. We illustrated the effectiveness of the indicators testing them on the Standard Map. First we showed that, varying the parameters K that regulate the chaoticity of the map, the parameters of the GEV fitted distribution for a small ensemble of initial conditions approach the theoretical values when $K >> 1$ that corresponds to chaotic motions. We have also showed numerically that GEV parameters are able to distinguish regular islands in chaotic sea for Standard Map parameter $K = 6.5$ and $K = 4.5$. In the latter case we have experienced the effect of superimposition of regular and chaotic motions that prevented us from obtaining the theoretically expected parameters for the chaotic regions when m was not big enough.

To conclude, the GEV parameters provide basically the same informations as other indicators such as the Divergence of orbits and the Reversibility error and they are easily accessible from a computational view point as they do not require the solution of the variational equations. With respect to the other indicators they highlight other relevant informations on the dynamics: the parameters are dependent on the local dimension of the attractor that can be derived from a fit to the empirical distribution once we are sure that we are dealing with a chaotic orbit. Furthermore, the extreme value distribution is itself interesting from both practical and theoretical point of view: as far as the observable are concerned it gives detailed informations about the closest return near a certain initial condition.

6. Aknowledgements

S.V. was supported by the CNRS-PEPS Project *Mathematical Methods of Climate Models*, and he thanks the GDRE Grefi-Mefi for having supported exchanges with Italy. DF and VL acknowledge the financial support of the EU FP7-ERC project NAMASTE Thermodynamics of the Climate System. DF acknowledges Martin Mestre for the support with the numerical simulations, Peter de Boeck and Jeroen Wouters for the useful comments and suggestions that improved the quality of the paper.

References

Balakrishnan, V., Nicolis, C. & Nicolis, G. [1995] "Extreme value distributions in chaotic dynamics," *J. Stat. Phys.* **80**, 307–336.

Buric, N., Rampioni, A. & Turchetti, G. [2005] "Statistics of Poincaré recurrences for a class of smooth circle maps," *Chaos Soliton. Fract.* **23**, 1829–1840.

Buric, N., Rampioni, A., Turchetti, G. & Vaienti, S. [2003] "Weak chaos and poincaré recurrences for area preserving maps," *J. Phys A-Math.Gen.* **36**, L209.

Cincotta, P. M., Giordano, C. M. & Simó, C. [2003] "Phase space structure of multi-dimensional systems by means of the mean exponential growth factor of nearby orbits," *Physica D* **182**, 151–178.

Coles, S., Heffernan, J. & Tawn, J. [1999] "Dependence measures for extreme value analyses," *Extremes* **2**, 339–365.

Collet, P. [2001] "Statistics of closest return for some non-uniformly hyperbolic systems," *Ergod. Theor. Dyn. Syst.* **21**, 401–420.

Faranda, D., Lucarini, V., Turchetti, G. & Vaienti, S. [2011a] "Extreme value distribution for singular measures," *Arxiv preprint arXiv:1106.2299* .

Faranda, D., Lucarini, V., Turchetti, G. & Vaienti, S. [2011b] "Numerical convergence of the block-maxima approach to the Generalized Extreme Value distribution," *J. Stat. Phys.* , 1–25.

Faranda, D., Mestre, M. & Turchetti, G. [2011c] "Analysis of round off errors with reversisibility test as a dynamical indicator," *Accepted for publication: Int. Jou. Bif. Chaos* .

Fisher, R. & Tippett, L. [1928] "Limiting forms of the frequency distribution of the largest or smallest member of a sample," *Proceedings of the Cambridge philosophical society*, p. 180.

Freitas, A. & Freitas, J. [2008] "On the link between dependence and independence in extreme value theory for dynamical systems," *Stat. Probab. Lett.* **78**, 1088–1093.

Freitas, A., Freitas, J. & Todd, M. [2009] "Hitting time statistics and extreme value theory," *Probab. Theory Relat. Fields* , 1–36.

Freitas, A., Freitas, J. & Todd, M. [2010] "Extremal Index, Hitting Time Statistics and periodicity," *Arxiv preprint arXiv:1008.1350* .

Freitas, A., Freitas, J. & Todd, M. [2011] "Extreme value laws in dynamical systems for non-smooth observations," *J. Stat. Phys.* , 1–19.

Gao, J. B. [1999] "Recurrence time statistics for chaotic systems and their applications," *Phys. Rev. Lett.* **83**, 3178–3181.

Ghil et al., M. [2011] "Extreme events: Dynamics, statistics and prediction," *Nonlinear Process Geophys.* **18**, 295–350.

Gnedenko, B. [1943] "Sur la distribution limite du terme maximum d'une série aléatoire," *Ann. Math.* **44**, 423–453.

Goździewski, K., Bois, E., Maciejewski, A. J. & Kiseleva-Eggleton, L. [2001] "Global dynamics of planetary systems with the MEGNO criterion," *Astron. Astrophys.* **378**, 569–586.

Gupta, C., Holland, M. & Nicol, M. [2009] "Extreme value theory for dispersing billiards and a class of hyperbolic maps with singularities." *preprint* .

Haiman, G. [2003] "Extreme values of the tent map process," *Stat. Probab. Lett.* **65**, 451–456.

Hirata, M., Saussol, B. & Vaienti, S. [1999] "Statistics of Return Times: A General Framework and New Applications," *Comm. Math. Phys.* **206**, 33–55.

Hosking, J. [1990] "L-moments: analysis and estimation of distributions using linear combinations of order statistics," *J. Roy. Stat. Soc. B Met.* **52**, 105–124.

Hu, H., Rampioni, A., Rossi, L., Turchetti, G. & Vaienti, S. [2004] "Statistics of Poincaré recurrences for maps with integrable and ergodic components," *Chaos* **14**, 160.

Kac, M. [1934] "On the notion of recurrence in discrete stochastic processes," *Proc. Nat. Acad. Sci. USA* **20**, 376–379.

Knuth, D. E. [1973] *The art of computer programming*, Vol. 2 (Addison-Wesley (Reading, MA)).

Laskar, J. [1999] "Introduction to frequency map analysis," *Hamiltonian systems with three or more degrees of freedom*, pp. 134–150.

Liverani, C., Marie, P. & Vaienti, S. [2007] "Random Classical Fidelity," *J. Stat. Phys.* **128**, 1079.

Nicolis, C., Balakrishnan, V. & Nicolis, G. [2006] "Extreme events in deterministic dynamical systems," *Phys. Rev. Lett.* **97**, 210602.

Pickands III, J. [1968] "Moment convergence of sample extremes," *Ann. Math. Stat.* **39**, 881–889.

Robutel, P. & Laskar, J. [2001] "Frequency Map and Global Dynamics in the Solar System I:: Short Period Dynamics of Massless Particles," *Icarus* **152**, 4–28.

Rosenstein, M. T., Collins, J. J. & De Luca, C. J. [1993] "A practical method for calculating largest Lyapunov exponents from small data sets," *Physica D* **65**, 117–134.

Skokos, C. [2010] "The lyapunov characteristic exponents and their computation," *Lect. Notes Phys., Berlin Springer Verlag*, pp. 63–135.

Skokos, C., Antonopoulos, C., Bountis, T. & Vrahatis, M. [2002] "Smaller alignment index (SALI): Detecting order and chaos in conservative dynamical systems," *Proc. of 4th GRACM Congress on Computational Mechanics* (Citeseer).

Skokos, C., Antonopoulos, C., Bountis, T. C. & Vrahatis, M. [2004] "Detecting order and chaos in Hamiltonian systems by the SALI method," *J. Phys. A-Math. Gen.* **37**, 6269.

Skokos, C., Bountis, T. C. & Antonopoulos, C. [2007] "Geometrical properties of local dynamics in Hamiltonian systems: The Generalized Alignment Index (GALI) method," *Physica D* **231**, 30–54.

Turchetti, G., Vaienti, S. & Zanlungo, F. [2010] "Relaxation to the asymptotic distribution of global errors due to round off," *Europhys. Lett.* **89**, 40006.

Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. [1985] "Determining Lyapunov exponents from a time series," *Physica D* **16**, 285–317.

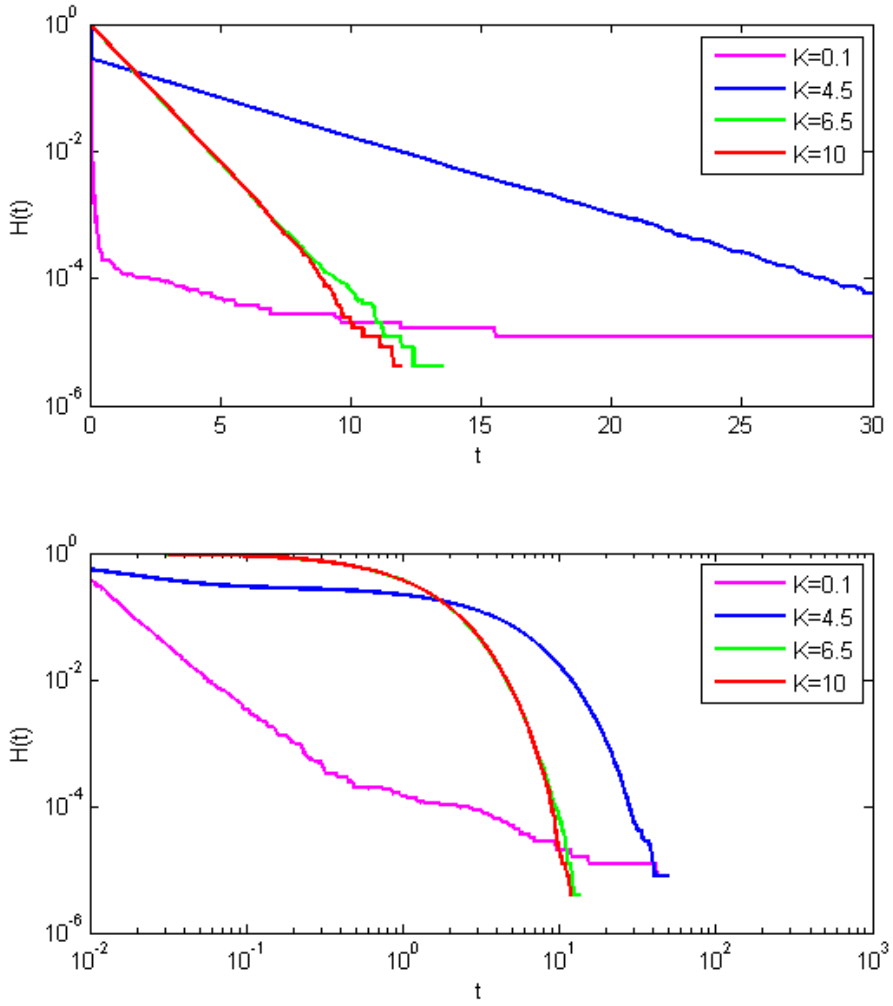


Fig. 1. Hitting time statistics on the Standard Map for the set $A = [0, \epsilon] \times [0, \epsilon]$ for different K, $\epsilon = 0.01$, in semilog scale (upper panel) and in log-log scale (lower panel).

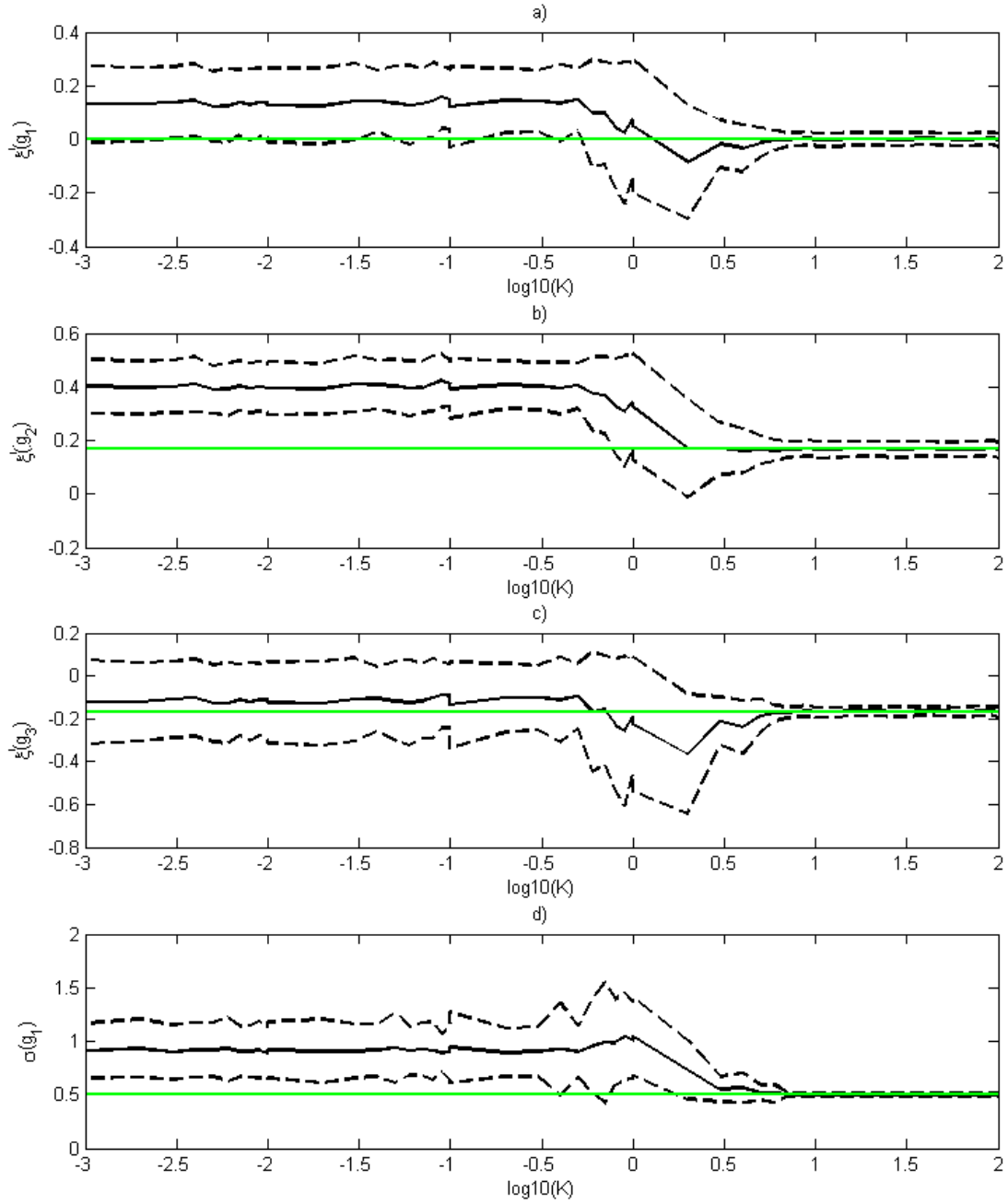


Fig. 2. Standard map: GEV parameters averaged over 500 different initial condition centered in $x_0 = 0.305, y_0 = 0.7340$ VS K. a) $\xi'(g_1)$, b) $\xi'(g_2)$, c) $\xi'(g_3)$, d) $\sigma(g_1)$.

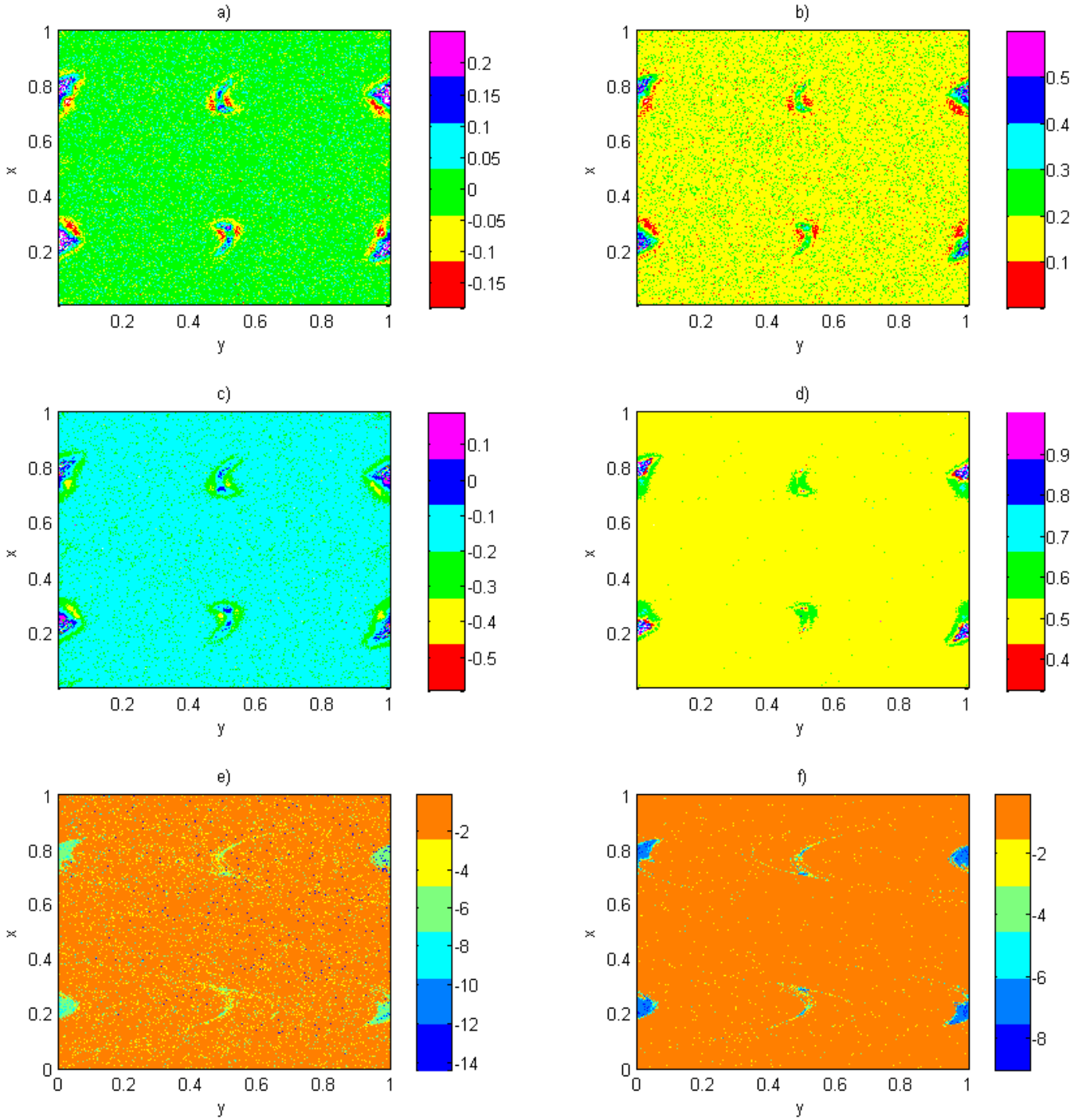


Fig. 3. Structure of the Standard Map for $K = 6.5$. **a)** $\xi'(g_1)$ for $m = n = 10^3$, theoretical value for chaotic orbits $\xi'(g_1) = 0$; **b)** $\xi'(g_2)$ for $m = n = 10^3$, theoretical value $\xi'(g_2) = 1/6$; **c)** $\xi'(g_3)$ for $m = n = 10^3$, theoretical value $\xi'(g_3) = -1/6$; **d)** $\sigma(g_1)$ for $m = n = 10^3$, theoretical value $\sigma(g_1) = 1/2$; **e)** $\log_{10}(R_{t=100})$; **f)** $\log_{10}(\Delta_{t=100})$

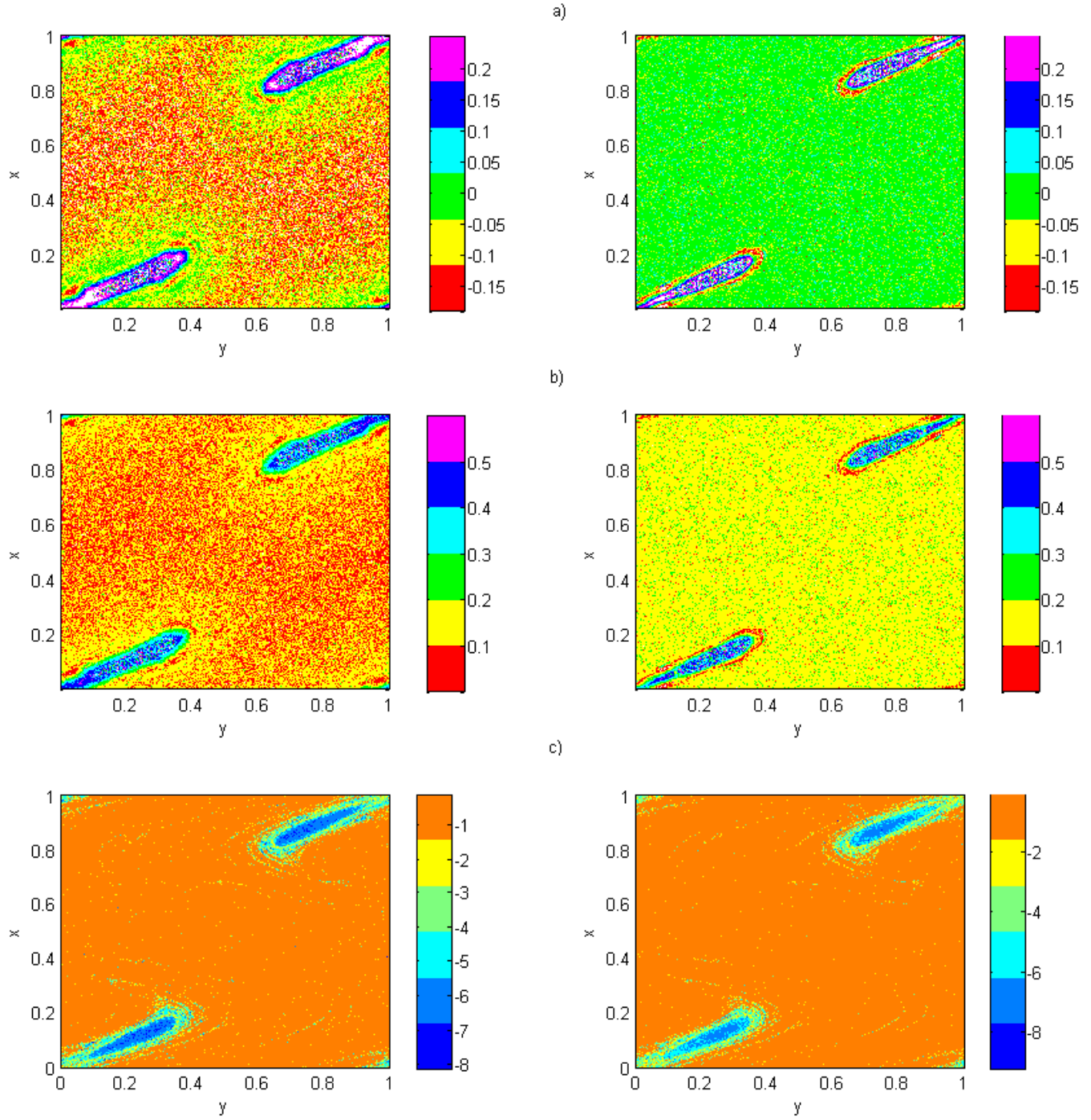


Fig. 4. Structure of the Standard Map for $K = 4.5$. **a)** Left-hand side: $\xi'(g_1)$ for $m = n = 10^3$, Right-hand side: $\xi'(g_1)$ for $m = 10^4, n = 10^3$, in both cases the theoretical value for chaotic orbits is $\xi'(g_1) = 0$; **b)** Left-hand side: $\xi'(g_2)$ for $m = n = 10^3$, Right-hand side: $\xi'(g_2)$ for $m = 10^4, n = 10^3$, theoretical value: $\xi'(g_2) = 1/6$; **c)** Left-hand side: $\log_{10}(R_{t=100})$, Right-hand side: $\log_{10}(\Delta_{t=100})$

Chapter 5

Article: Extreme value theory for singular measures

CHAOS **22**, 023135 (2012)

Extreme value theory for singular measures

Valerio Lucarini,[1,2,a)] Davide Faranda,[1,2,b)] Giorgio Turchetti,[3,c)] and Sandro Vaienti[4,d)]
[1]*Klimacampus, Institute of Meteorology, University of Hamburg, Grindelberg 5, 20144 Hamburg, Germany*
[2]*Department of Mathematics and Statistics, University of Reading, Whiteknights, P.O. Box 220, Reading RG6 6AX, United Kingdom*
[3]*Department of Physics, University of Bologna, INFN-Bologna Via Irnerio 46, Bologna, 40126, Italy*
[4]*UMR-6207, Centre de Physique Théorique, CNRS, Universités d'Aix-Marseille I,II, Université du Sud Toulon-Var and FRUMAM (Fédération de Recherche des Unités de Mathématiques de Marseille), CPT, Luminy, Case 907, 13288 Marseille Cedex 09, France*

(Received 11 June 2011; accepted 1 May 2012; published online 13 June 2012)

In this paper, we perform an analytical and numerical study of the extreme values of specific observables of dynamical systems possessing an invariant singular measure. Such observables are expressed as functions of the distance of the orbit of initial conditions with respect to a given point of the attractor. Using the block maxima approach, we show that the extremes are distributed according to the generalised extreme value distribution, where the parameters can be written as functions of the information dimension of the attractor. The numerical analysis is performed on a few low dimensional maps. For the Cantor ternary set and the Sierpinskij triangle, which can be constructed as iterated function systems, the inferred parameters show a very good agreement with the theoretical values. For strange attractors like those corresponding to the Lozi and Hènon maps, a slower convergence to the generalised extreme value distribution is observed. Nevertheless, the results are in good statistical agreement with the theoretical estimates. It is apparent that the analysis of extremes allows for capturing fundamental information of the geometrical structure of the attractor of the underlying dynamical system, the basic reason being that the chosen observables act as magnifying glass in the neighborhood of the point from which the distance is computed. [http://dx.doi.org/10.1063/1.4718935]

The existence of extreme value laws for dynamical systems preserving an absolutely continuous invariant measure or a singular continuous invariant measure has been recently proven if strong mixing properties or exponential hitting time statistics on balls are satisfied. In this context, in a previous work, we have proposed an algorithmic way to study extrema by using a block-maxima approach for the time series of specific observables of dynamical systems possessing an absolutely continuous invariant measure. Such observables boil down to functions of the distance of the trajectory of the dynamical system from a given point of the attractor. In this work, we test our algorithm for maps that do not have an absolutely continuous invariant measure and show that the parameters of the cumulative distribution function of maxima, which is in all cases described by a member of the family of the generalised extreme value (GEV) distributions, can be related to the scaling exponent of the measure of a ball centered around the point from which the distance is computed. Such a scaling exponent turns out to be the Hausdorff dimension of the measure (also known as information dimension). Even if we cannot estimate analytically the asymptotic behavior of the measure of the balls, the agreement with the numerical simulations we have carried out for different maps suggests the validity of our proposed scaling in terms of the information dimension. Our conjecture has been tested with numerical experiments on different low dimensional maps such as the Cantor ternary set, the Sierpinskij triangle, iterated function system (IFS) with non-uniform weights and strange attractors such as Lozi and Hénon. In all cases considered, there is a good agreement between the theoretical parameters and the inferred ones although, in the case of strange attractors which exhibit multifractal structures, the convergence is slower. Extremes can be thought as geometric indicators of the local properties of the attractor. As a side note, we emphasize that the statistical inference of the GEV parameters has been performed using the L-moments procedure, which, as opposed to more common maximum likelihood method, allows for overcoming the problem of dealing with a singular continuous invariant measure.

I. INTRODUCTION

A. Classical extreme value theory

The extreme value Theory (EVT), first developed by Fisher and Tippett (1928) and formalized by Gnedenko (1943) with the goal of studying the maxima of series of independent and identically distributed stochastic variables, has progressively attracted a wider and wider interest in

[a)]Electronic mail: valerio.lucarini@zmaw.de.
[b)]Electronic mail: davide.faranda@zmaw.de.
[c)]Electronic mail: turchett@bo.infn.it.
[d)]Electronic mail: vaienti@cpt.univ-mrs.fr.

many different scientific fields and in many practical applications as robust statistical model for understanding and possibly forecasting events that occur with very small probability but can be extremely relevant in terms of impacts (Coles, 2001). We mention the case of floods (Gumbel, 1941; Sveinsson and Boes, 2002; and Friederichs and Hense, 2007), insurance losses (Brodin and Kluppelberg, 2006 and Cruz, 2002); earthquakes (Sornette *et al.*, 1996; Cornell, 1968; and Burton, 1979); meteorological and climate events (Smith, 1989; Katz and Brown, 1992; Felici *et al.*, 2007; Vitolo *et al.*, 2009b; and Altmann *et al.*, 2006). An extensive review of the techniques and applications related to the EVT is presented in Ghil *et al.* (2011). The power of the EVT and its seamless applicability in a large variety of fields lies in the fact that it defines in a universal way what an extreme is, as opposed to adopting somewhat subjective definitions of an extreme as something which is *very large*. The actual analysis of the extremes of a data series is usually performed by adopting either one of two conceptually complementary strategies, the peak-over-threshold (POT) or the block-maxima (BM) approach. The former consists in looking at exceedances over asymptotically higher and higher thresholds (Todorovic and Zelenhasic, 1970) and takes into consideration the generalized Pareto distribution (GPD) as statistical model (Smith, 1984; Davison, 1984; and Davison and Smith, 1990). The BM approach represents a very natural way to look at extremes realized within fixed time intervals: it is based on dividing the data series of some observable into bins of equal length and on selecting the maximum (or the minimum) value in each of them (Coles, 2001; Felici *et al.*, 2007; Katz and Brown, 1992; Katz, 1999; and Katz *et al.* 2005). In this case, the statistical model for the BM is the GEV distribution (Coles, 2001). A strong connection exists between the two methodologies, as we have that if BM obeys the GEV distribution, then exceedances over some high threshold will have an associated GPD, and, moreover, the parameters of the GEV and GPD feature a very close connection. As a result, several practical methods (e.g., Hill and Pickands estimators) (Leadbetter *et al.*, 1983) developed for estimating the parameters of the GEV distribution of the extremes of a given time series are actually based upon comparing the GPD fits at various thresholds.

In this paper, we will focus on the GEV approach and refer to it when discussing EVT. Let us first briefly review the Gnedenko (1943) fomalisation of the theory of extremes. We study the convergence of the maxima of i.i.d. variables

$$X_0, X_1, ..X_{m-1}$$

by looking at the cumulative distribution function (cdf) $F(x)$ of the form

$$F(x) = P\{a_m(M_m - b_m) \leq x\},$$

where a_m and b_m are normalizing sequences and $M_m = \max\{X_0, X_1, ..., X_{m-1}\}$. The previous definition of the cdf may be rewritten as $F(u_m) = P\{M_m \leq u_m\}$, where $u_m = x/a_m + b_m$. Under general hypotheses on the nature of the parent distribution of the X_j's, Gnedenko (1943) showed that the cdf of the maxima, up to an affine change of variable, adheres to one of the following three laws:

- Type 1 (Gumbel).

$$E(x) = \exp(-e^{-x}), -\infty < x < \infty \tag{1}$$

- Type 2 (Fréchet).

$$E(x) = \begin{cases} 0, x \leq 0 \\ \exp(-x^{-\xi}), \text{ for some } \xi > 0, x > 0 \end{cases} \tag{2}$$

- Type 3 (Weibull).

$$E(x) = \begin{cases} \exp(-(-x)^{-\xi}), \text{ for some } \xi < 0, x \leq 0 \\ 1, x > 0. \end{cases} \tag{3}$$

Defining the right endpoint x_F of a cdf $F(x)$ as

$$x_F = \sup\{x : F(x) < 1\}, \tag{4}$$

it is possible to compute the normalizing sequences a_m and b_m using the following corollary:

Corollary (Gnedenko): The normalizing sequences a_m and b_m to be used to achieve the convergence of the cdf normalized maxima $P\{a_m(M_m - b_m) \leq x\} \to F(x)$ as $m \to \infty$ may be taken as follows:

- *Type 1:* $a_m = [G(\gamma_m)]^{-1}, \quad b_m = \gamma_m$;
- *Type 2:* $a_m = \gamma_m^{-1}, \quad b_m = 0$;
- *Type 3:* $a_m = (x_F - \gamma_m)^{-1}, \quad b_m = x_F$;

where

$$\gamma_m = F^{-1}(1 - 1/m) = \inf\{x; F(x) \geq 1 - 1/m\}, \tag{5}$$

and

$$G(t) = \int_t^{x_F} \frac{1 - F(u)}{1 - F(t)} du, \quad t < x_F. \tag{6}$$

In (Faranda *et al.*, 2011a) we have shown that this approach is equivalent to fitting the distribution of unnormalized values of M_m to the GEV distribution, whose cdf can be written as

$$F_G(x; \mu, \sigma, \xi') = \exp\left\{-\left[1 + \xi'\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi'}\right\}, \tag{7}$$

which holds for $1 + \xi'(x - \mu)/\sigma > 0$. For all values of m, the fitted values of $\mu \in \mathbb{R}$ (location parameter) and $\sigma > 0$ (scale parameter) are related to the scaling constants b_m, and a_m as follows:

$$\mu = b_m, \quad \sigma = \frac{1}{a_m}.$$

Instead, $\xi' \in \mathbb{R}$ is the shape parameter, also known as the tail index, and is related to the parameter ξ given above as $\xi' = 1/\xi$, except for the limiting case $\xi' = 0$. When $\xi' \to 0$,

we have a type 1 (Gumbel) distribution, when ξ' is positive, we have a type 2 (Fréchet) distribution, and, when ξ' is negative, we have a type 3 (Weibull) distribution.

B. Extreme value theory for dynamical systems

The classical EVT lays the foundation for constructing and inferring the statistical properties of the extremes of time series generated as a result of stochastic processes. Obviously, it is of crucial relevance, for both mathematical reasons and for devising a working framework to be used in applications, to understand under which circumstances the time series of observables of deterministic dynamical systems can be treated using the EVT. Empirical studies show that in some cases, the extremes of dynamical observables of chaotic systems obey up to a high degree of precision the GEV statistics, even if it is apparent that the asymptotic convergence is highly dependent on the considered observables (Felici *et al.*, 2007; Vannitsem 2007; and Vitolo *et al.*, 2009a; 2009b). Instead, Balakrishnan *et al.* (1995) and, more recently, Nicolis *et al.* (2006) showed that when considering a dynamical system featuring a regular (periodic or quasi-periodic) motion, the extremes of a generic dynamical observable do not obey any statistics compatible with those of GEV distributions.

The first rigorous mathematical approach aimed at extending EVT in the context of dynamical systems goes back to the pioneering paper (Collet, 2001). Along this line, relevant contributions have subsequently come from Freitas and Freitas (2008), Freitas *et al.* (2010a; 2010b), and Gupta *et al.* (2011). The bottom line of all of these investigations is based upon constructing from the stationary stochastic process given by the dynamical system and by the chosen observable a new stationary independent sequence whose maxima can be proved to obey one of the classical three extreme value laws. Subsequently, it is obtained that the original sequence of maxima obeys the same extreme value law.

We briefly recapitulate this approach and introduce the notation relevant for our analysis. We consider a dynamical system $(\Omega, \mathcal{B}, \nu, f)$, where Ω is the invariant set in some manifold, usually $\mathbb{R}^d$, $\mathcal{B}$ is the Borel σ-algebra, $f : \Omega \to \Omega$ is a measurable map, and ν is a probability f-invariant Borel measure. The stationary stochastic process given by the dynamical system is of the form $X_m = g \circ f^m$, for any $m \in \mathbb{N}$, where the observable g has values in $\mathbb{R} \cup \pm\infty$ and achieves a global maximum at the point $\zeta \in \Omega$. We therefore study the partial maxima $M_m = \max\{X_0, \ldots, X_{m-1}\}$, and, in particular, we look for the normalising sequences $\{a_m\}, \{b_m\} \in \mathbb{R}^+, m \in \mathbb{N}$ such that $\nu\{x; a_m(M_m - b_m) \le t\} = \nu\{x; M_m \le u_m\} = \nu(M_m \le u_m)$ converges to a non-degenerate distribution function. Here $u_m = \frac{t}{a_m} + b_m$ is such that $m\nu(X_0 > u_m) \to \tau$, for some positive τ depending eventually on t. We refer to (Leadbetter *et al.*, 1983) for a clear and complete description of this construction.

Our goal is now to associate to our process a new i.i.d. sequence $\tilde{X}_0, \ldots, \tilde{X}_{m-1}$, whose cdf is the same as that of X_0 and consider the partial maxima $\tilde{M}_m = \max\{\tilde{X}_0, \ldots, \tilde{X}_{m-1}\}$. The idea is to prove that, after suitable normalization, the cdf of such a maximum converges to one of the three laws in Eqs. (1)–(3). In this case, the cdf of the original partial maxima M_m obeys Eqs. (1)–(3), if we are able to prove that

$$\lim_{m\to\infty} \nu(\tilde{M}_m \le u_m) = \lim_{m\to\infty} \nu(M_m \le u_m).$$

This construction is successful if the original stochastic process X_m obeys two sufficient conditions called D_2 and D'. (We briefly state here the two. If $X_m, m \ge 0$ is the considered stochastic process, we can define $M_{j,l} \equiv \max\{X_j, X_{j+1}, \ldots, X_{j+l}\}$ and we put $M_{0,m} = M_m$. The condition $D_2(u_m)$ holds for the sequence X_m if for all integers l, t, m we have $|\nu(X_0 > u_m, M_{t,l} \le u_m) - \nu(X_0 > u_m)\nu(M_{t,l} \le u_m)| \le \gamma(m,t)$, where $\gamma(m,t)$ is non-increasing in t for each m and $m\gamma(m,t_m) \to 0$ as $m \to \infty$ for some sequence $t_m = o(m), t_m \to \infty$. Instead, the condition $D'(u_m)$ holds for the sequence X_m if $\lim_{l\to\infty} \limsup_m m \sum_{j=1}^{[m/l]} \nu(X_0 > u_m, X_j > u_m) = 0$. These conditions basically require a sort of independence of the stochastic dynamical sequence in terms of uniform mixing condition on the distribution functions. In particular, the condition D_2, introduced in its actual form by Freitas and Freitas (2008), can be checked directly by estimating the rate of decay of correlations for general Hölder observables.

Another crucial aspect of the sequence of papers mentioned above is the careful selection of the g observables. The observables are expressed as functions $g = g(\mathrm{dist}(x, \zeta))$, whose argument is the Euclidean distance of x from a given point ζ. Moreover, the functional shape of the g's is such that g achieves a global maximum at $x = \zeta$ for almost all points $\zeta \in \Omega$. It is especially convenient to define three classes of observables g_1, g_2, g_3. The g_i's, $i = 1, 2, 3$ observables are constructed so that they ensure the existence of a non-degenerate limit distribution for the partial maxima (Freitas *et al.*, 2010b and Holland *et al.*, 2012) and, at the same time, lead to establishing extreme value laws of type i, respectively, for the normalised sequences of maxima. This will be described accurately in Sec. II.

Recently, a major advance in this field has been obtained through the establishment of a connection between the extreme value laws and the statistics of first return and hitting times, see the papers by Freitas *et al.* (2010b) and Freitas *et al.* (2011). In particular, they showed that for dynamical systems preserving an absolutely continuous invariant measure or a singular invariant measure ν, the existence of an exponential hitting time statistics on balls around ν-almost any point ζ implies the existence of extreme value laws for the same observables of type $g_i, i = 1, 2, 3$ mentioned above. The converse is also true, namely, if we have an extreme value law which applies to the observables of type $g_i, i = 1, 2, 3$ achieving a maximum at ζ, then we have exponential hitting time statistics to balls with center ζ. Recently, these results have been generalized to local returns around balls centered at periodic points (Freitas *et al.*, 2010a). These new results are especially promising because the study of the extremes of suitable observables of dynamical systems bypasses the investigation of the cumbersome D_2 and D'

conditions on the underlying dynamical systems, while the possibility to apply the EVT is linked to properties that have a more intuitive appeal and that can be readily checked from the numerical outputs.

C. This work

In a previous work (Faranda *et al.*, 2011a), we have proposed analytic results and algorithms for studying, using the BM approach, EVT on dynamical systems which possess an absolutely continuous invariant measure and satisfy the mixing properties given by conditions D_2 and D', or, equivalently, obey an exponential hitting time statistics. We have established the best *experimental* conditions to observe convergence to the analytical results by highlighting how the deviations from theoretical behavior depend on the number of maxima and on the size of the bin of observations over which each maximum is selected. In another work of (Faranda *et al.* 2011b), reversing the arguments described here, we have shown that extremes can be used as dynamical indicators of whether or not the underlying dynamical system has regular or non-regular motion by checking to what extent the extremes of the g_i $i=1, 2, 3$ observables obey the EVT.

We now present an outline of some aspects of the methodology and main results presented here, in order to help the reader to follow the subsequent discussions, derivations, and analyses of numerical results. In this work, we take on from (Faranda *et al.*, 2011a) and test our methodology for maps possessing a singular invariant measure ν. We remind that in the context of dynamical system, the invariant measure plays the role of the probability measure on the space of events. In this respect, the general theory of extremes will continue to apply no matter whether such a probability is absolutely continuous or singular with respect to Lebesgue.

Nonetheless, an interesting additional point emerges when the invariant measure is singular. In order to get the values of ξ and of a_m and b_m for finite m one should know how the measure of the ball $B_r(\zeta)$ behaves as a function of r and of ζ. Choosing the usual observables g_i, the cumulative distribution function F of the extremes will be related to the scaling of the measure $\nu(B_r(\zeta))$ of a ball $B_r(\zeta)$ of radius r and centered at almost ν-all points ζ. We note that for absolutely continuous measure such a scaling is often available, and we refer to our previous paper (Faranda *et al.*, 2011a). Instead, when singular measures are considered, we are not aware of any analytic result allowing to derive rigorously the expansion of $\nu(B_r(\zeta))$ for small values of r. This will prevent us from computing rigorously the normalising sequences a_m for type 1 observables g_1. Instead, we will be able to get the limiting values of b_m for type 1 and the limiting values of a_m for types 2 and 3. Moreover, it is not possible to compute rigorously the exponent ξ. Our heuristic estimates for the constants and the parameters are obtained by simply approximating $\nu(B_r(\zeta))$ as $\sim r^{D(\zeta)}$, where $D(\zeta)$ is the local dimension of the measure ν at the point ζ. Whenever this limit holds for ν-almost any choice of the point ζ, the corresponding limit is proved to be equal to the Hausdorff dimension of the measure $\nu, HD(\nu)$, defined as the infimum of the Hausdorff dimension of the measurable sets of full ν measure, see Young (1982). This limit is also called the information dimension (which we refer to as Δ in Eckmann and Ruelle (1985)). It is interesting to note that the existence of the above-mentioned limit ν-almost everywhere could be proved for a large class of dynamical systems, especially hyperbolic, and it could be expressed in terms of suitable ratios of the entropies and of the Lyapunov exponents of the measure ν, see Kaplan and Yorke (1979) and Young (1982), for the two-dimensional case and Ledrappier and Young (1985a; 1985b), for the multidimensional case. A good survey of this matter, which also contains the references to the 1-D case investigated by several people, is in the already quoted Eckmann and Ruelle (1985). The agreement with the numerical simulations suggests that the scaling argument is indeed a good choice and implies the possibility of providing a direct proof of the EVT for our observables. Therefore, from the fitted parameters of the GEV distributions, it is possible to deduce $D(\zeta)$, and so, for all practical purposes, also Δ. Therefore, the extremes can be thought of as geometric indicators of the local properties of the attractor.

In order to understand whether the dynamical system we are considering admits an EVT theory, as explained before, we can either check whether that either the conditions D_2 and D' apply or the existence of an exponential return time statistics.

The latter property is easy to prove in the case of the IFSs considered in Sec. III B. These systems, in fact, correspond to expanding maps (since they verify the so-called *open set condition*). In this case, the validity of the exponential return time statistics for balls can be obtained using the technique proposed in Bessis *et al.* (1987). This is the case also for the attractor of the Hénon system when the parameters studied by Benedicks and Carlesson are chosen: Chazottes and Collet have recently established the existence of a Poissonian statistics for the number of visits in balls around generic points with respect to the Sinai-Ruelle-Bowen (SRB) measure (Ruelle, 1989). Our numerical computation will instead concern the *usual* Hénon attractor, but we conjecture that the same behavior is found also in this case. Finally, we consider the Lozi attractor, for which Gupta *et al.* (2011) proved the existence of the extreme value distributions for the observables constructed with the functions g_i and for balls around almost any point with respect to the SRB measure. As a final remark, we stress that the results by Freitas-Freitas and Todd have been proved under the assumption that $\nu(B_r(\zeta))$ is a continuous function of r. This is indeed valid for all the examples considered here.

This paper is organized as follows: in Sec. II, we present the analytical results for the EVT in maps with singular measures and derive the asymptotic behavior of normalising sequences, as well as their link with GEV parameter fitted from finite datasets. In Sec. III, we present the numerical experiments that we have carried out for both singular measures generated with IFSs and for invariant singular measures corresponding to the attractors of the Baker transformation and of the Hènon and Lozi maps, and describe the numerical

procedures used for the statistical inference of the GEV parameters. Eventually, in Sec. IV, we present our conclusion and perspectives for future work.

II. EXTREME VALUE THEORY FOR MAPS WITH SINGULAR MEASURES

A. Definitions and remarks

Let us consider a dynamical system $(\Omega, \mathcal{B}, \nu, f)$, where Ω is the invariant set in some manifold, usually $\mathbb{R}^d$, $\mathcal{B}$ is the Borel σ-algebra, $f : \Omega \to \Omega$ is a measurable map, and ν a probability f-invariant Borel measure. As anticipated in Sec. I, we consider the stationary stochastic process $X_0, X_1, \ldots$ given by

$$X_m(x) = g(\mathrm{dist}(f^m(x), \zeta)) \quad \forall m \in \mathbb{N}, \tag{8}$$

where "dist" is a distance on the ambient space, $\zeta \in \Omega$ is a given point and g is an observable function. We define the partial maximum M_m as

$$M_m = \max\{X_0, \ldots, X_{m-1}\}. \tag{9}$$

The invariant measure ν of the dynamical system plays the role of the probability measure. We also suppose that our system verifies the conditions D_2 and D', or, that obeys and exponential hitting time statistics, as briefly described in the introduction. Therefore, the statistical properties of M_m converge asymptotically, after suitable rescaling, to what prescribed by EVT. We can define three types of observables g_i, $i = 1, 2, 3$ for which it is straightforward to derive the corresponding Type i extreme value law for the normalized maxima

$$g_1(x) = -\log(\mathrm{dist}(x, \zeta)), \tag{10}$$

$$g_2(x) = \mathrm{dist}(x, \zeta)^{-1/\alpha}, \tag{11}$$

$$g_3(x) = C - \mathrm{dist}(x, \zeta)^{1/\alpha}, \tag{12}$$

where C is a constant and $\alpha > 0 \in \mathbb{R}$.

We highlight the main properties of these observables. The functions (10)–(12) share the following properties: (i) they are defined on the positive semi-axis $[0, \infty]$ with values into $\mathbb{R} \cup \{+\infty\}$; (ii) the global maximum can be found in 0, and its value is possibly equal to $+\infty$; (iii) in a neighborhood V of 0, they are strictly decreasing bijections, whose image we denote as W.

The Gnedenko corollary says that the different kinds of extreme value laws are determined by the distribution of

$$F(u) = \nu(X_0 \le u) \tag{13}$$

and by the right endpoint of F, x_F. Therefore, we need to compute and to control the measure $\nu(B_r(\zeta))$ of a ball of radius r around the point ζ. At this regard we invoke, and assume, the existence of the following limit:

$$\lim_{r \to 0} \frac{\log \nu(B_r(\zeta))}{\log r} = D(\zeta), \quad \text{for } \zeta \text{ for } \nu - \text{a.e.} \tag{14}$$

Moreover, we assume that $\nu(B_r(\zeta))$ is a continuous function of r (see Freitas *et al.* (2011) for a discussion of this condition which shows that all the examples considered in the present paper comply with this requirement). When the limit (14) exists on a metric space equipped with the Borel σ-algebra and a probability measure ν, it gives the Hausdorff dimension of the measure or information dimension Δ, and one can prove that it is also the infimum of the Hausdorff dimension taken over all the set of ν measure 1 (Young, 1982). Note that, in some cases, the information dimension Δ can be explicitly computed. For a very general class of one-dimensional maps with positive metric entropy, Δ is equal to the ratio between the metric entropy and the (positive) Lyapunov exponent of ν (Ledrappier, 1981). For two dimensional hyperbolic diffeomorphisms, Δ is equal to the product of the metric entropy times the difference of the reciprocal of the positive and of the negative Lyapunov exponents (Young, 1982). The information dimension is a lower bound of the Hausdorff dimension of the support of the measure ν and is an upper bound of the correlation dimension (Pesin, 1998; Hentschel and Procaccia,1983; Grassberger, 1983; Bessis *et al.*, 1988, Bessis *et al.*, 1987; and Cutler and Dawson, 1989).

B. Limiting behavior of the extreme value theory parameters

We summarize the three basic assumptions for the next considerations, and refer to an orbit with initial condition in the attractor and a point ζ also belonging to the attractor.

- Assumption 1a: Our orbit obeys to conditions D_2 and D'; or
- Assumption 1b: We have an exponential HTS in the neighborhood of ζ;
- Assumption 2: The measure of the ball $B_r(\zeta)$ is a continuous function of r for ν-almost all points ζ and has no atoms;
- Assumption 3: The limit (14) exists (and its value is called $D(\zeta) = \Delta$) at ν-almost all points ζ.

It is now possible to compute rigorously a few of the expected parameters for the three types of observables.

1. Case 1: $g_1(x) = -\log(\mathrm{dist}(x, \zeta))$

Substituting Eq. (8) into Eq. (13), we obtain that

$$\begin{aligned} 1 - F(u) &= 1 - \nu(g(\mathrm{dist}(x, \zeta)) \le u) \\ &= 1 - \nu(-\log(\mathrm{dist}(x, \zeta)) \le u) \\ &= \nu(\mathrm{dist}(x, \zeta) < e^{-u}) = \nu(B_{e^{-u}}(\zeta)) \end{aligned} \tag{15}$$

$$x_F = \sup\{u; F(u) < 1\}.$$

In order to use Gnedenko corollary, we need to compute x_F. In this case, $x_F = +\infty$ as we will explain in the proof below. According to Corollary 1.6.3 in Leadbetter *et al.* (1983) for type 1 $a_m = [G(\gamma_m)]^{-1}$ and $b_m = \gamma_m = F^{-1}(1 - \frac{1}{m})$. We now

show how to get the limiting value of γ_m; a similar proof will hold for types 2 and 3.

Proposition 1. Let us suppose that our system verifies assumptions 1–3 above and let us consider the observable g_1; then

$$\lim_{m\to\infty} \frac{\log m}{\gamma_m} = \Delta$$

Proof. Given the choice of the observable, we have: $1 - F(\gamma_m) = \nu(B_{e^{-\gamma_m}}(\zeta)) = \frac{1}{m}$. Since the measure is not atomic and it varies continuously with the radius, we have necessarily that $\gamma_m \to \infty$ when $m \to \infty$. We define $\delta > 0$ and δ small enough so that there is $m_{\delta,\zeta}$ depending on δ and on ζ, such that for any $m \geq m_{\delta,\zeta}$ we have

$$-\delta\gamma_m \leq \log \nu(B_{e^{-\gamma_m}}(\zeta)) + \Delta\gamma_m \leq \delta\gamma_m. \tag{16}$$

Thanks to the fact that $\log m - \Delta\gamma_m = -[\log \nu(B_{e^{-\gamma_m}}(\zeta)) + \Delta\gamma_m]$ and thanks to the bounds (16), we derive

$$-\delta\gamma_m \leq \log m - \Delta\gamma_m \leq \delta\gamma_m,$$

which proves the Proposition.

The previous proposition does not allow us to derive the values of γ_m and of b_m, which is equal to γ_m for type 1 observables. Instead, we obtain rigorously a limiting behavior, so that we can derive

$$\gamma_m = b_m \sim \frac{1}{\Delta}\log m.$$

The exact results for finite values of m could be obtained only by knowing the functional dependence of $\nu(B_r(\zeta)$ on the radius r and the center ζ. We have been able to perform this derivation in the case of non-trivial absolutely continuous invariant measure in our previous paper (Faranda *et al.*, 2011).

Moreover, we are not able to derive a rigorous limiting behavior for $a_m = [G(\gamma_m)]^{-1}$. The only rigorous results we can derive is that $G(\gamma_m) = o(\gamma_m)$. This can be proved by adapting the previous proof of the Proposition 1 to another result (see Leadbetter *et al.* (1983)), which says that for type 1 observables one has $\lim_{m\to\infty} m(1 - F\{\gamma_m + xG(\gamma_m)\}) = e^{-x}$ for all real x. By choosing $x = 1$, we derive the previous domination result. In the following and again for numerical purposes, we will take

$$a_m = [G(\gamma_m)]^{-1} \sim \frac{1}{\Delta}.$$

This follows easily by replacing in formula (6) $\nu(B_r(\zeta)) \sim r^{\Delta}$ for small values of r.

2. Case 2: $g_2(x) = \text{dist}(x, \zeta)^{-1/\alpha}$

In this case, we have

$$\begin{aligned} 1 - F(u) &= 1 - \nu(\text{dist}(x,\zeta)^{-1/\alpha} \leq u) \\ &= 1 - \nu(\text{dist}(x,\zeta) \geq u^{-\alpha}) \\ &= \nu(B_{u^{-\alpha}}(\zeta)), \end{aligned} \tag{17}$$

and $x_F = +\infty$. Since $b_m = 0$, we need to compute only a_m, which is the reciprocal of $\gamma_m = F^{-1}(1 - 1/m)$. By adapting Proposition 1, we derive that

$$\lim_{m\to\infty} \frac{\log m}{\log \gamma_m} = \alpha\Delta.$$

This allows us to use the approximation $a_m \sim \frac{1}{m^{\frac{1}{\alpha\Delta}}}$. The exponent ξ for type 2 observables is given by the following limit (see Leadbetter *et al.*, 1983, Theorem 1.6.2):

$$\lim_{t\to\infty}(1 - F(tx))/(1 - F(t)) = x^{-\xi}, \xi > 0, x > 0.$$

The approximation $\nu(B_r(\zeta)) \sim r^{\Delta}$ gives that $\xi \sim \alpha\Delta$. This result is crucial for estimating the exponent $\xi' = 1/\xi$ appearing in the definition of GEV cdf.

3. Case 3: $g_3(x) = C - \text{dist}(x, \zeta)^{1/\alpha}$

We have

$$\begin{aligned} 1 - F(u) &= 1 - \nu(C - \text{dist}(x,\zeta)^{1/\alpha} \leq u) \\ &= \nu(B_{(C-u)^{\alpha}}(\zeta)). \end{aligned} \tag{18}$$

In this case $x_F = C < \infty$ and $a_m = (C - \gamma_m)^{-1}$; $b_m = C$. The previous proposition leads us to derive that $\lim_{n\to\infty} \frac{\log m}{-\alpha \log(C-\gamma_m)} = \Delta$, which gives the asymptotic scalings $\gamma_m \sim C - \frac{1}{m^{\frac{1}{\alpha D}}}$; $a_m \sim m^{\frac{1}{\alpha\Delta}}$; $b_m = C$. Finally, the exponent ξ is given again following Theorem 1.6.2 in Leadbetter *et al.* (1983) by the formula:

$$\lim_{h\to 0}(1 - F(C - hx))/(1 - F(C - h)) = x^{\xi}, \xi < 0, x < 0,$$

which, considering our usual approximation, results into $\xi \sim -\alpha\Delta$.

C. Generalised observables of type i, i = 1, 2, 3

The prototypical observables g_i, $i = 1, 2, 3$ previously introduced allow for computing explicitly the EVT of mixing dynamical systems. Nonetheless, those results can be generalised by considering a broader class of observables, the basic reason being that we are only interested in asymptotic properties. Whereas in the rest of the paper, we will consider explicitly only the observables given in Eqs. (10)–(12), it is worth mentioning the properties required to extend the theory beyond such a rather limited set of observables. Following Freitas *et al.* (2010b), we state below which is the property required to a g_i-observables so that the statistical properties of its maxima are described by an EVT type i distribution:

- Type 1: There is a strictly positive function $p : W \to \mathbb{R}$ such that $\forall y \in \mathbb{R}$ we have

$$\lim_{s\to g_1(0)} \frac{g_1^{-1}(s + yp(s))}{g_1^{-1}(s)} = e^{-y}.$$

- Type 2: $g_2(0) = +\infty$ and there exists $\beta > 0$ such that $\forall y > 0$ we have

$$\lim_{s\to\infty} \frac{g_2^{-1}(sy)}{g_2^{-1}(s)} = y^{-\beta}.$$

- Type 3: $g_3(0) = D < +\infty$ and there exists $\gamma > 0$ such that $\forall y > 0$ we have

$$\lim_{s\to 0} \frac{g_3^{-1}(D - sy)}{g_3^{-1}(D - s)} = y^{\gamma}.$$

III. NUMERICAL EXPERIMENTS

A. Procedure for statistical inference

In this paper, we consider numerical experiments performed on several maps featuring a singular invariant measure. The methodology followed here mirrors what described in Faranda *et al.* (2011a). For each map, we run a long simulation up to k iterations starting from a given initial condition on the attractor, as experimentally defined by a preliminary long integration. Then, at each iteration step, we compute the observables g_1, g_2, g_3 defined in Eqs. (10)–(12) and divide the k-long time series into n bins each containing $m = k/n$ observations. We then select for each bin the maximum, thus ending up with a series of n maxima, which provides the basis for our statistical inference. At this stage, we fit the n-maxima to the statistical model provided by the GEV family of distributions and compare the best fit to the theoretical results presented in the previous sections.

In Faranda *et al.* (2011a), the fitting procedure was based on the classical method of maximum likelihood estimation (MLE). Unfortunately, this method is not efficient when dealing with singular measures. Therefore, in this paper, we resort to an L-moments estimation procedure, which, being based upon the computation of integrals rather than upon the solution of a variational problem, is more robust. The L-moments are analogous to ordinary moments, but are computed from linear combinations of the data values, arranged in increasing order. Asymptotic approximations to sampling distributions are better for L-moments than for ordinary moments (Hosking 1990). We have derived from the L-moments the best estimates of the GEV parameters as described in Hosking (1990), while we have computed the 95% confidence intervals of the parameters using a bootstrap procedure. In order to double-check the procedure adopted here, we have recomputed the results presented in Faranda *et al.* (2011a) with the L-moments method (not shown), finding that the MLE and L-moments method have comparable precision.

In principle, using the L-moments procedure, we could fit the data to any kind of known cdf. In order to validate the use of the GEV model, we have proceeded as follows:

- *A priori*: The choice of a GEV model arises naturally if the assumptions presented in Sec. II are satisfied.
- *A posteriori*: We can verify the goodness of fit to the GEV family by applying some parametric or non-parametric tests commonly used in statistical inference procedures. Using the Kolmogorov Smirnov test (Lilliefors 1967), we have quantified the deviation between the empirical cdf and the fitted GEV cdf, obtaining in all cases a positive outcome.

We summarize below what are the implications of the conjecture proposed above in terms of fits of GEV parameters obtained from numerical experiments. Since we consider a fixed length of the time series $k = n \cdot m$, the following relationships can be obtained by simply replacing $m = k/n$ in the equations derived in Sec. II. By replacing m with n we are just rewriting the relationship in terms of k which is the constant total length of the series. This does not affect in any way the results that have been checked both against n and against m (not shown in the figures), see the discussion in Faranda *et al.* (2011). For the g_1-type observable, we have

$$\sigma = \frac{1}{\Delta}, \quad \mu \sim \frac{1}{\Delta}\ln(k/n), \quad \xi' = 0. \tag{19}$$

For the g_2-type observable

$$\sigma \sim n^{-1/(\alpha\Delta)}, \quad \mu \sim n^{-1/(\alpha\Delta)}, \quad \xi' = \frac{1}{\alpha\Delta}. \tag{20}$$

Note that, as discussed in Beirlant *et al.* (2004), two suitable choices for the suitable normalising sequence b_m are, in principle, possible, given by $b_m = 0$ or $b_m = c \cdot m^{-\xi'}$ where $c \in \mathbb{R}$ is a positive constant. We obtain that in all cases, the experimental procedure we use automatically selects $b_m \sim c \cdot m^{-\xi'}$, hence the m-dependence of μ given above. Finally, for the g_3-type observable, we have

$$\sigma \sim n^{1/(\alpha\Delta)}, \quad \mu = C, \quad \xi' = -\frac{1}{\alpha\Delta}. \tag{21}$$

B. Iterated function systems and Cantor sets

A Cantor set can be obtained as an attractor of a suitable defined IFS. An IFS is a finite family of contractive maps $\{f_1, f_2, ..., f_s\}$ acting on a compact normed space Ω with norm $|\cdot|$ and possessing a unique compact limit set (the attractor) $K \in \Omega$ which is non-empty and invariant by the IFS, namely

$$K = \bigcup_{i=1}^{s} f_i(K).$$

We will put a few restrictions on the IFS in order to see it as the *inverse* of a genuine dynamical system; we will shortly explain why this change of perspective will help us to compute observables on fractal sets. We refer to the seminal paper by Barnsley and Demko (1985) for a comprehensive outlook on the concepts and results we are going to use here. First, we assume that all the f_i's are strict contractions, i.e., there is a number $0 < \lambda < 1$ such that $\forall i = 1, \ldots, s$ we have that $|f_i(x) - f_i(y)| < \lambda|x - y| \forall x, y \in \Omega$. Furthermore, we suppose that each f_i is one-to-one on the attractor K and that $\forall i = 1, \ldots, s$ we have $f_i(K) \cap f_j(K) = \emptyset, i \neq j$. This is the crucial *open set condition*. Therefore, we can define one measurable map $T : K \to K$ by $T(x) = f_i^{-1}(x)$ for $x \in f_i(K)$: the attractor K will be the invariant set for T. Therefore,

the transformation T plays the role of a usual dynamical system.

In order to attain a complete statistical description of a dynamical system we need to construct an invariant probability measure. In particular, we require such a measure be ergodic, if we want to meaningfully compute the maxima of the time series constructed with the observables g_i. The time series of the observables come from a forward orbit of an initial point chosen according to the invariant measure.

The construction of the invariant ergodic measure goes as follows. First, we associate to each map f_i a positive weight p_i in such a way that $\sum_{i=1}^{s} p_i = 1$. Then, it is possible to prove the existence of a unique measure ν (called *balanced*) which enjoys the following properties:

- the measure ν is supported on the attractor K and is invariant for the map T associated the considered IFS;
- for any measurable set B in Ω we have

$$\nu(B) = \sum_{i=1}^{s} p_i \nu(f_i^{-1}(B)).$$

- Defining $(Sg)(x) = \sum_{i=1}^{s} p_i g(f_i(x))$ for a continuous functions g on Ω and for *any* point $x \in \Omega$, then we have

$$\lim_{n\to\infty} S^n g(x) = \int_{\Omega} g d\nu.$$

The last property is especially important: it holds also for the characteristic function of a set, provided that the boundary of this set has zero ν-measure. Moreover, it corresponds to a sort of ergodic theorem, as it states that the backward orbit constructed by applying to any point in Ω the maps f_i with weights p_i will populate the attractor K as the forward orbit according to the transformation T associated with the IFS of a point chosen almost everywhere according to ν in K. Some examples relevant for our discussion are presented below.

1. Uniform weights

We consider the Cantor ternary set, which can be constructed as the attractor K_1 of the IFS $\{f_1, f_2\}$ defined as

$$\begin{cases} f_1(x) = x/3 \text{ with weight } p_1 \\ f_2(x) = (x+2)/3 \text{ with weight } p_2 \end{cases}, \tag{22}$$

where $x \in [0,1]$, and we set $p_1 = p_2 = 1/2$. Therefore, at each time step, we have the same probability to iterate $f_1(x)$ or $f_2(x)$. Equivalently, the previous IFS can be written as

$$x_{t+1} = (x_t + b)/3, \tag{23}$$

where, at each time step, we select randomly with equal probability b to be either 0 or 2. Instead, the so-called Sierpisnkij triangle can be constructed as the attractor K_2 of the following IFS:

$$\begin{cases} x_{t+1} = (x_t + v_{p,1})/2 \\ y_{t+1} = (y_t + v_{p,2})/2 \end{cases}, \tag{24}$$

where, at each iteration, we assign to the number p with equal probability 1/3 one of the values 1,2, or 3, and, subsequently, iterate the map 24 substituting the elements $v_{p,1}$ and $v_{p,2}$ of the following matrix:

$$v = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

As well known, the information dimension is $\Delta = \log(2)/\log(3)$ for the Cantor ternary set and $\Delta = \log(3)/\log(2)$ for the Sierpinskij triangle (Sprott, 2003).

In order to obtain a suitable center ζ belonging to the fractal attractor, we proceed as follows. We take a point $x \in \Omega$ and we consider ζ as one of the preimages $f^{-t}(x)$, with t much larger than the sequence of observed events. By construction, $f^{-t}(x)$ is closer and closer to the invariant Cantor set and approaches a generic point with respect to the balanced measure ν.

We start by considering the empirical cdf $F(u)$ of the extrema for a g_3-observable. An example is shown in Fig. 1. The histogram is obtained iterating the map in Eq. (22) for $k = 5 \times 10^7$ iterations, $\zeta \simeq 0.775, \alpha = 4$, $C = 10$. In this case, we have divided the time series of k data into bins, each containing $m = 5000$ data, so that we obtain a total of $n = 10\,000$ maxima. As claimed in Sec. II, the empirical cdf is a singular continuous function, and this is due to the structure of the Cantor set, and the fitted GEV cdf is an excellent continuous approximation which averages out the holes of the Cantor set. The results are qualitatively similar for the other observables and other initial conditions on the attractor.

We need to provide a more robust quantitative evaluation of the suitability of the approach proposed here. In order to check that, effectively, the parameters of the GEV distribution obtained by L-moments estimation are related to the information dimension of the Cantor ternary set and of the Sierpinskij triangle, we have considered an ensemble of 10^4 different realizations of Eqs. (22) and (24), each one of

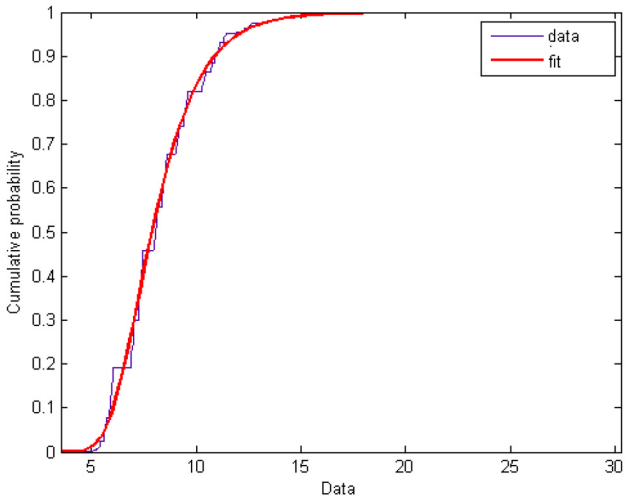


FIG. 1. Empirical (blue) and fitted GEV (red) cdf for the IFS that generates a Cantor ternary set, $\zeta \simeq 0.775$, g_3 observable function.

length $k = 10^7$, and each one starting from the same initial condition. In order to check the effectiveness of the inference algorithm, we have considered different values of n, number of extracted maxima, and m, length of the bins, keeping in mind that $k = n \cdot m$. As discussed in Faranda *et al.* (2011), we need to ensure that, at the same time, we are dealing with *true* extremes (m must be large enough) and that we have a sample of extremes sufficiently large to perform effectively the statistical inference (n must be large enough). Rather general considerations suggest a good convergence is observed when $n, m > 1000$. We take into account these (soft) bounds and consider only (n, m) pairs that satisfy them.

In Figs. 2–4, we present the results of the fits of the GEV parameters for the extremes of the observables g_1, g_2, and g_3, respectively. In each figure, from top to bottom, we present the estimates for ξ', σ, and μ as a function of n, respectively. On the left (right) hand side, the plots refer to the outputs of the IFS that generate the Cantor ternary set (Sierpisnkij triangle). In all the cases considered, the numerical results agree to a high degree of accuracy with the analytical estimates given in Eqs. (19)–(21). The results shown here refer to the initial conditions (and centers of the balls) $\zeta = 1/3$ for Cantor ternary set and $\zeta \simeq (0.02, 0.40)$ for Sierpinskij triangle, but we have verified that, as expected, the choice of the initial conditions is immaterial as long as it is chosen on the attractor. The black lines correspond to the mean value of the estimate of the parameter over the different realizations of the map, while the black dotted lines represent one standard deviation over such ensemble.

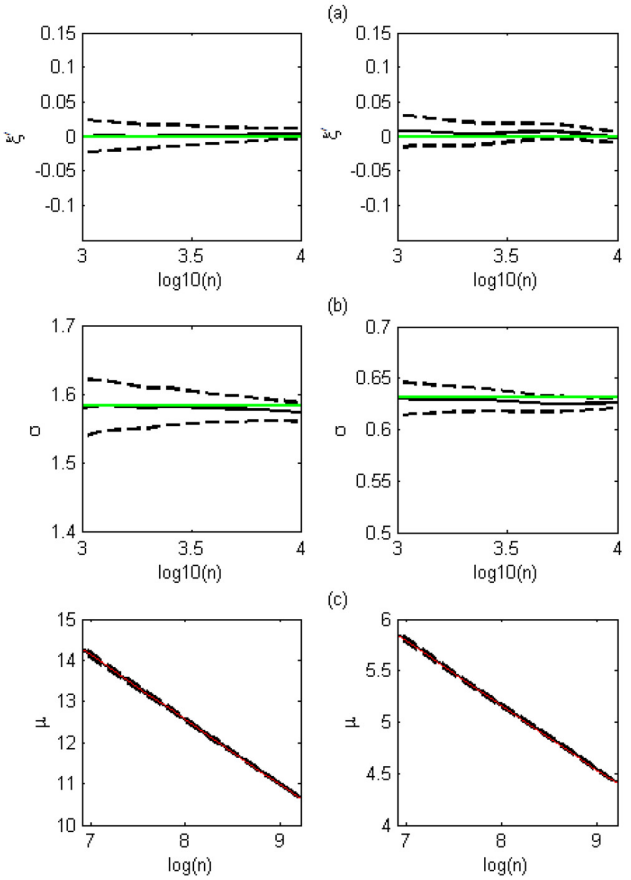


FIG. 2. g_1 observable. (a) ξ' vs $\log_{10}(n)$; (b) σ vs $\log_{10}(n)$; (c) μ vs $\log(n)$. Left: Cantor ternary set. Right: Sierpinskij triangle. The dashed lines represent one standard deviation from the mean values (black solid lines), the red lines represent the logarithmic linear fit, the green lines are the theoretical values.

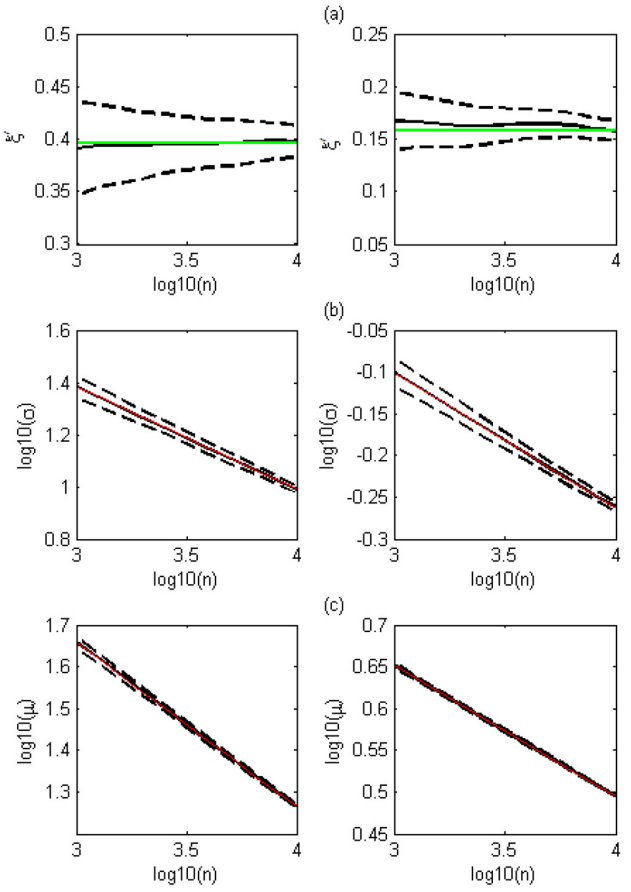


FIG. 3. g_2 observable (a) ξ' vs $\log_{10}(n)$; (b) $\log_{10}(\sigma)$ vs $\log_{10}(n)$; (c) $\log_{10}(\mu)$ vs $\log_{10}(n)$. Left: Cantor ternary set. Right: Sierpinskij triangle. The dashed lines represent one standard deviation from the mean values (black solid lines), the red lines represent he power law fits, green lines are the theoretical values.

When considering a g_1 observable function, we expect to find $\xi' = 0$, see Eq. (19). The numerical outputs are in excellent agreement with this prediction for both IFS considered, as shown in Figure 2(a). For the scale parameter σ, a similar agreement is achieved with respect to the theoretically derived values. In fact, we have that $\sigma = 1/\Delta$. This implies that $\sigma = \log(3)/\log(2) \simeq 1.5850$ for the Cantor tertiary set and that $\sigma = \log(2)/\log(3) \simeq 0.6309$ for the Sierpinskij triangle. These values are represented in Fig. 2(b) with a green line. Eventually, the experimental values for the location parameter μ agree convincingly with the logarithm decay with n given in Eq. (19). A linear fit of μ in respect to $\log(n)$ is shown with a red line in Figure 2(c) for both IFSs.

The agreement between the theoretical results and the numerical evidences is confirmed also for g_2-type and g_3-type observables, as shown in Figs. 3 and 4, respectively. Results are presented for the case $\alpha = 4$ for both for Sierpinskij triangle and Cantor ternary set IFSs and,

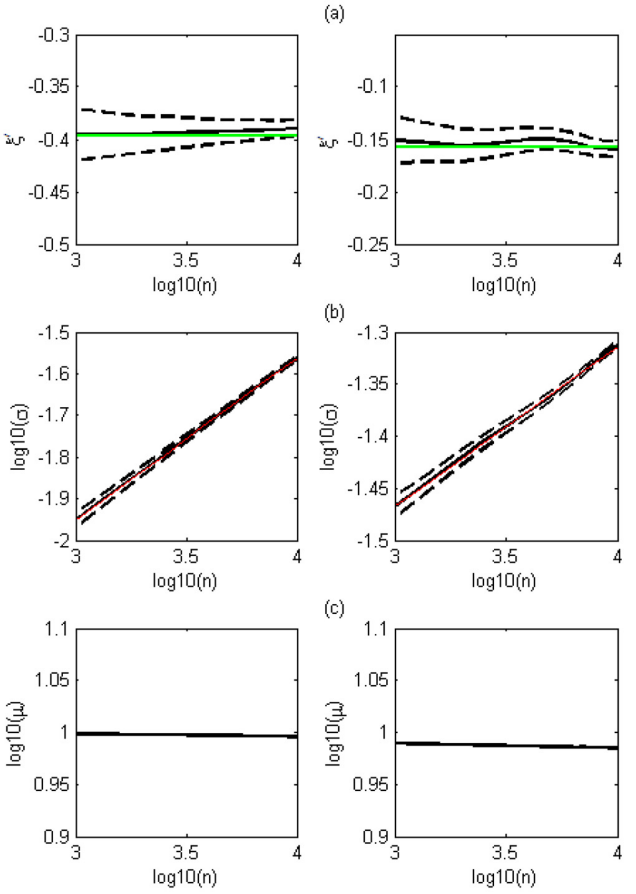


FIG. 4. g_3 observable. (a) ξ' vs $\log_{10}(n)$; (b) $\log_{10}(\sigma)$ vs $\log_{10}(n)$; (c) $\log_{10}(\mu)$ vs $\log_{10}(n)$. Left: Cantor ternary set. Right: Sierpinskij triangle. The dashed lines represent one standard deviation from the mean values (black solid lines), the red lines represent the power law fits, the green lines are the theoretical values.

additionally, in the case of g_3-type observables, we choose $C = 10$.

For both g_2-type and g_3-type observables, we find an excellent agreement between the experimental and theoretical values for the ξ'-values given in Eqs. (20) and (21), respectively, as shown by the fact that the green lines in Figs. 3(a) and 4(a) lie well within the confidence interval of the inferred parameters' values. Similarly, the inferred values for the parameter μ in the case of the g_3-observable are weakly dependent on n, so that we find—see Fig. 4(c)—an excellent agreement with the theoretical prediction given in Eq. (21) of finding for all values of n the constant value $\mu = C = 10$.

In Figs. 3(b), 3(c), and 4(b), a log-log scale is used to highlight the agreement between the inferred values and the power-law behavior described by Eqs. (20) and (21). The uncertainty is in all cases given as one standard deviation of parameters' estimates over the ensemble of realizations. The best power-law fits of the experimental data are shown by the red lines, which are almost exactly overlapping with the green lines, corresponding to the theoretical values. Equations (20) and (21) imply that the angular coefficients κ of the green lines are such that for all Figs. 3(b), 3(c), and 4(b) we have that $|\kappa| = 1/(\alpha\Delta)$.

Interestingly, by using Eqs. (19)–(21), we can estimate in multiple ways the information dimension Δ of the attractors from the inferred values of the GEV distributions parameters, plus, in the case of g_2 and g_3 observables, from the choice of α. In Table I we report our results: we find an overall excellent agreement of the estimates among themselves and with the theoretical values. This suggests the rather relevant fact that analysing the extremes of the observables g_i for a mixing dynamical system provides a way to study the geometrical properties of its attractor, the fundamental reason being that the looking at the maxima of the observables g_i amounts to using a magnifying glass in the neighborhood of the center of the ball ζ, because we select the close recurrences of the orbit nearby ζ.

As a side note, we wish to underline that other tests have been done computing the statistics using parameter $\alpha = 5, 6, 7, 8$ for g_2 and g_3 observables for both the Cantor tertiary set and the Sierpinskij triangle IFSs, with no notable deviation from what given in Eqs. (19)–(21). Instead, serious problems in terms of numerical convergence arise for the Cantor tertiary set IFS when using $\alpha = 2$ and $\alpha = 3$ for the observables g_2 and g_3. This issue is possibly due to the fact that L-moments method works efficiently only if $\xi' \in [-0.5, 0.5]$, while for $\alpha \leq 3$, the theoretically expected value for the shape parameter $|\xi'| > 0.5$.

2. IFS with non-uniform weights

Let us now generalize the problem discussed in Sec. III B 1 by considering the case of the IFS

$$f_k(x) = a_k + \lambda_k x \quad x \in [0,1] \quad k = 1, 2, ..., s, \tag{25}$$

where each f_i is iterated with (different) probability w_i. In this case, it is possible to compute the information dimension as the ratio between the metric entropy and the Lyapunov exponent of the associated balanced measure (Barnsley 2000)

$$\Delta = \frac{w_1 \log w1 + \ldots + w_s \log w_s}{w_1 \log \lambda_1 + \ldots + w_s \log \lambda_s}. \tag{26}$$

We hereby generalize the Cantor tertiary set IFS by considering the following IFS:

$$\begin{cases} f_1(x) = x/3 & \text{with weight } w \\ f_2(x) = (x+2)/3 & \text{with weight } 1-w \end{cases}, \tag{27}$$

where the weight w is varied between 0.35 and 0.65 with steps of 0.01. For $w = 0.5$, we obtain the same results shown in Sec. II, while for different weights we can check Eq. (26). Note that, in this general case, the value of Δ can be more efficiently obtained by taking the average of $D(\zeta)$ over many different ζ belonging to the attractor of the system (27) distributed, asymptotically, according to the physical measure.

Equations (19)–(21) and Table I suggest various different and, in principle, equivalent ways to estimate Δ also in the more complex context of a multifractal set. In Fig. 5, we present the estimates for the dimension Δ obtained from Eq. (19) as

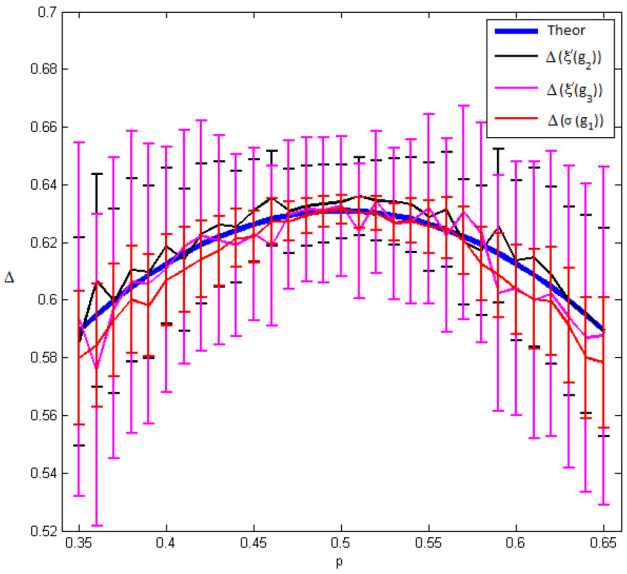


FIG. 5. Estimates of Δ for the modified Cantor tertiary set IFS given in Eq. (27) as a function of the weight parameter w. The inferred results are obtained using the estimates of the GEV parameters for g_i observables. See text for further details.

$$\Delta(\sigma(g_1)) = \frac{1}{\langle \sigma(g_1) \rangle}, \tag{28}$$

as well as from Eqs. (20) and (21) as follows:

$$\Delta(\xi'(g_i)) = \frac{1}{\alpha |\langle \xi'(g_i) \rangle|}, \quad i = 2, 3 \tag{29}$$

where in all expressions above the brackets $\langle \bullet \rangle$ indicate an average on different sample points ζ taken in the attractor of the IFS (asymptotically) according to the balanced measure. For the rest of the numerical computations, we set $\alpha = 5$. The parameters have been estimated using 1000 different initial conditions on the support of the attractor, and for 30 realizations of each sample point ζ. Finally, the block-maxima approach is implemented by choosing with $n = m = 1000$, which, as described above, is within the range of robustness of the inference procedure, and the error bars, which represent the standard deviation of the mean, are computed using the standard error propagation rules.

The agreement between the theoretical information dimension Δ (blue line) and the experimental data is clear for all the values of the weight w and, quite consistently, for all the inference procedures given in Eqs. (28) and (29). The uncertainty increases the more w departs from 0.5, the reason being that the larger the difference between w and 0.5, the harder it is to compute precisely $D(\zeta)$, as the multifractal nature of the attractor is more pronounced. For all values of w, the best agreement between the inferred and theoretical value of Δ is obtained by deriving Δ from the estimates of the σ parameter for the (g_1)-type observable. It is not clear why for finite samples the convergence is better for this specific estimator of Δ, possibly because the logarithm modulation of the distance highlights proper extrema while weighting less possible outliers. Nevertheless, we remark again that all the estimators are in agreement with the theoretical estimate.

C. Non-trivial singular measures

In Sec. III B, we have analysed the relatively simple cases of Cantor sets generated with IFSs. In order to provide further support to our conjectures, we now present some application of our theory to the output of dynamical systems possessing less trivial singular measures. We analyze three relevant examples of two-dimensional maps: the Baker map, the Hènon map, and the Lozi map.

1. The Baker map

The Baker map is defined as follows:

$$x_{t+1} = \begin{cases} \gamma_a x_t \mod 1 & \text{if } y_t < \alpha, \\ 1/2 + \gamma_b x_t \mod 1 & \text{if } y_t \geq \alpha, \end{cases}$$
$$y_{t+1} = \begin{cases} \dfrac{y_t}{\alpha} \mod 1 & \text{if } y_t < \alpha, \\ \dfrac{y_t - \alpha}{1 - \alpha} \mod 1 & \text{if } y_t \geq \alpha, \end{cases} \tag{30}$$

where we consider the classical values for the parameters: $\alpha = 1/3, \gamma_a = 1/5$ and $\gamma_b = 1/4$. Rigorous analytical results are available for the estimate of the information dimension Δ (Kaplan and Yorke,1979). For the parameters' values considered here, we have that $\Delta \simeq 1.4357$.

We have performed the same analysis detailed in Sec. III B, but taking into account an important difference. This map is invertible and its invariant set is an attractor given by the Cartesian product of a segment, along the y-axis, and a one-dimensional Cantor set, along the x-axis. The system possesses an invariant SRB measure, which can practically be constructed by taking ergodic sums for any point sitting on the basin of attraction. In order to compute the center of the balls on the attractor, we proceed in a similar manner as described for IFS, namely, we take any point x in the basin of attraction and we iterate it t times with t much bigger than the sequence of observed events. Then, we take the point $f^t(x)$ as ζ: as t get larger, $f^t(x)$ is closer and closer to the attractor and distributed according to the SRB measure. In our set-up, we choose $\alpha = 4$ for g_2 and g_3, and $C = 10$ for g_3. The results are shown in Figs. 6–8: the black continuous lines represent the parameter average over different initial conditions and the black dashed lines represent the standard deviation of the distribution of the estimated parameters.

The expected theoretical values for ξ' are within one standard deviation of the results obtained from the fits of the tail parameter for all the three observables g_i's. The agreement seems to be better when we increase n, even if this, in our experimental set-up, corresponds to a decrease of m. Such a behavior is quite interesting as it seems that we obtain a much better convergence to theoretical values if $n \simeq 10^4$, while in all the other examples, there is no such a difference between $n = 10^3$ and $n = 10^4$. A similar consideration can be made for the dependence on n of the parameter σ inferred from a g_1-type observable, shown in Fig. 6(b), while Fig. 8(c) shows that the parameter μ obtained from the g_3-type observable closely approximates $C = 10$ for all the values of n explored here.

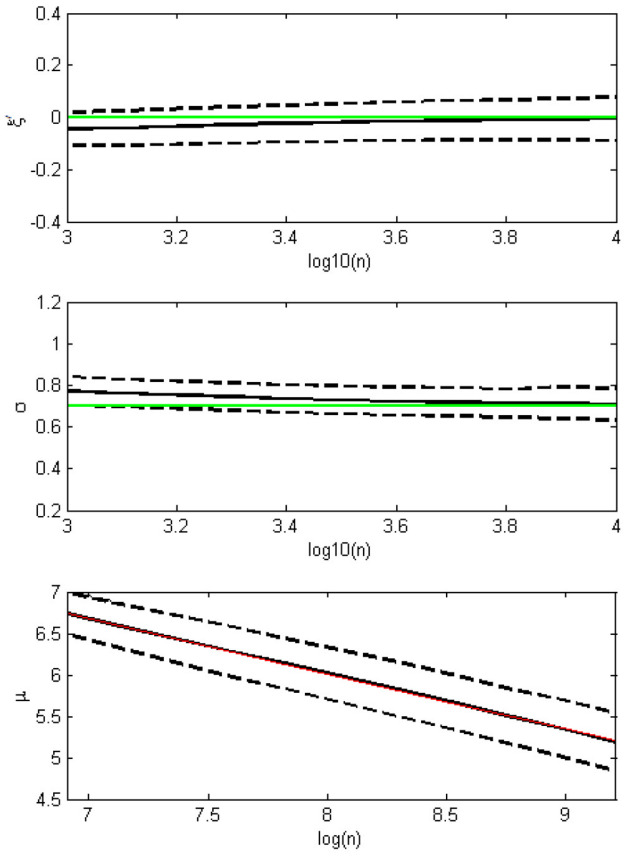


FIG. 6. g_1 observable. (a) ξ' vs $\log_{10}(n)$; (b) σ vs $\log_{10}(n)$; (c) μ vs $\log(n)$. Baker map. The dashed lines represent one standard deviation from the mean values (black solid lines), the red line represents the logarithmic linear fit, the green lines give the theoretical values.

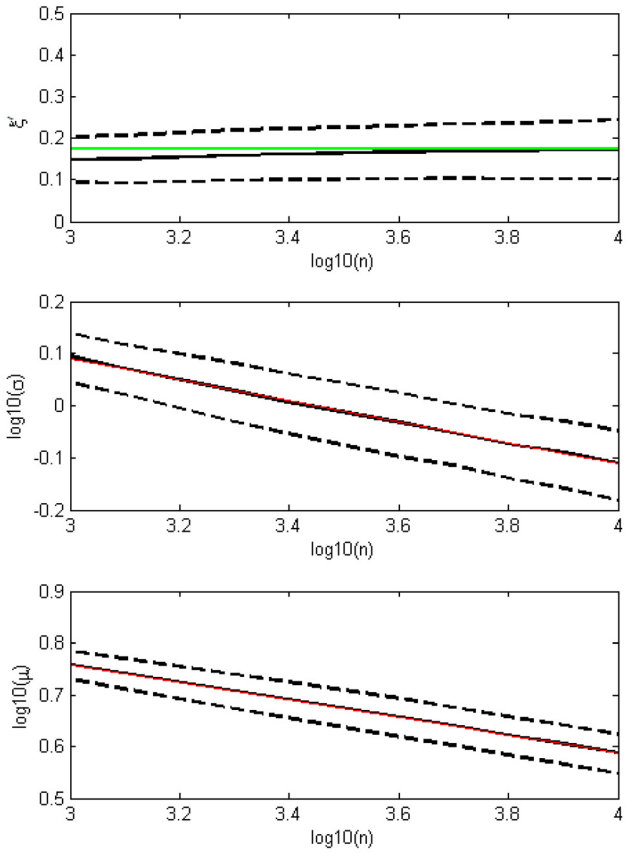


FIG. 7. g_2 observable (a) ξ' vs $\log_{10}(n)$; (b) $\log_{10}(\sigma)$ vs $\log_{10}(n)$; (c) $\log_{10}(\mu)$ vs $\log_{10}(n)$. Baker map. The dashed lines represent one standard deviation from the mean values (black solid lines), the red lines represent the power law fits, the green lines give the theoretical values.

A semilog-linear plot clearly depicts the dependence on n of the estimates of μ of a g_1-type observable in Fig. 6(c), while log-log plots of the GEV inferred parameters against n are shown in Figs. 7(b), 7(c), and 8(b). Along the lines of the discussion following Figs. 2–4, we have that on the average (with respect to the choice of the points ζ) the theoretical results given in Eqs. (19)–(21) and the best fits to those law obtained with the experimental data have a truly outstanding agreement. Obviously, the uncertainties are larger than in the case of the IFS attractors because in the case of the Baker map self-similarity is not exactly obeyed in the attractor, while local properties are relevant. Along the lines of Table I, we can invert the expressions given in Eqs. (19)–(21) and derive various independent estimates of the information dimension Δ from the GEV parameters estimates extremes of the observables g_i, suitably averaged over many centers ζ along the lines of Eqs. (28) and (29). The agreement between the theoretical value of Δ and the experimentally retrieved estimates is fairly good enough for all the options given in Table II, with the value for Δ deduced from the expression of μ in Eqs. (19) and (20) faring best.

2. The Hènon and Lozi maps

The Hènon map is defined as

$$\begin{aligned} x_{t+1} &= y_t + 1 - a x_t^2 \\ y_{t+1} &= b x_t, \end{aligned} \tag{31}$$

while in the Lozi map the term ax_t^2 is substituted with the term $a|x_t|$

$$\begin{aligned} x_{t+1} &= y_t + 1 - a|x_t| \\ y_{t+1} &= b x_t. \end{aligned} \tag{32}$$

We consider the classical set of parameter $a = 1.4$, $b = 0.3$ for the Hènon map and $a = 1.7$ and $b = 0.5$ for the Lozi map. Young (1985) proved the existence of the SRB measure for the Lozi map, whereas for the Hènon map no such rigorous proof exists, even if convincing numerical results suggest its existence (Badii and Politi, 1987). Note that Benedicks and Young (1993) proved the existence of an SRB measure for the Hènon map with a different set of parameters using the approach outlined in Benedicks and Carleson (1991). Using the classical Young results, which makes use of the Lyapunov exponents, we obtain an exact result for Δ for the Lozi attractor

$$\Delta \simeq 1.40419.$$

Instead, in the case of the Hènon attractor, we consider the numerical estimate provided by Grassberger and Procaccia (1983)

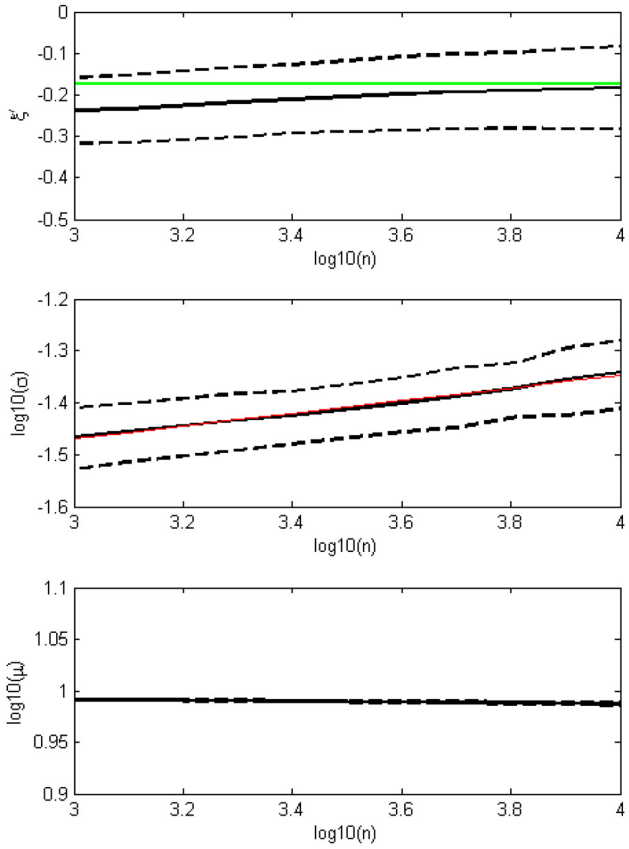


FIG. 8. g_3 observable. (a) ξ' vs $\log_{10}(n)$; (b) $\log_{10}(\sigma)$ vs $\log_{10}(n)$; (c) $\log_{10}(\mu)$ vs $\log_{10}(n)$. Baker map. The dashed lines represent one standard deviation from the mean values (black solid lines), the red lines represent the power law fits, the green lines give the theoretical values.

$$\Delta = 1.25826 \pm 0.00006.$$

As in the previous cases, the GEV distribution is computed with L-moments methods by varying n and m and averaging the distribution parameters over 1000 different sample points ζ chosen with the same strategy as described before for the Baker map. The results are presented in Figs. 9–11, with the plots on the left-hand side referring to the Hènon map, while those on the right-hand side refer to the Lozi map. In general, the performance of these two maps is rather similar in terms

TABLE I. Estimates for the information dimension Δ for the Cantor tertiary set and Sierpinskij triangle obtained through Eqs. (19)–(21) via the inferred values of the GEV distributions. The g_2 and g_3 observables are specified by $\alpha = 4$; for the g_3 observable the value $C = 10$ is also selected.

Parameter $[g_i]$	Cantor	Sierpinskij
Theory	$\log(2)/\log(3) \simeq 0.6309$	$\log(3)/\log(2) \simeq 1.5850$
$\sigma\,[g_1]$	0.635 ± 0.006	1.57 ± 0.03
$\mu\,[g_1]$	0.64 ± 0.01	1.59 ± 0.01
$\xi'\,[g_2]$	0.628 ± 0.007	1.56 ± 0.02
$\sigma\,[g_2]$	0.63 ± 0.01	1.56 ± 0.02
$\mu\,[g_2]$	0.64 ± 0.01	1.59 ± 0.01
$\xi'\,[g_3]$	0.642 ± 0.008	1.60 ± 0.02
$\sigma\,[g_3]$	0.64 ± 0.01	1.62 ± 0.01

TABLE II. Estimates for the information dimension Δ for the Baker, Hènon and Lozi maps obtained through Eqs. (19)–(21) via the inferred values of the GEV distributions. The g_2 and g_3 observables are specified by $\alpha = 4$; for the g_3 observable the value $C = 10$ is also selected.

Parameter $[g_i]$	Baker map	Hènon map	Lozi map
Theory	1.4357	1.2582	1.4042
$\sigma\,[g_1]$	1.43 ± 0.03	1.21 ± 0.02	1.39 ± 0.01
$\mu\,[g_1]$	1.48 ± 0.03	1.23 ± 0.02	1.40 ± 0.01
$\xi'\,[g_2]$	1.41 ± 0.02	1.24 ± 0.02	1.41 ± 0.01
$\sigma\,[g_2]$	1.39 ± 0.04	1.35 ± 0.07	1.38 ± 0.02
$\mu\,[g_2]$	1.47 ± 0.02	1.24 ± 0.01	1.40 ± 0.01
$\xi'\,[g_3]$	1.45 ± 0.02	1.28 ± 0.02	1.43 ± 0.01
$\sigma\,[g_3]$	1.56 ± 0.08	1.15 ± 0.07	1.42 ± 0.01

of agreement between the theoretical values and the experimental estimates.

When considering ξ', the numerical results are in good agreement with the theoretical estimates. Nevertheless, the parameters have a rather large spread, comparable with that observed for the Baker map, which indicates a slower convergence towards the expected values with respect to what is observed for the IFS case. The estimates of ξ', though, are more stable with respect to different choices of n with

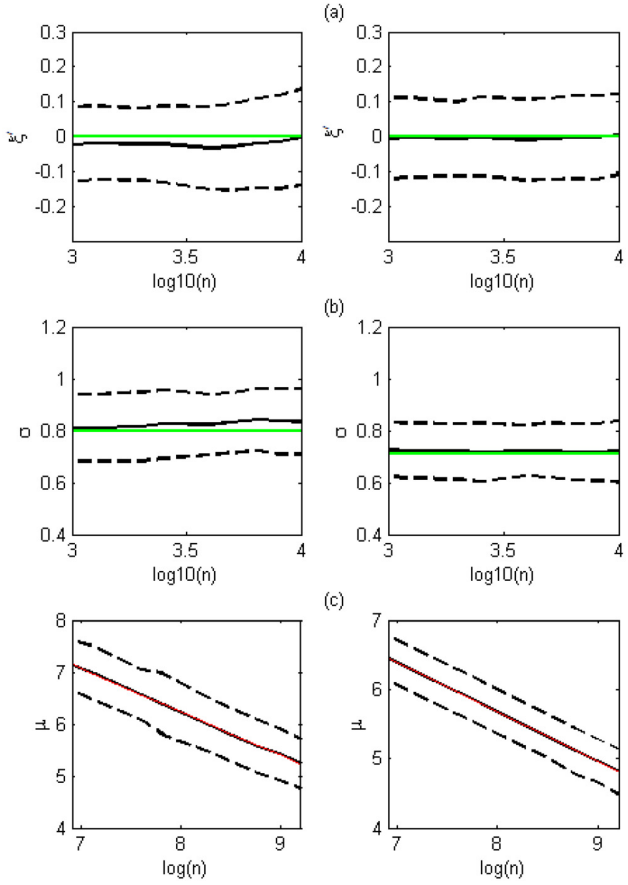


FIG. 9. g_1 observable. (a) ξ' vs $\log_{10}(n)$; (b) σ vs $\log_{10}(n)$; (c) μ vs $\log(n)$. Left: Hènon map, Right: Lozi map. The dashed lines represent one standard deviation from the mean values (black solid lines), the red lines represent the logarithmic linear fits, the green lines give the theoretical values.

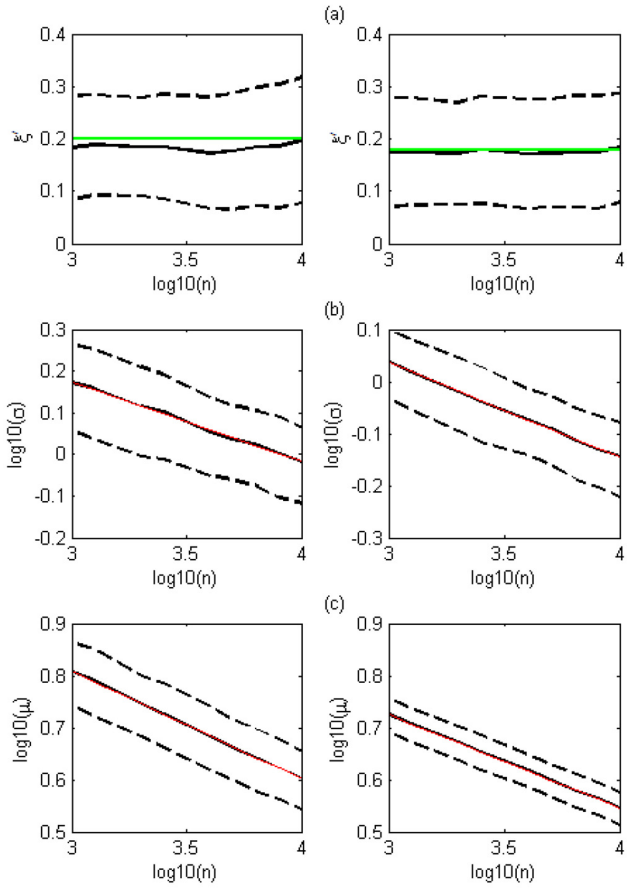


FIG. 10. g_2 observable (a) ξ' vs $\log_{10}(n)$; (b) $\log_{10}(\sigma)$ vs $\log_{10}(n)$; (c) $\log_{10}(\mu)$ vs $\log_{10}(n)$. Left: Hènon map, Right: Lozi map. Dotted lines represent one standard deviation from the mean values (black solid lines), the red lines represent he power law fits, the green lines give the theoretical values.

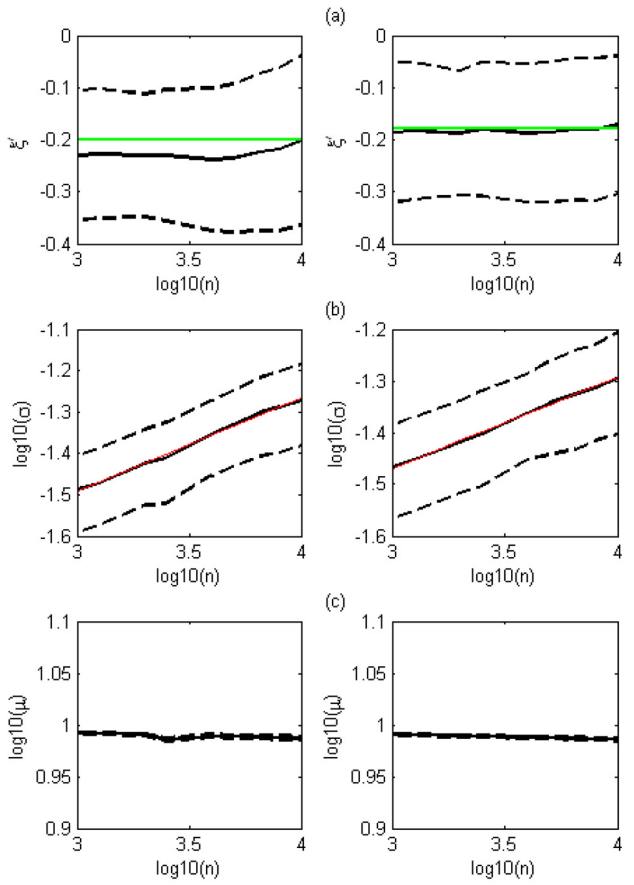


FIG. 11. g_3 observable. (a) ξ' vs $\log_{10}(n)$; (b) $\log_{10}(\sigma)$ vs $\log_{10}(n)$; (c) $\log_{10}(\mu)$ vs $\log_{10}(n)$. Left: Hènon map, Right: Lozi map. Dotted lines represent one standard deviation from the mean values (black solid lines), the red lines represent he power law fits, the green lines give the theoretical values.

respect to the Baker map. The same applies for the other two parameters, σ for a g_1-type observable (see Fig. 9) and μ for a g_3-type observable (Fig. 11), whose theoretical values given in Eqs. (19) and (21) do not depend on n.

Analogously to the case of the Baker map, the agreement between the experimentally inferred and the theoretical dependence on n of the other parameters is excellent. The estimates of μ of a g_1-type observable given in Fig. 9(c) feature a very pronounced linearity with the logarithm of n, with the angular coefficient matching almost exactly the theoretically predicted value given in Eq. (19). Moreover, the log-log plots of the GEV inferred parameters vs. n given in Figs. 7(b), 7(c), and 8(b) clarify the presence of an excellent agreement between the experimental value and the theoretical predictions of power-law dependence presented in Eqs. (19)–(21). All of this applies when we consider the averages estimates with respect to many points ζ distributed according to the invariant measure. The signature of the local fluctuations is reflected in the presence of relatively large uncertainties around the average values. Table II contains the estimates of the information dimension Δ for the attractors of the Hènon attractor and Lozi maps obtained by inverting the expressions given in Eqs. (19)–(21) as discussed before for the Baker map. While the agreement is fairly good for most independent estimates of Δ, the best results are obtained, just as in the case of the Baker map, starting from the parameter μ in the case of g_1-type and g_2-type observables.

While overall more than satisfactory, the agreement between the true Δ and its estimates in clearly worse for the three maps considered here than in the case of the IFS systems. The slower convergence for these maps may be related to the difficulties experienced computing the dimension with all box-counting methods, as shown in Grassberger and Procaccia (1983) and Badii and Politi (1987). In that case, it has been proven that the number of points that are required to cover a fixed fraction of the attractor support diverges faster than the number of boxes itself for this kind of non-uniform attractor. In our case, the situation is similar since we consider balls around the initial condition ζ. As pointed out, the best result for the dimension is achieved using the parameters provided by g_1 observable since the logarithm modulation of the distance captures actual extremes while weighting less possible outliers.

IV. FINAL REMARKS

Extreme value theory is attracting a lot of interest both in terms of extending pure mathematical results and in terms of applications to many fields of social and natural science. As an example, in geophysical applications, it is crucial to

have a tool for quantifying, understanding, and forecasting climatic extremes and events such as strong earthquakes and floods see Coles (2001); Kotz and Nadarajah (2000); and Beirlant *et al.* (2004); see also the recent informative review by Ghil *et al.*, 2011.

Whereas the classical EVT deals with stochastic processes, it is of obvious relevance to understand, whereas it could rigorously be used to study the outputs of deterministic dynamical systems. Recently, the existence of extreme value laws for dynamical systems preserving an absolutely continuous invariant measure or a singular invariant measure has been proven if strong mixing properties or exponential hitting time statistics on balls are satisfied. An extensive summary of the main theoretical results, detailed numerical investigations in this directions, as well as the explanation of the fundamental reason why these results do not apply for regular dynamical systems, have been recently given by Faranda *et al.*, 2011.

Nonetheless, it is crucial to consider the case of dynamical systems, whose invariant measure is not absolutely continuous with respect to Lebesgue, as this is the general case of many models of non-equilibrium, forced and dissipative systems (Ruelle, 1989). In this work, we have extended the results presented in Faranda *et al.* (2011) to this class of dynamical systems using the BM approach. Note that in another paper, we have shown that using the POT method, which leads to the GPD family of distributions rather than to the GEV's, the difference between regular and mixing systems is immaterial (Lucarini *et al.*, 2012).

In this paper, we have showed that it is indeed possible to formulate an EVT for this kind of systems when special observable functions of the distance between the iterated orbit and the initial condition are chosen. The three classical extreme value types for the limit distribution laws for maxima and the GEV distributions are absolutely continuous functions. We show that GEV models can be used to study the distribution of maxima, whose empirical cdf, in general, is singular continuous and has *plateaux* just in correspondence of the holes of the Cantor set, whenever this one is featured in the invariant set. This could be easily explained by the very nature of our observables which measure the distance with respect to a given point: there are distances which are not allowed when such distances are computed from points in the holes. It should be stressed that such a cumulative distribution function, which is a sort of devil staircase and therefore a singular continuous function, in any way could converge to a GEV distribution.

The strength of our approach relies in the possibility of inferring the values of the normalising constants needed for the EVT theory and the qualitative properties of the distribution (e.g., whether it belongs to the type 1, type 2, or type 3 family) by a GEV fitting procedure on the unnormalized data. This method, whose strength has been emphasized in Faranda *et al.* (2011a) in the case of a.c.i.m., works also for singular measures, provided we keep in mind that in this case the fitting procedure contains a sort of extrapolation needed to smooth out the gaps of the Cantor sets.

The observables studied in the context of EVT for dynamical systems are related to distances between the orbit and its initial condition, where such condition must be generic. It is interesting to observe that on Cantor sets the notion of *generic point* is not so obvious as for smooth manifolds which support Lebesgue measure: in this latter case, in fact, one could suppose that each point accessible for numerical iterations is generic with respect to an invariant measure which is in turn absolutely continuous. This notion of genericity is restored on a singular attractor by considering the SRB measure. In the case of iterated function systems, where uncountably many measures are available, we have a precise manner to identify them and this will be reflected in the different dimensions produced by the numerical computation of the parameters of the GEV. The possibility to discriminate among different singular measures having all the same topological support is another indication of the validity and of the efficiency of our approach.

We have also shown that the parameters of the distribution are intimately related to the information dimension of the invariant set. Actually, there are several independent ways to deduce the information dimension of the attractor from the estimates of the GEV parameters. We have tested our conjecture with numerical experiments on different low-dimensional maps such as: the Cantor ternary set, the Sierpinskij triangle, IFS with non-uniform weights, and non-trivial strange attractors such as those pertaining to the Baker, Lozi, and Hènon maps. The estimates of Δ are in agreement with the theoretical values in all cases considered. It is interesting to observe that the algorithm described with the selection of maxima acts like a magnifying glass on the neighborhood of the initial condition. In this way, we have both a powerful tool to study and highlight the fine structure of the attractor (in particular deriving the local dimension $D(\zeta)$ around a generic point ζ), but, on the other hand, we can obtain global properties averaging on different initial conditions. Therefore, the extremes can be thought of as geometric indicators of the local properties of the attractor. Even if we are dealing with very simple maps for which many properties are known, it is clear from numerical experiments that it is not so obvious to observe a good convergence to the GEV distribution. Even if we are able to compute very large statistics and the results are consistent with theoretical values, the error range is wide if compared to the experiment for maps with a.c.i.m. measures that we have carried out in Faranda *et al.* (2011a). This should be taken into consideration each time this statistics is applied in a predictive way to systems, whose attractor has multifractal properties. To compute the errors, we have taken into account the uncertainty on the fits and the uncertainty on the ensemble.

In the case of an experimental temporal series, for which the underlying dynamics is unknown, a classical problem is to obtain the dimensionality of the attractor of the dynamical systems which generated it. This can be achieved through the so called Ruelle-Takens delay embedding where, starting from the time series of an observable $O(n)$, we can construct the multivariate vectors in a P-dimensional space

$$\phi(n) = [O(n), O(n+1), ..., O(n+P-1)]$$

and study the geometrical properties using the recurrence qualification analysis (Marwan *et al.*, 2007). The minimum

value of P needed to reconstruct the actual dimension d^* is given by $P = [2d^*] + 2$. Using the procedure described in this paper, we could increase P until the value of the dimensionality of the reconstructed system stabilizes. In this way, we could get d^*. Otherwise, by starting with a large embedding dimension P, we could directly determine d^*, thus deducing the minimum value of P to be used in the recurrence quantification analysis. We will test this strategy in a future publication.

The theory and the algorithm presented in this work and in Faranda *et al.* (2011a) allow to study in detail the recurrence of an orbit around a point: this is due to the particular choice of the observables that require to compute distances between initial and future states of the system. Understanding the behavior of a dynamical system in a neighborhood of a particular initial condition is of great interest in many applications. As an example, in weather forecast and climate, it is important to study the recurrence of patterns (the so called analogues). In principle, applying the extreme value statistics to the output of meteorological models, will make it possible to infer dynamical properties related to the closest return towards a certain weather pattern. EVT will give information not only about the probability distribution of the extrema but also about the scaling of the measure of a ball centered on the chosen initial condition providing an insight to the dynamical structure of the attractor.

ACKNOWLEDGMENTS

S.V. was supported by the CNRS-PEPS Project *Mathematical Methods of Climate Models* and by the French ANR *PERTURBATIONS*. He also thanks the GDRE Grefi-Mefi for having supported exchanges with Italy. V.L. and D.F. acknowledge the financial support of the EU FP7-ERC project NAMASTE "Thermodynamics of the Climate System". The authors are grateful to the anonymous reviewers and to Jorge Freitas for many helpful comments.

Altmann, E. G., Hallerberg, S., and Kantz, H., "Reactions to extreme events: Moving threshold model," Physica A **364**, 435–444 (2006).

Badii, R., and Politi, A., "Renyi dimensions from local expansion rates," Phys. Rev. A **35**(3), 1288 (1987).

Balakrishnan, V., Nicolis, C., and Nicolis, G., "Extreme value distributions in chaotic dynamics," J. Stat. Phys. **80**(1), 307–336 (1995).

Barnsley, M. F., *Fractals Everywhere* (Morgan Kaufmann, 2000).

Barnsley, M. F., and Demko, S., "Iterated function systems and the global construction of fractals," Proc. R. Soc. London, Ser. A **399**(1817), 243 (1985).

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (with contributions by D. De Waal and C. Ferro), *Statistics of Extremes: Theory and Applications* (John Wiley & Sons, Inc., Chichester, 2004).

Benedicks, M., and Carleson, L., "The dynamics of the Hénon map," Ann. Math. **133**(1), 73–169 (1991).

Benedicks, M., and Young, L.-S., "Sinai-Bowen-Ruelle measures for certain Hénon maps," Invent. Math. **112**, 541–576 (1993).

Bessis, D., Fournier, J. D., Servizi, G., Turchetti, G., and Vaienti, S., "Mellin transforms of correlation integrals and generalized dimension of strange sets," Phys. Rev. A **36**(2), 920 (1987).

Bessis, D., Paladin, G., Turchetti, G., and Vaienti, S., "Generalized dimensions, entropies, and liapunov exponents from the pressure function for strange sets," J. Stat. Phys. **51**(1), 109–134 (1988).

Brodin, E., and Kluppelberg, C., "Extreme value theory in finance," in *Encyclopedia of Quantitative Risk Assessment*, edited by B. Everitt and E. Melnick (Wiley, 2008).

Burton, P. W., "Seismic risk in southern Europe through to India examined using Gumbel's third distribution of extreme values," Geophys. J. R. Astron. Soc. **59**(2), 249–280 (1979).

Coles, S., *An Introduction to Statistical Modeling of Extreme Values* (Springer, 2001).

Collet, P., "Statistics of closest return for some non-uniformly hyperbolic systems," Ergod. Theory Dyn. Syst. **21**(02), 401–420 (2001).

Cornell, C. A., "Engineering seismic risk analysis," Bull. Seismol. Soc. Am. **58**(5), 1583 (1968).

Cruz, M. G., *Modeling, Measuring and Hedging Operational Risk* (John Wiley & Sons, 2002).

Cutler, C. D., and Dawson, D. A., "Estimation of dimension for spatially distributed data and related limit theorems," J. Multivariate Anal. **28**(1), 115–148 (1989).

Davison, A. C., "Modelling excesses over high thresholds, with an application," in *Statistical Extremes and Applications*, edited by J. Tiago de Oliveira (Reidel, 1984), pp. 461–482.

Davison, A. C., and Smith, R. L., "Models for exceedances over high thresholds," J. R. Stat. Soc. Ser. B (Methodol.) **52**(3), 393–442 (1990).

Eckmann, J. P., and Ruelle, D., "Ergodic theory of chaos and strange attractors," Rev. Mod. Phys. **57**, 617–656 (1985).

Faranda, D., Lucarini, V., Turchetti, G., and Vaienti, S., "Numerical convergence of the block-maxima approach to the generalized extreme value distribution," J. Stat. Phys. **145**(5), 1156–1180 (2011a).

Faranda, D., Lucarini, V., Turchetti, G., and Vaienti, S., "Generalized extreme value distribution parameters as dynamical indicators of stability," Int. J. Bifurcation Chaos (to be published, 2011b).

Felici, M., Lucarini, V., Speranza, A., and Vitolo, R., "Extreme value statistics of the total energy in an intermediate complexity model of the mid-latitude atmospheric Jet. Part I: Stationary case," J. Atmos. Sci. **64**, 2137–2158 (2007).

Fisher, R. A., and Tippett, L. H. C.,"Limiting forms of the frequency distribution of the largest or smallest member of a sample," Proc. Cambridge Philosophical Society **24**, 180–90 (1928).

Freitas, A. C. M., and Freitas, J. M., "On the link between dependence and independence in extreme value theory for dynamical systems," Stat. Probab. Lett. **78**(9), 1088–1093 (2008).

Freitas, A. C. M., Freitas, J. M., and Todd, M., "Extremal index, hitting time statistics and periodicity," e-print arXiv:1008.1350 (2010a).

Freitas, A. C. M., Freitas, J. M., and Todd, M., "Hitting time statistics and extreme value theory," Probab. Theory Relat. Fields **147**, 675–710 (2010b).

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

Friederichs, P., and Hense, A., "Statistical downscaling of extreme precipitation events using censored quantile regression," Mon. Weather Rev. **135**(6), 2365–2378 (2007).

Ghil, M., *et al.*, "Extreme events: Dynamics, statistics and prediction," Nonlinear Processes Geophys. **18**, 295–350 (2011).

Gnedenko, B., "Sur la distribution limite du terme maximum d'une série aléatoire," Ann. Math. **44**(3), 423–453 (1943).

Grassberger, P., and Procaccia, I., "Characterization of strange attractors," Phys. Rev. Lett. **50**(5), 346–349 (1983).

Gumbel, E. J., "The return period of flood flows," Ann. Math. Stat. **12**(2), 163–190 (1941).

Gupta, C., Holland, M., and Nicol, M., "Extreme value theory and return time statistics for dispersing billiard maps and flows, Lozi maps and Lorenz-like maps," Ergod. Theory Dyn. Syst. **31**, 1363–1390 (2011).

Hentschel, H. G. E., and Procaccia, I., "The infinite number of generalized dimensions of fractals and strange attractors," Physica D **8**(3), 435–444 (1983).

Holland, M., Nicol, M., and Török, A., "Extreme value distributions for non-uniformly hyperbolic dynamical systems," Trans. Am. Math. Soc. **364**, 661–688, (2012).

Hosking, J. R. M., "L-moments: Analysis and estimation of distributions using linear combinations of order statistics," J. R. Stat. Soc. Ser. B (Methodol.) **52**(1), 105–124 (1990).

Kaplan, J., and Yorke, J. A., "Chaotic behavior of multidimensional difference equations," Lect. Notes Math. **730**, 204–227 (1979).

Katz, R. W., "Extreme value theory for precipitation: Sensitivity analysis for climate change," Adv. Water Resour. **23**(2), 133–139 (1999).

Katz, R. W., and Brown, B. G., "Extreme events in a changing climate: Variability is more important than averages," Climatic Change **21**(3), 289–302 (1992).

Katz, R. W., Brush, G. S., and Parlange, M. B., "Statistics of extremes: Modeling ecological disturbances," Ecology **86**(5),1124–1134 (2005).

Kotz, S., and Nadarajah, S., *Extreme Value Distributions: Theory and Applications* (World Scientific, 2000).

Leadbetter, M. R., Lindgren, G., and Rootzen, H., *Extremes and Related Properties of Random Sequences and Processes* (Springer, New York, 1983).

Ledrappier, F., "Some relations between dimension and lyapounov exponents," Commun. Math. Phys. **81**(2), 229–238 (1981).

Ledrappier, F., and Young, L. S., "The metric entropy of diffeomorphisms, part I: Characterization of measures satisfying pesin's entropy formula," Ann. Math. **122**, 509–539 (1985a).

Ledrappier, F., and Young, L. S., "The metric entropy of diffeomorphisms, part II: Relations between entropy, exponents and dimension," Ann. Math. **122**, 540–574 (1985b).

Lilliefors, H. W., "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," J. Am. Stat. Assoc. **62**(318), 399–402 (1967).

Lucarini, V., Faranda, D., and Wouters, J., "Universal behavior of extreme value statistics for selected observables of dynamical systems," J. Stat. Phys. **147**(1), 63–73 (2012).

Marwan, N., Carmen Romano, M., Thiel, M., and Kurths, J., "Recurrence plots for the analysis of complex systems," Phys. Rep. **438**(5–6), 237–329 (2007).

Nicolis, C., Balakrishnan, V., and Nicolis, G., "Extreme events in deterministic dynamical systems," Phys. Rev. Lett. **97**(21), 210602 (2006).

Pesin, B. Y., *Dimension Theory in Dynamical Systems: Contemporary Views And Applications* (Chicago lectures in mathematics) (University of Chicago Press, 1998).

Ruelle, D., *Chaotic Evolution and Strange Attractors* (Cambridge University Press, 1989).

Smith, R. L., "Threshold methods for sample extremes," in *Statistical Extremes and Applications*, Vol. 131 (Vimeiro, 1984), pp. 621–638.

Smith, R. L., "Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone," Stat. Sci. **4**(4), 367–377 (1989).

Sornette, D., Knopoff, L., Kagan, Y. Y., and Vanneste, C., "Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes," J. Geophys. Res. **101**(B6), 13883, doi:10.1029/96JB00177 (1996).

Sprott, J. C., *Chaos and Time-Series Analysis* (Oxford University Press, 2003).

Sveinsson, O. G. B., and Boes, D., "Regional frequency analysis of extreme precipitation in northeastern colorado and fort collins flood of 1997," J. Hydrol. Eng. **7**, 49 (2002).

Todorovic, P., and Zelenhasic, E., "A stochastic model for flood analysis," Water Resour. Res. **6**(6), 1641–1648 (1970).

Vannitsem, S., "Statistical properties of the temperature maxima in an intermediate order quasi-geostrophic model," Tellus, Ser. A **59**(1), 80–95 (2007).

Vitolo, R., Holland, M. P., and Ferro, C. A. T., "Robust extremes in chaotic deterministic systems," Chaos **19**, 043127 (2009a).

Vitolo, R., Ruti, P. M., Dell'Aquila, A., Felici, M., Lucarini, V., and Speranza, A., "Accessing extremes of mid-latitudinal wave activity: Methodology and application," Tellus A **61**(1), 35–49 (2009b).

Young, L. S., "Dimension, entropy and Lyapunov exponents," Ergod. Theory Dyn. Syst. **2**(01), 109–124 (1982).

Young, L. S., "Bowen-Ruelle measures for certain piecewise hyperbolic maps," Am. Math. Soc. **287**(1), 41–48 (1985).

Chapter 6

Article: Universal Behaviour of Extreme Value Statistics for Selected Observables of Dynamical systems

J Stat Phys (2012) 147:63–73
DOI 10.1007/s10955-012-0468-z

Universal Behaviour of Extreme Value Statistics for Selected Observables of Dynamical Systems

Valerio Lucarini · Davide Faranda · Jeroen Wouters

Received: 2 October 2011 / Accepted: 17 March 2012 / Published online: 28 March 2012

Abstract The main results of the extreme value theory developed for the investigation of the observables of dynamical systems rely, up to now, on the block maxima approach. In this framework, extremes are identified with the block maxima of the time series of the chosen observable, in the limit of infinitely long blocks. It has been proved that, assuming suitable mixing conditions for the underlying dynamical systems, the extremes of a specific class of observables are distributed according to the so called Generalised Extreme Value (GEV) distribution. Direct calculations show that in the case of quasi-periodic dynamics the block maxima are not distributed according to the GEV distribution. In this paper we show that considering the exceedances over a given threshold instead of the block-maxima approach it is possible to obtain a Generalised Pareto Distribution also for extremes computed in systems which do not satisfy mixing conditions. Requiring that the invariant measure locally scales with a well defined exponent—the local dimension—, we show that the limiting distribution for the exceedances of the observables previously studied with the block maxima approach is a Generalised Pareto distribution where the parameters depend only on the local dimensions and the values of the threshold but not on the number of observations considered. We also provide connections with the results obtained with the block maxima approach. In order to provide further support to our findings, we present the results of numerical experiments carried out considering the well-known Chirikov standard map.

1 Introduction

Extreme value theory was originally introduced by Fisher and Tippett [1] and formalised by Gnedenko [2], who showed that the distribution of the maxima of a sample of independent

V. Lucarini (✉) · D. Faranda · J. Wouters
Klimacampus, Institute of Meteorology, University of Hamburg, Grindelberg 5, 20144, Hamburg, Germany
e-mail: valerio.lucarini@zmaw.de

V. Lucarini
Department of Mathematics and Statistics, University of Reading, Reading, UK

identically distributed (i.i.d.) stochastic variables converges under very general conditions to a member of the so-called Generalised Extreme Value (GEV) distribution. The attention of the scientific community to the problem of understanding extreme value theory is growing, as this theory is crucial in a wide class of applications for defining risk factors such as those related to instabilities in the financial markets and to natural hazards related to seismic, climatic and hydrological extreme events. Even if the probability of extreme events decreases with their magnitude, the damage that they may bring increases rapidly with the magnitude as does the cost of protection against them. From a theoretical point of view, extreme values of observables are related to large fluctuations of the corresponding underlying system. An extensive account of recent results and relevant applications is given in [3].

The traditional block-maxima (BM) approach for the statistical inference of extremes is related to the original results by Gnedenko [2]: we partition the experimental time series into bins of fixed length, we extract the maximum of each bin, and fit the selected data to the GEV distribution family using methods such as maximum likelihood estimation (MLE) or L-moments. See [4] for a detailed account of this methodology. The selection of just one maximum in a fixed period may lead to the loss of relevant information on the large fluctuations of the system, especially when there are many large values in a given period [5]. This problem can be taken care of by considering several of the largest order statistics instead of just the largest one. For such maxima distributions we expect convergence to the Generalised Pareto Distribution (GPD), which was adopted by Pickands III [6] and Balkema and De Haan [7] in order to model the exceedances over a given threshold. This is usually referred to as the peak over threshold (POT) approach. Also this approach has been widely adopted for studying empirically natural extreme phenomena such as those related to waves, winds, temperatures, earthquakes and floods [8–10].

Both the BM and POT approaches were originally designed to study extreme values for series of i.i.d. variables. In this case it is well known that a strong connection exists between the two methodologies, as we have that if block maxima obey the GEV distribution, then exceedances over some high threshold will have an associated GPD. Moreover, the shape parameter of the GPD and that of the corresponding GEV distribution are identical [11]. As a result, several practical methods (e.g. Hill's and Pickands' estimators) developed for estimating the shape parameter of the GEV distribution of the extremes of a given time series are actually based upon comparing the GPD fits at various thresholds [12, 13]. In practical terms, it appears that, while the BM and POT approaches provide equivalent information in the asymptotic limit of infinitely long time series, the GPD statistics is more robust when realistic, finite time series are considered (see, e.g., [14]).

In recent years, especially under the influence of the rapid development of numerical modelling in the geophysical sciences and of its applications for the investigation of the socio-economic impacts of extreme events, it has become of great relevance to understand whether it is possible to apply the extreme value theory on the time series of observables of deterministic dynamical systems. Several papers have addressed this issue at a general level. A first important result is that when a dynamical system has a regular (periodic of quasi-periodic) behaviour, we do not expect, in general, to find convergence to GEV distributions for the extremes of any observable. These results have been presented by Lacroix [15]. Numerical experiments on climate models of various degrees of complexity have shown that the speed of convergence (if any) of the statistical properties of the extremes definitely depends on the chosen climatic variable of interest [4, 16–18].

A mathematical approach to extreme value theory in dynamical systems was proposed in the landmark paper by Collet [19], which has paved the way for the recent results obtained in the last few years [20–23]. The starting point of all of these investigations has been to associate to the stationary stochastic process given by the dynamical system, a new stationary

independent sequence which obeys one of the classical three extreme value laws introduced by Gnedenko [2]. The assumptions which are necessary to observe a GEV distribution in dynamical systems rely on the choice of suitable observables (specific functions of the distance between the orbit and the initial condition, chosen to be on the attractor) and the fulfillment of particular mixing conditions that guarantee the independence of subsequent maxima. Recent studies have shown that the resulting parameters of the GEV distributions can be expressed as simple functions of the local (around the initial condition) dimension of the attractor, and numerical investigations have clarified the conditions under which convergence to the theoretical GEV distributions can be satisfactorily achieved when considering finite time series [24–26].

In this paper, we wish to attempt to present a new angle on the extreme value theory for dynamical systems by using the POT rather than the BM approach. We choose the same class of observables presented in [20–26] and show that, assuming that locally the measure scales with a local (or point-wise) dimension [27], it is possible to obtain by direct integration a GPD for the threshold exceedances when considering a generic orbit of a dynamical system. The parameters will depend only on the choice of the threshold and, more importantly, on the local dimension, regardless of the nature of the flow, be it chaotic or regular. We will also discuss why in our set-up the POT approach is more efficient than the BM for estimating extreme value properties for a finite time series. Note that Castillo and Hadi [5] had already pointed out that in the case of periodic or quasi-periodic motion the BM approach to the evaluation of the extreme value statistics is inefficient, basically because in the limit of very large blocks, we tend to observe always the same maximum in all bins. To support our analytical results we provide numerical experiments that we carry out considering the classic Chirikov standard map [28].

This paper is organised as follows. In Sect. 2 we recapitulate the extreme value theory for dynamical systems obtained using the BM approach. In Sect. 3 we present our general results obtained using the POT approach. In Sect. 4 we provide support to our investigation by examining the results of the numerical simulations performed on the standard map. In Sect. 5 we present our final remarks and future scientific perspectives.

2 Block Maxima Approach: Generalised Extreme Value Distributions in Dynamical Systems

Gnedenko [2] studied the convergence of maxima of i.i.d. variables $X_0, X_1, \ldots, X_{m-1}$ with cumulative distribution function (cdf) $F(x) = P\{M_m \leq x\}$ where

$$M_m = \max\{X_0, \ldots, X_{m-1}\}. \tag{1}$$

Under general hypothesis on the nature of the parent distribution of data, Gnedenko [2] showed that the asymptotic distribution of maxima belongs to a single family of generalised distribution called GEV distribution whose cdf can be written as:

$$F_{GEV}(x; \mu, \alpha, \kappa) = \mathrm{e}^{-t(x)} \tag{2}$$

where

$$t(x) = \begin{cases} (1 + \kappa(\frac{x-\mu}{\alpha}))^{-1/\kappa} & \text{if } \kappa \neq 0, \\ \mathrm{e}^{-(x-\mu)/\alpha} & \text{if } \kappa = 0. \end{cases} \tag{3}$$

This expression holds for $1+\kappa(x-\mu)/\alpha > 0$, where $\kappa \in \mathbb{R}$ is the shape parameter (also called the tail index), $\mu \in \mathbb{R}$ is the location parameter and $\alpha > 0$ the scale parameter. In Faranda *et al.* [25] we have shown that the parameters μ and α are such that $b_m \to \mu$ and $a_m \to 1/\alpha$, for m large enough, where b_m and a_m are the normalising sequences described in Freitas *et al.* [21]. When $\kappa \to 0$, the distribution corresponds to a Gumbel type (Type 1 distribution). When the index is positive, it corresponds to a Fréchet (Type 2 distribution); when the index is negative, it corresponds to a Weibull (Type 3 distribution).

Let us consider a dynamical system $(\Omega, \mathcal{B}, \nu, f)$, where Ω is the invariant set in some manifold, usually $\mathbb{R}^d$, $\mathcal{B}$ is the Borel σ-algebra, $f : \Omega \to \Omega$ is a measurable map and ν an f-invariant Borel measure. Among the invariant measures that a dynamical systems possess we will always refer to the natural (invariant) measure which can be thought as the one really accessible when numerical experiments are performed. In order to adapt the extreme value theory to dynamical systems, following [20–23], we consider the stationary stochastic process $X_0, X_1, \ldots$ given by:

$$X_m(x) = g\big(\mathrm{dist}\big(f^m(x), \zeta\big)\big) \quad \forall m \in \mathbb{N} \tag{4}$$

where 'dist' is a distance in the ambient space Ω, ζ is a given point and g is an observable function. The partial maximum in the BM approach is defined as in Eq. (1). Defining $r = \mathrm{dist}(x, \zeta)$, we consider the three classes of observables $g_i, i = 1, 2, 3$:

$$g_1(r) = -\log(r), \tag{5}$$

$$g_2(r) = r^{-\beta}, \tag{6}$$

$$g_3(r) = C - r^{\beta} \tag{7}$$

where C is a constant and $\beta > 0 \in \mathbb{R}$. Using the observable g_i we obtain convergence of the statistics of the block maxima of their time series obtained by evolving the dynamical system to the Type i distribution if one can prove two sufficient conditions called D_2 and D', which imply a type of independence of the series of extremes resulting from the mixing of the underlying dynamics [20, 29]. The conditions cannot be simply related to the usual concepts of strong or weak mixing, but are indeed not obeyed by observables of systems featuring a regular dynamics.

A connection also exists between the existence of extreme value laws and the statistics of first return and hitting times, which provide information on how fast the point starting from the initial condition ζ comes back to a neighbourhood of ζ, as shown by Freitas *et al.* [21] and Freitas *et al.* [30]. In particular, they proved that for dynamical systems possessing an invariant measure ν, the existence of an exponential hitting time statistics on balls around ν-almost any point ζ implies the existence of extreme value laws for one of the observables of type $g_i, i = 1, 2, 3$, described above. The converse is also true, namely if we have an extreme value law which applies to the observables of type $g_i, i = 1, 2, 3$, achieving a maximum at ζ, then we have exponential hitting time statistics to balls with center ζ. Recently these results have been generalised to local returns around balls centered at periodic points [22] by requiring conditions reminiscent of D_2 and D', but taking into account the fact that some trajectories stay close to the periodic trajectory.

In Faranda *et al.* [24–26] we analysed both from an analytical and numerical point of view the extreme value distribution in a wide class of low dimensional maps. We divided the time series of length k of the g_i observables into n bins each containing the same number m of observations, and selected the maximum (or the minimum) value in each of them [31]. We showed that at leading order (the formulas are asymptotically correct for $m, k \to \infty$),

the GEV parameters in mixing maps can be written in terms of m (or equivalently n) and the local dimension of the attractor D. We have:

- g_1-type observable:

$$\alpha = \frac{1}{D}, \qquad \mu \sim \frac{1}{D}\ln(k/n), \qquad \kappa = 0; \tag{8}$$

- g_2-type observable:

$$\alpha \sim n^{-\frac{\beta}{D}}, \qquad \mu \sim n^{-\frac{\beta}{D}}, \qquad \kappa = \frac{\beta}{D}; \tag{9}$$

- g_3-type observable:

$$\alpha \sim n^{\frac{\beta}{D}}, \qquad \mu = C, \qquad \kappa = -\frac{\beta}{D}. \tag{10}$$

Moreover, we clearly showed that other kinds of distributions not belonging to the GEV family are observed for quasi-periodic and periodic motions.

3 The Peak over Threshold Approach: Generalised Pareto Distributions in Dynamical Systems

We define an exceedance as $z = X - T$, which measures by how much X exceeds the threshold T. As discussed above, under the same conditions under which the block maxima of the i.i.d. stochastic variables X obey the GEV statistics, the exceedances z are asymptotically distributed according to the Generalised Pareto Distribution [11]:

$$F_{GPD}(z;\xi,\sigma) = \begin{cases} 1-(1+\frac{\xi z}{\sigma})^{-1/\xi} & \text{for } \xi \neq 0, \\ 1-\exp(-\frac{z}{\sigma}) & \text{for } \xi = 0, \end{cases} \tag{11}$$

where the range of z is $0 \leq z \leq -\sigma/\xi$ if $\xi < 0$ and $0 \leq z < \infty$ if $\xi \geq 0$. We consider the same set up described in the previous section and take into account the observables $g = g(\text{dist}(x,\zeta)) = g(r)$, such that g achieves a maximum g_{max} for $r = 0$ (finite or infinite) and is strictly decreasing bijection in a neighbourhood of 0. We study the exceedance above a threshold T defined as $T = g(r^*)$. We obtain an exceedance every time the distance between the orbit of the dynamical system and ζ is smaller than r^*. Therefore, we define the exceedances $z = g(r) - T$. By Bayes' theorem, we have that $P(r < g^{-1}(z+T)|r < g^{-1}(T)) = P(r < g^{-1}(z+T))/P(r < g^{-1}(T))$. In terms of ν, an invariant measure of the system, we have that the probability $H_{g,T}(z)$ of observing an exceedance of at least z given that an exceedance occurs is given by:

$$H_{g,T}(z) \equiv \frac{\nu(B_{g^{-1}(z+T)}(\zeta))}{\nu(B_{g^{-1}(T)}(\zeta))}. \tag{12}$$

Obviously, the value of the previous expression is 1 if $z = 0$. In agreement with the conditions given on g, the expression contained in Eq. (12) monotonically decreases with z and vanishes when the radius is given by $g^{-1}(g_{max})$. Note that the corresponding cdf is given by $F_{g,T}(z) = 1 - H_{g,T}(z)$. In order to address the problem of extremes, we have to consider small radii. At this regard we will invoke, and assume, the existence of the following limit

$$\lim_{r\to 0}\frac{\log \nu(B_r(\zeta))}{\log r} = D(\zeta), \quad \text{for } \zeta \text{ chosen } \nu\text{-a.e.}, \tag{13}$$

where $D(\zeta)$ is the local dimension of the attractor [27]. Whenever this limit holds for ν-almost any choice of the point ζ, the corresponding limit is proved to be equal to the Hausdorff dimension of the measure, defined as the infimum of the Hausdorff dimension of the measurable sets of full ν measure, see [32]. This limit is also called the information dimension in [33]. It is interesting to notice that the existence of the previous limit ν-almost everywhere could be proved for a large class of dynamical systems, especially hyperbolic, and it could be expressed in terms of suitable ratios of the entropies and of the Lyapunov exponents of the measure, see again [32], for the two dimensional case and [34, 35], for the multidimensional case. Therefore, we rewrite the following expression for the tail probability of exceedance:

$$H_{g,T}(z) \sim \left(\frac{g^{-1}(z+T)}{g^{-1}(T)} \right)^{D} \tag{14}$$

where we have dropped the ζ dependence of D. By substituting g with specific observable we are considering, we obtain explicitly the corresponding extreme value distribution law.

By choosing an observable of the form given by either g_1, g_2, or g_3, we derive as extreme value distribution law one member of the Generalised Pareto Distribution family given in Eq. (11). Results are detailed below:

- g_1-type observable:

$$\sigma = \frac{1}{D}, \qquad \xi = 0; \tag{15}$$

- g_2-type observable:

$$\sigma = \frac{T\beta}{D}, \qquad \xi = \frac{\beta}{D}; \tag{16}$$

- g_3-type observable:

$$\sigma = \frac{(C-T)\beta}{D}, \qquad \xi = -\frac{\beta}{D}. \tag{17}$$

The previous expressions show that there is a simple algebraic link between the parameters of the GPD and the local dimension of the attractor around the point ζ. This implies that the statistics of extremes provides us with a new algorithmic tool for estimating the local fine structure of the attractor. These results show that it is possible to derive general properties for the extreme values of the observables g_1, g_2, or g_3 independently on the qualitative properties of the underlying dynamics, be the system periodic, quasi-periodic, or chaotic. Therefore, by taking the POT instead of the BM approach, we are able to overcome the mixing conditions (or the requirements on the properties of the hitting time statistics) proposed in [19–23]. In [24–26] we had proposed that the link between the extreme value theory and the local properties of the invariant measure in the vicinity of the point ζ can be explained by the fact that selecting the extremes of the observables g_1, g_2, or g_3 amounts to performing a zoom around ζ. The detailed analysis of the extremes using the BM approach allows to understand whether the underlying dynamics is mixing or not. Instead, the POT approach does not provide such dynamical information, but it rather gives a very efficient way to estimate the local dimension in all cases, including for systems that do not satisfy mixing conditions. The reason of this is clear if we observe that in the POT approach the cdf depends only on the properties of the natural measure in the neighbourhood of ζ (technically, on the properties of X_0 only) and on the threshold, whereas in the BM approach the cdf depends on the time evolution of the system as it results by Eq. (1). This restricts the class of systems whose

maxima converge to a GEV distribution to those satisfying the mixing conditions detailed in Freitas *et al.* [21].

3.1 Relationship Between the BM and POT Approaches

The relation between GEV and GPD parameters have been already discussed in literature in case of i.i.d. variables [13, 14, 36, 37]. Coles [13] and Katz *et al.* [36] have proven that the cdf of the GEV defined as $F_{GEV}(z;\mu,\alpha,\kappa)$ can be asymptotically written as that of GPD under a high enough threshold as follows:

$$\begin{aligned} F_{GEV}(z;\mu,\alpha,\kappa) &\sim F_{GPD}(z;T,\sigma,\xi) \\ &= 1-\left[1-\xi\left(\frac{z-T}{\sigma}\right)\right]^{1/\xi} \end{aligned} \tag{18}$$

where $\xi=\kappa$, $\sigma=\alpha+\xi(T-\mu)$, and $T=\mu+\frac{\sigma}{\xi}(\lambda^{-\xi}-1)$, with $\ln(\alpha)=\ln(\sigma)+\xi\ln(\lambda)$. In the present case, we have to compare Eqs. (8)–(10) for GEV with Eqs. (15)–(17) for GPD, keeping in mind that the GEV results hold only under the mixing conditions discussed before. While it is immediate to check that $\kappa=\xi$, the other relationships are valid in the limit of large n, as expected.

4 Numerical Investigation

The standard map [38] is an area-preserving chaotic map defined on the bidimensional torus, and it is one of the most widely-studied examples of dynamical chaos in physics. The corresponding mechanical system is usually called a kicked rotator. It is defined as:

$$\begin{cases} y_{t+1}=y_t-\frac{K}{2\pi}\sin(2\pi x_t) & \mod 1, \\ x_{t+1}=x_t+y_t+1 & \mod 1. \end{cases} \tag{19}$$

The dynamics of the map given in Eq. (19) can be regular or chaotic. For $K\ll 1$ the motion follows quasi periodic orbits for all initial conditions, whereas if $K\gg 1$ the motion turns to be chaotic and irregular. An interesting behaviour is achieved when $K\sim 1$: in this case we have coexistence of regular and chaotic motions depending on the chosen initial conditions [39].

We perform for various values of K ranging from $K=10^{-4}$ up to $K=10^2$ an ensemble of 200 simulations, each characterised by a different initial condition ζ randomly taken on the bidimensional torus, and we compute for each orbit the observables g_i, $i=1,2,3$. Choosing as point ζ exactly the initial condition ensures that the orbit is really passing through the point ζ which is desirable for quasi-periodic and periodic motions. In each case, the map is iterated until obtaining a statistics consisting 10^4 exceedances, where the threshold $T=7\cdot 10^{-3}$ and $\beta=3$. We have checked that all the results are indeed robust with respect to the choice of the threshold and of the value of β. For each orbit, we fit the statistics of the 10^4 exceedances values of the observables to a GPD distribution, using a MLE estimation [5] implemented in the MATLAB© function *gpdfit* [40]. The results are shown in Fig. 1 for the inferred values of ξ and σ and should be compared with Eqs. (15)–(17). When $K\ll 1$, we obtain that the estimates of ξ and σ are compatible with a dimension $D=1$ for all the initial conditions: we have that the ensemble spread is negligible. Similarly, for $K\gg 1$, the estimates for ξ and σ agree remarkably well with having a local dimension

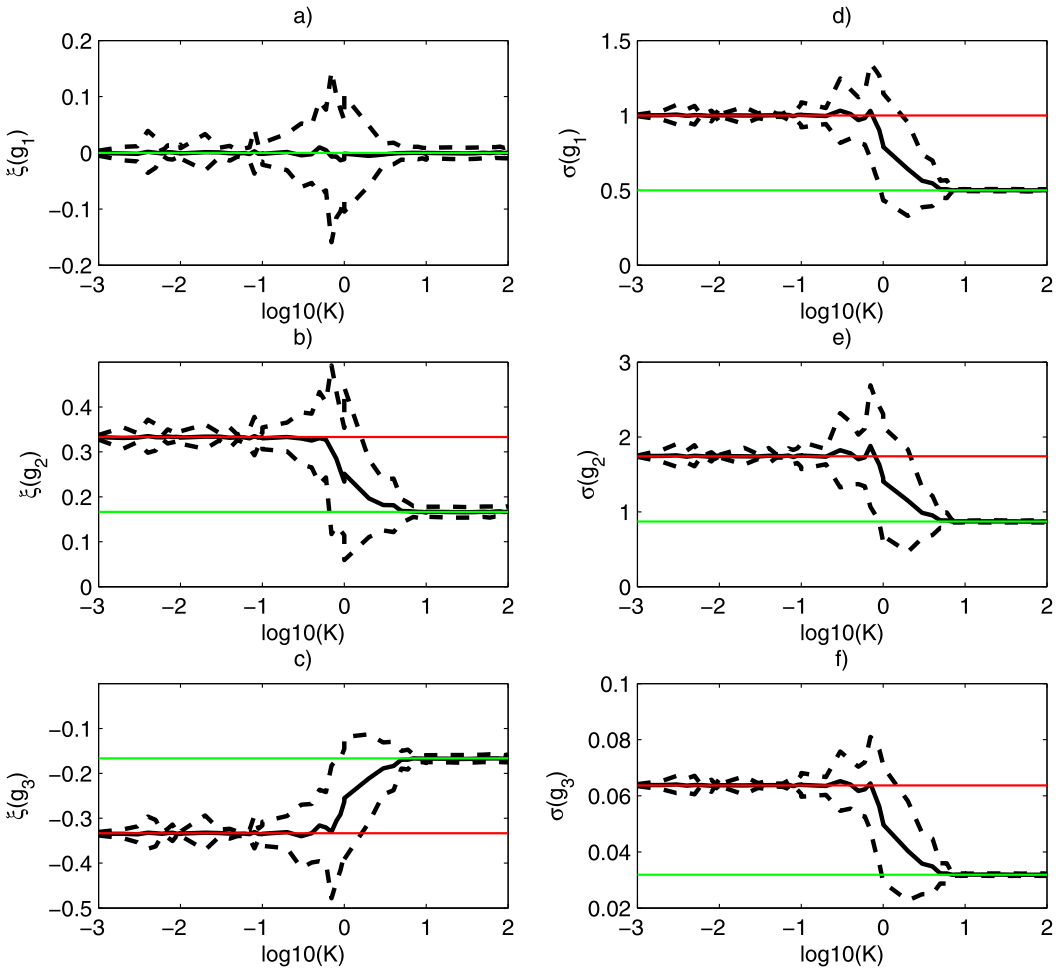


Fig. 1 GPD parameters for the observables g_i, $i = 1, 2, 3$, computed over orbits of the standard map, for various values of the constant K. For each value of K, results refer to an ensemble of 200 randomly chosen initial conditions ζ. The notation $p(g_i)$ indicates the parameter p computed using the extreme value statistics of the observable g_i. **(a)** $\xi(g_1)$ VS K, **(b)** $\xi(g_2)$ VS K, **(c)** $\xi(g_3)$ VS K, **(d)** $\sigma(g_1)$ VS K, **(e)** $\sigma(g_2)$ VS K, **(f)** $\sigma(g_3)$ VS K. *Black solid lines*: ensemble-average value. *Black dotted lines*: ensemble spread evaluated as one standard deviation of the ensemble. *Green lines*: theoretical values for regular orbits. *Red lines*: theoretical values for chaotic orbits (Color figure online)

$D = 2$ for all the initial conditions. In the transition regime, which occurs for $K \simeq 1$, the ensemble spread is much higher, because the scaling properties of the measure is different among the various initial conditions. As expected, the ensemble averages of the parameters change monotonically from the value pertaining to the regular regime to that pertaining to the chaotic regime with increasing values of K. Basically, this measures the fact that the so-called regular islands shrink with K. Note that in the case of the observable g_1, the estimate of the ξ is robust in all regimes, even if, as expected, in the transition between low and high values of K the ensemble spread is larger. These results can also be compared with the analysis presented in Faranda *et al.* [24], where we used the BM approach. In that case, the values obtained in the regular regions were inconsistent with the GEV findings, the very reason being that the dynamics was indeed not mixing. Here, it is clear that the statistics can be computed in all cases, and we have a powerful method for discriminating regular from chaotic behaviours through the analysis of the inferred local dimension.

5 Conclusions

The growing attention of the scientific community in understanding the behaviour of extreme values have led, in the recent past, to the development of an extreme value theory for dynamical systems. In this framework, it has been shown that the statistics of extreme value can be linked to the statistics of return in a neighbourhood of a certain initial conditions by choosing special observables that depend on the distance between the iterated trajectory and some fixed point of the phase space. Until now, rigorous results have been obtained assuming the existence of an invariant measure for the dynamical systems and the fulfillment of independence requirement on the series of maxima achieved by imposing D' and D_2 mixing conditions, or, alternatively, assuming an exponential hitting time statistics [20–23]. The parameters of the GEV distribution obtained choosing as observables the function g_i, $i = 1, 2, 3$, defined above depend on the local dimension of the attractor D and numerical algorithms to perform statistical inference can be set up for mixing systems having both absolutely continuous and singular invariant measures [25, 26, 41, 42].

Taking a complementary point of view, in this paper we have studied the statistics of exceedances for the same class of observables and derived the limiting distributions assuming only the existence of an invariant measure and the possibility to define a local dimension D around the point ζ of interest. To prove that the limiting distribution is a GPD we did not use any further conditions. In particular no assumptions on the mixing nature of the maxima sequence have been made. This means that a GPD limiting distribution holds for the statistics of exceedance also for non-mixing dynamical systems and it depends only on the threshold value and on the local dimension once we choose the observables g_i, $i = 1, 2, 3$, but not on the number of observations. Other functions can converge to the limiting behaviour of the GPD family if they asymptotically behave like the g_i's (compare the discussion in [21]). Nonetheless, this requires, analytically, to perform separately the limit for the threshold T going to g_{max} and that for the radius of the ball going to zero. In practical terms, this requires, potentially, much stricter selection criteria for the exceedances when finite time series are considered.

We note that, as the parameter ξ is inversely proportional to D, one can expect that each time we analyse systems of intermediate or high dimensionality, the distributions for g_2 and g_3 observables will be virtually indistinguishable from what obtained considering the g_1 observable: the $\xi = 0$ is in this sense an *attracting* member of the GPD family. This may also explain, at least qualitatively, why the Gumbel ($k = 0$ for the GEV family) distribution is so efficient in describing the extremes of a large variety of natural phenomena [3].

The universality of this approach allows to resolve the debate on whether there exists or not a general way to obtain information about extreme values for quasi-periodic motions where the GEV limit does not apply [15]. This is due to the fact that considering several of the largest order statistics instead of just the largest one we can study orbits where numerous exceedance are observed in a given block, as it happens for systems with periodic or quasi-periodic behaviours. As an example, one may consider a system with multiple commensurable frequencies: choosing a block length larger than or equal to the smallest common period, we select always the same value for any considered observable. On the other hand for mixing maps we find, as expected, an asymptotic equivalence of the results obtained via the BM and via the POT approach.

Whereas the BM approach allows to reconstruct dynamical features of the dynamical systems discriminating between chaotic and regular behaviour, the POT approach provides a way to reconstruct local properties of invariant measures for any kind of dynamical systems: once a threshold is chosen and a suitable exceedance statistics is recorded, we can compute

the local dimension for different initial conditions taken on the attractor. This is true also in the opposite direction: if the knowledge of the exact value for the local dimension is available, once we chose a small enough radius (threshold), it is possible to compute a priori the properties of extremes without doing any further computations. In fact, the expression for the parameters (15)–(17) do not contain any dependence on the properties of the dynamics except the local dimension.

Besides the analytical results, we have proved that POT approach is easily accessible for numerical investigations. The algorithm used to perform numerical simulations is versatile and computationally accessible: unlike the BM algorithm that requires a very high number of iterations to obtain unbiased statistics, using the POT approach we can fix a priori a value for the threshold and the number of maxima necessary to construct the statistics. With the simulations carried out on the standard maps, we obtain meaningful results with a much smaller statistics with respect to what observed when considering the BM approach.

We hope that the present contribution may provide a tool that is not only useful for the analysis of extreme events itself, but also for characterising the dynamical structure of attractors by giving a robust way to compute the local dimensions, with the new possibility of embracing also the case of quasi-periodic motions.

Future investigations will include the systematic study of the impact on the extreme value statistics of adding stochastic noise to regular and chaotic deterministic dynamical systems, and the use of the Ruelle response theory [43, 44] to study the modulation of the statistics of extremes due to changes in the internal or external parameters of the system, especially in view of potentially relevant applications in geophysics such as in the case of climate studies [45, 46].

Acknowledgements The authors acknowledge various useful exchanges with R. Blender and K. Fraedrich, and the financial support of the EU-ERC project NAMASTE-Thermodynamics of the Climate System.

References

1. Fisher, R., Tippett, L.: Proc. Camb. Philos. Soc. **24**, 180 (1928)
2. Gnedenko, B.: Ann. Math. **44**, 423 (1943)
3. Ghil, M., et al.: Nonlinear Process. Geophys. **18**, 295 (2011)
4. Felici, M., Lucarini, V., Speranza, A., Vitolo, R.: J. Atmos. Sci. **64**, 2137 (2007)
5. Castillo, E., Hadi, A.: J. Am. Stat. Assoc. **92**, 1609–1620 (1997)
6. Pickands, J. III: Ann. Stat. **3**, 119–131 (1975). ISSN 0090-5364
7. Balkema, A., De Haan, L.: Ann. Prob. **2**, 792–804 (1974)
8. Simiu, E., Heckert, N.: J. Struct. Eng. **122**, 539 (1996)
9. Bayliss, A., Jones, R.: Peaks-over-threshold flood database. Report 121, Institute of Hydrology, Crowmarsh Gifford (1993)
10. Pisarenko, V., Sornette, D.: Pure Appl. Geophys. **160**, 2343 (2003)
11. Leadbetter, M., Lindgren, G., Rootzen, H.: Extremes and Related Properties of Random Sequences and Processes. Springer, New York (1983)
12. Embrechts, P., Kluppelberg, C., Mikosh, T.: Modelling Extremal Events. Springer, Berlin (1997)
13. Coles, S.: An Introduction to Statistical Modeling of Extreme Values. Springer, Berlin (2001)
14. Ding, Y., Cheng, B., Jiang, Z.: Adv. Atmos. Sci. **25**, 507 (2008)
15. Lacroix, Y.: Isr. J. Math. **132**, 253 (2002)
16. Kharin, V., Zwiers, F., Zhang, X.: J. Climate **18**, 5201 (2005)
17. Vannitsem, S.: Tellus A **59**, 80 (2007). ISSN 1600-0870
18. Vitolo, R., Ruti, P., Dell'Aquila, A., Felici, M., Lucarini, V., Speranza, A.: Tellus A **61**, 35 (2009). ISSN 1600-0870
19. Collet, P.: Ergod. Theory Dyn. Syst. **21**, 401 (2001). ISSN 0143-3857
20. Freitas, A., Freitas, J.: Stat. Probab. Lett. **78**, 1088 (2008). ISSN 0167-7152
21. Freitas, A., Freitas, J., Todd, M.: Prob. Theory Related Fields **147**, 675–710 (2010)

22. Freitas, A., Freitas, J., Todd, M.: Preprint. arXiv:1008.1350 (2010)
23. Gupta, C., Holland, M., Nicol, M.: Lozi-like maps, and Lorenz-like maps. Preprint (2009)
24. Faranda, D., Lucarini, V., Turchetti, G., Vaienti, S.: Int. Jou. Bif. Chaos (in press). arXiv:1107.5972 (2011)
25. Faranda, D., Lucarini, V., Turchetti, G., Vaienti, S.: J. Stat. Phys. **145**, 1156–1180 (2012). doi:10.1007/s10955-011-0234-7
26. Lucarini, V., Faranda, D., Turchetti, G., Vaienti, S.: Preprint. arXiv:1106.2299 (2011)
27. Bandt, C.: In: Gazeau, J., Nešetřil, J., Rovan, B. (eds.) Physics and Theoretical Computer Science: From Numbers and Languages to (Quantum) Cryptography Security, pp. 91–112. IOS Press, Amsterdam (2007)
28. Chirikov, B., Shepelyansky, D.: Scholarpedia **3**, 3550 (2008)
29. Bruin, H., Todd, M.: Stoch. Dyn. **9**, 81 (2009)
30. Freitas, A., Freitas, J., Todd, M.: J. Stat. Phys. **142**, 108–126 (2011)
31. Coles, S., Heffernan, J., Tawn, J.: Extremes **2**, 339 (1999). ISSN 1386–ISSN 1999
32. Young, L.-S.: Ergod. Theory Dyn. Syst. **2**, 109–124 (1982)
33. Eckmann, J., Ruelle, D.: Rev. Mod. Phys. **57**, 617 (1985)
34. Ledrappier, F., Young, L.S.: Ann. Math. **122**, 509 (1985)
35. Ledrappier, F., Young, L.S.: Ann. Math. **122**, 540 (1985)
36. Katz, R., Brush, G., Parlange, M.: Ecology **86**, 1124 (2005). ISSN 0012-9658
37. Malevergne, Y., Pisarenko, V., Sornette, D.: Appl. Financ. Econ. **16**, 271 (2006)
38. Chirikov, B.: Research concerning the theory of nonlinear resonance and stochasticity. Preprint 267. Institute of Nuclear Physics, Novosibirsk (1969)
39. Ott, E.: Chaos in Dynamical Systems. Cambridge University Press, New York (2002)
40. Martinez, W., Martinez, A.: Computational Statistics Handbook with MATLAB. CRC Press, Boca Raton (2002)
41. Vitolo, R., Holland, M., Ferro, C.: Chaos: an interdisciplinary. J. Nonlinear Sci. **19**, 043127 (2009)
42. Holland, M., Vitolo, R., Rabassa, P., Sterk, A., Broer, H.: Physica D **241**, 497–513 (2012). arXiv:1107.5673
43. Ruelle, D.: Phys. Lett. A **245**, 220 (1998)
44. Ruelle, D.: Nonlinearity **22**, 855 (2009)
45. Abramov, R., Majda, A.: Nonlinearity **20**, 2793 (2007)
46. Lucarini, V., Sarno, S.: Nonlinear Process. Geophys. **18**, 7 (2011)

Chapter 7

Article: Extreme Value Statistics for dynamical systems with noise

EXTREME VALUE STATISTICS FOR DYNAMICAL SYSTEMS WITH NOISE

DAVIDE FARANDA, JORGE MILHAZES FREITAS, VALERIO LUCARINI, GIORGIO TURCHETTI, AND SANDRO VAIENTI

Abstract. We study the distribution of maxima (*Extreme Value Statistics*) for sequences of observables computed along orbits generated by random transformations. The underlying, deterministic, dynamical system can be regular or chaotic. In the former case, we will show that by perturbing rational or irrational rotations with additive noise, an extreme value law will appear, regardless of the intensity of the noise, while unperturbed rotations do not admit such limiting distributions. In the case of deterministic chaotic dynamics, we will consider observables specially designed to study the recurrence properties in the neighbourhood of periodic points. The exponential limiting law for the distribution of maxima is therefore modified by the presence of the *extremal index*, a positive parameter not larger than one, whose inverse gives the average size of the clusters of extreme events. The theory predicts that such a parameter is unitary when the system is perturbed randomly. We perform sophisticated numerical tests to assess how strong is the impact of noise level, when finite time series are considered. We find agreement with the asymptotic theoretical results but also non-trivial behaviour in the finite range. Finally, we focus on a prototypical dissipative system such as the Hénon map with standard values for the parameters. In a previous paper, we showed that it is possible to relate the parameters describing the type of the limiting extreme value laws with the local dimension of the attractor of the system. Here, we find that while it is well known that adding an arbitrary weak noise results into smoothing the invariant measure, it is far from trivial to detect such an effect. Our results suggest that in many applications where finite datasets can be produced or analysed one must be careful in assuming that the smoothing nature of noise prevails over the underlying deterministic dynamics.

2010 *Mathematics Subject Classification.* 37A50, 60G70, 37B20, 60G10, 37A25, 37H99.

Key words and phrases. Random dynamical systems, Extreme Values, Hitting Times Statistics, Extremal Index.

Acknowledgements SV thanks the French ANR *Perturbations* and the CNRS-PEPS *Mathematical Methods for Climate Models* for support. SV thanks H. Aytach for useful discussions. DF acknowledges Jeroen Wouters for useful discussions. DF and VL acknowledge the support of the NAMASTE project. NAMASTE has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement No. 257106. JMF is thankful to A.C.M. Freitas and M. Todd for fruitful conversations. JMF was partially supported by FCT (Portugal) grant SFRH/BPD/66040/2009, by FCT project PTDC/MAT/099493/2008 and CMUP, which is financed by FCT through the programs POCTI and POSI, with national and European Community structural funds. JMF and SV are also supported by FCT project PTDC/MAT/120346/2010, which is financed by national and European Community structural funds through the programs FEDER and COMPETE. .

1. Introduction

The main purpose of this paper is to study the extremal behaviour of randomly perturbed dynamical systems. By extremal behaviour we mean its statistical performance regarding the existence of Extreme Value Laws (EVLs), or in other words, the existence of distributional limits for the partial maxima of stochastic processes arising from such systems. In many aspects of natural and social sciences and engineering, the statistical properties of the extremes of a system are usually relevant tied with actual risk assessment. This is an element of why the theory of extremes has received such a great deal of attention over the years. The motivation for considering randomly perturbed systems follows from the fact that, very often, dynamical systems are used to model human activities or natural phenomena and the fact that errors made by observations usually take a random character which can be well described by the random perturbation formalism. On the other hand, random noise is often added in numerical modelling in order to represent the lack of knowledge on (or practical impossibility of representing) some of the processes taking place in the system of interest, often characterised by small spatial and/or temporal scales, whose explicit representation is virtually impossible. See [31] for a comprehensive discussion on this issue in a geophysical setting. Also, noise is often considered as a "good" component to add when performing numerical simulations, because of its ability to smoothen the invariant measure and basically remove unphysical solutions [28].

Very recently, in [2], the authors proved the first results (up to our knowledge) regarding the existence of EVLs for piecewise expanding systems which are randomly perturbed by additive noise. They observed that adding noise has a "smoothing" effect in terms of eliminating all possible clustering of exceedances. The distributional properties of the maxima of stationary sequences is driven by the appearance of exceedances of high thresholds. For independent and identically distributed (i.i.d.) processes, the exceedances appear scattered through the time line. For dependent processes this may not necessarily hold and clustering of exceedances may occur. The Extremal Index (EI), which we will denote by $0 \leq \theta \leq 1$, is a parameter that quantifies the amount of clustering. No clustering means that $\theta = 1$ and strong clustering means that θ is close to 0. For deterministic hyperbolic systems, it has been shown, in [17, 12, 19, 2], that a dichotomy holds: either an EI $0 < \theta < 1$ is present at the (repelling) periodic points, or the EI is always 1 at every other point. We remark that Hirata [18] had already observed the two types of behaviour but not the dichotomy. Coming back to the "smoothing" effect of adding noise, to be more precise, in [2], the authors show that, for the randomly perturbed systems considered, the EI is always 1. The main tool used there was decay of correlations against L^1 observables, which held for the unperturbed and perturbed systems considered. We remark that this property implies an excellent mixing behaviour of both systems.

One of the main achievements of this paper is the extension of the results in [2], by showing that adding noise still has a smoothing effect when the unperturbed system is not mixing at all, or even worse: when the system is actually periodic. To be more precise, for unperturbed systems, we consider rotations on the circle (irrational or not) and show

that by adding absolutely continuous noise with respect to the Haar measure, we can always prove the existence of EVLs with the EI being 1 everywhere. The analysis of these systems will be made using Fourier transforms to compute decay of correlations of specific observables. Observe that for the original unperturbed systems the lack of mixing does not allow to prove EVLs in the usual sense. Hence, here, we have a more drastic transition which motivates the question of whether it is possible to distinguish numerically the real nature of the underlying (unperturbed) systems when we look at the extremal statistics of the randomly perturbed data.

Another main goal of this paper is to support the analytical discussion with numerical simulations devised to show how, for finite samples, the extremal behaviour may or not follow the asymptotic results. This has obvious relevance in terms of applications. One of the key aspects surfacing from the numerical analysis of extremes that we carried out is that the estimation of some dynamical and geometrical properties of the underlying physical measure strongly depends on a combination of the intensity of the noise and the length of the data sample.

There are two main experimental issues deriving from stochastic perturbations. The first concerns with the general claim that the addition of noise has a "smoothing" effect over the physical measure and enhances the chaoticity of the underlying deterministic dynamical systems. As we have mentioned, we expect to see this in terms of extremal behaviour, at least asymptotically. This claim will be weighed up against the simulations showing that we may need to integrate the system for a very long time to observe any changes in the measure structure. Moreover, for higher dimensional systems featuring stable and unstable manifolds, the stochastic perturbations should be directed on the stable direction so that the "smoothing" effect of the measure is more effective.

The second issue is related to the level of noise needed to modify the deterministic properties and, in this case in particular, the sensitivity of the extremal type behaviour to the the noise level.

To discuss this latter issue we will use the fact that the numerical round off is comparable to a random noise on the last precision digit [20]. This observation will allow us to claim that, for systems featuring periodic or quasi periodic motions (as the rotations on the circle), such a noise level is not sufficient for producing detectable changes regarding the observed extremal behaviour.

On the other hand, the analysis for generic points of chaotic systems, such as the full shift on three symbols, will show that the intrinsic chaotic behaviour of the map dims the effect of the noise and makes the perturbed system indistinguishable from a deterministic one. If, instead, the analysis is carried out at periodic points, the effect of the perturbation will be easily noticeable in the disruption of periodicity and its consequent clustering of exceedances, which can be detected by estimation of the EI.

The disruption of periodicity will be further analysed through simulations on the quadratic map. We will show, for finite sample behaviour, how the EI depends on the intensity of

the noise. We will conclude the analysis suggesting a way to analyse the disruption of periodicity in general systems where the nature of the noise may be unknown.

We will also present some numerical simulations on the Hénon map, which features a stable and an unstable directions, homoclinic tangencies and the coexistence of two different attractors. In a previous paper we have shown that it is possible to relate the EVL parameters to the local dimension of the attractor of a system possessing a singular invariant physical measure [29]. This latter experiment will be the gateway to go beyond the maps presented in this paper and to suggest a general procedure to analyse issues related to the stochastic perturbations of dynamical systems in order to frame the analysis carried out here in a more general setup. In fact, we find that it is far from trivial to detect the smoothing effect of noise on the invariant measure when finite time series are considered.

We would like to remark that such considerations have great relevance in the context of Axiom A systems and having statistical mechanical applications in mind when addressing the problem of the applicability of the fluctuation-dissipation theorem (FDT). While in the deterministic case the non-smooth nature of the measure along the stable manifold makes it impossible to apply straightforwardly the FDT (see discussion in [33, 34, 35, 27, 30]), the addition of even a very small amount of random forcing in principle "cures" the singularities of the measure and makes sure that in principle the FDT can be used to related forced and free motions (see [1, 21]). Our results suggest that one must be careful in assuming that such a desirable effect is really detectable when finite datasets are analysed.

2. Extreme values, the system and the perturbation

Our main purpose is to study the extremal behaviour of randomly perturbed dynamical systems. This will be done by analysing the partial maxima of stationary stochastic processes arising from such systems. We will start by presenting the main concepts regarding the extreme values on a general framework of stationary stochastic processes. Then we will turn to the construction of stationary stochastic processes deriving from dynamical systems and their random perturbations by additive noise, which will be the object of our study.

2.1. Extreme values – definitions and concepts.

Let $X_0, X_1, \ldots$ be a stationary stochastic process. We denote by F the cumulative distribution function (d.f.) of X_0, *i.e.*, $F(x) = \mathbb{P}(X_0 \leq x)$. Given any d.f. F, let $\bar{F} = 1 - F$ and u_F denote the right endpoint of the d.f. F, *i.e.*, $u_F = \sup\{x : F(x) < 1\}$. We say we have an exceedance of the threshold $u < u_F$ whenever

$$U(u) := \{X_0 > u\} \tag{2.1}$$

occurs.

We define a new sequence of random variables $M_1, M_2, \ldots$ given by

$$M_n = \max\{X_0, \ldots, X_{n-1}\}. \tag{2.2}$$

Definition 1. We say that we have an *Extreme Value Law* (EVL) for M_n if there is a non-degenerate d.f. $H : \mathbb{R} \to [0,1]$ with $H(0) = 0$ and, for every $\tau > 0$, there exists a sequence of levels $u_n = u_n(\tau)$, $n = 1, 2, \ldots$, such that

$$n\mathbb{P}(X_0 > u_n) \to \tau, \quad \text{as } n \to \infty, \tag{2.3}$$

and for which the following holds:

$$\mathbb{P}(M_n \leq u_n) \to \bar{H}(\tau), \quad \text{as } n \to \infty. \tag{2.4}$$

where the convergence is meant in the continuity points of $H(\tau)$.

The motivation for using a normalising sequence u_n satisfying (2.3) comes from the case when $X_0, X_1, \ldots$ are independent and identically distributed (i.i.d.). In this i.i.d. setting, it is clear that $\mathbb{P}(M_n \leq u) = (F(u))^n$, where F is the d.f. of X_0, *i.e.*, $F(x) := \mathbb{P}(X_0 \leq x)$. Hence, condition (2.3) implies that

$$\mathbb{P}(M_n \leq u_n) = (1 - \mathbb{P}(X_0 > u_n))^n \sim \left(1 - \frac{\tau}{n}\right)^n \to \mathrm{e}^{-\tau}, \tag{2.5}$$

as $n \to \infty$. Moreover, the reciprocal is also true (see [24, Theorem 1.5.1] for more details). Note that in this case $H(\tau) = 1 - \mathrm{e}^{-\tau}$ is the standard exponential d.f.

When $X_0, X_1, X_2, \ldots$ are not independent but satisfy some mixing condition $D(u_n)$ introduced by Leadbetter in [22] then something can still be said about H. Let $F_{i_1,\ldots,i_n}$denote the joint d.f. of $X_{i_1}, \ldots, X_{i_n}$, and set $F_{i_1,\ldots,i_n}(u) = F_{i_1,\ldots,i_n}(u, \ldots, u)$.

Condition $(D(u_n))$**.** We say that $D(u_n)$ holds for the sequence $X_0, X_1, \ldots$ if for any integers $i_1 < \ldots < i_p$ and $j_1 < \ldots < j_k$ for which $j_1 - i_p > m$, and any large $n \in \mathbb{N}$,

$$\left|F_{i_1,\ldots,i_p,j_1,\ldots,j_k}(u_n) - F_{i_1,\ldots,i_p}(u_n)F_{j_1,\ldots,j_k}(u_n)\right| \leq \gamma(n,m),$$

where $\gamma(n, m_n) \xrightarrow[n\to\infty]{} 0$, for some sequence $m_n = o(n)$.

If $D(u_n)$ holds for $X_0, X_1, \ldots$ and the limit (2.4) exists for some $\tau > 0$ then there exists $0 \leq \theta \leq 1$ such that $\bar{H}(\tau) = \mathrm{e}^{-\theta\tau}$ for all $\tau > 0$ (see [23, Theorem 2.2] or [24, Theorem 3.7.1]).

Definition 2. We say that $X_0, X_1, \ldots$ has an *Extremal Index* (EI) $0 \leq \theta \leq 1$ if we have an EVL for M_n with $\bar{H}(\tau) = \mathrm{e}^{-\theta\tau}$ for all $\tau > 0$.

The notion of the EI was latent in the work of Loynes [26] but was established formally by Leadbetter in [23]. It gives a measure of the strength of the dependence of $X_0, X_1, \ldots$, so that $\theta = 1$ indicates that the process has practically no memory while $\theta = 0$, conversely, reveals extremely long memory. Another way of looking at the EI is that it gives some indication on how much exceedances of high levels have a tendency to "cluster". Namely, for $\theta > 0$ this interpretation of the EI is that θ^{-1} is the mean number of exceedances of a high level in a cluster of large observations, *i.e.*, is the "mean size of the clusters".

In the dependent case, as shown in [22], if $D(u_n)$ holds and in addition an anti clustering condition $D'(u_n)$ (see definition below) also holds, one can show that the EI is 1, which

means that $\bar{H}(\tau) = \mathrm{e}^{-\tau}$, as in the independent case. However, since the rates of mixing for dynamical systems are usually given by decay of correlations of observables in certain given classes of functions, it turns out that condition $D(u_n)$ is too strong to be checked for chaotic systems whose mixing rates are known only through decay of correlations. For that reason, motivated by Collet's work [8], in [14] the authors suggested a condition $D_2(u_n)$ which together with $D'(u_n)$ was enough to prove the existence of an exponential EVL ($\bar{H}(\tau) = \mathrm{e}^{-\tau}$) for maxima around non-periodic points z. In fact, [14, Theorem 1] states that if the following conditions hold for $X_0, X_1, \ldots$ then there exists an EVL for M_n and $H(\tau) = 1 - e^{-\tau}$.

Condition ($D_2(u_n)$)**.** We say that $D_2(u_n)$ holds for the sequence $X_0, X_1, \ldots$ if for all ℓ, t and n

$$|\mathbb{P}\left(X_0 > u_n \cap \max\{X_t, \ldots, X_{t+\ell-1} \leq u_n\}\right) - \mathbb{P}(X_0 > u_n)\mathbb{P}(M_\ell \leq u_n)| \leq \gamma(n,t), \quad (2.6)$$

where $\gamma(n,t)$ is decreasing in t for each n and $n\gamma(n,t_n) \to 0$ when $n \to \infty$ for some sequence $t_n = o(n)$.

Now, let $(k_n)_{n\in\mathbb{N}}$ be a sequence of integers such that

$$k_n \to \infty \quad \text{and} \quad k_n t_n = o(n). \quad (2.7)$$

Condition ($D'(u_n)$)**.** We say that $D'(u_n)$ holds for the sequence X_0, X_1, X_2, ... if there exists a sequence $\{k_n\}_{n\in\mathbb{N}}$ satisfying (2.7) and such that

$$\lim_{n\to\infty} n \sum_{j=1}^{\lfloor n/k_n \rfloor} \mathbb{P}(X_0 > u_n, X_j > u_n) = 0. \quad (2.8)$$

Condition $D_2(u_n)$ is much weaker than the original $D(u_n)$, and it is easy to show that it follows easily from sufficiently fast decay of correlations (see [14, Section 2]).

In the results mentioned above, condition $D'(u_n)$ prevented the existence of clusters of exceedances, which implies for example that the EVL was a standard exponential $\bar{H}(\tau) = \mathrm{e}^{-\tau}$. However, when $D'(u_n)$ does not hold, clustering of exceedances is responsible for the appearance of a parameter $0 < \theta < 1$ in the EVL which now is written as $\bar{H}(\tau) = \mathrm{e}^{-\theta\tau}$.

In [17], the authors established a connection between the existence of an EI less than 1 and periodic behaviour. This was later generalised for rare events point processes in [16] . Under the presence of periodic phenomena, the inherent rapid recurrence creates clusters of exceedances which makes it easy to check that condition $D'(u_n)$ fails (see [17, Section 2.1]). This was a serious obstacle for the application to dynamical systems since the theory developed up to [17] was based on Collet's important observation that $D'(u_n)$ could be used not only in the usual way as in Leadbetter's approach, but also to compensate the weakening of the original $D(u_n)$, which allowed the application to chaotic systems with sufficiently fast decay of correlations. To overcome this difficulty, in [17], the authors considered the annulus

$$Q_p(u) := \{X_0 > u,\ X_p \leq u\} \quad (2.9)$$

resulting from removing from $U(u)$ the points that were doomed to return after p steps, which form the smaller ball $U(u) \cap f^{-p}(U(u))$. In [17, 16], the occurrence of $Q_p(u)$ was named an *escape* since it corresponds to the realisations escaping the influence of the underlying periodic phenomena and exit the ball $U(u)$ after p iterates. Then, the main crucial observation in [17] is that the limit law corresponding to no entrances up to time n into the ball $U(u_n)$ was equal to the limit law corresponding to no entrances into the annulus $Q_p(u_n)$ up to time n (see [17, Proposition 1]). This meant that, roughly speaking, the role played by the balls $U(u)$ could be replaced by that of the annuli $Q_p(u)$, with the advantage that points in $Q_p(u)$ were no longer destined to return after just p steps.

Based in this last observation and motivated by the behaviour at repelling periodic points, in [17], the authors proposed certain general conditions (for general stationary stochastic processes) designed to prove the existence of an EI less than 1 under the presence of some sort of underlying periodic phenomenon. The first one establishes exactly the type of periodic behaviour assumed, namely:

Condition ($\mathrm{SP}_{p,\theta}(u_n)$)**.** We say that $X_0, X_1, X_2, \ldots$ satisfies condition $\mathrm{SP}_{p,\theta}(u_n)$ for $p \in \mathbb{N}$ and $\theta \in [0,1]$ if

$$\lim_{n\to\infty} \sup_{1\le j<p} \mathbb{P}(X_j > u_n | X_0 > u_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}(X_p > u_n | X_0 > u_n) \to (1-\theta) \tag{2.10}$$

and moreover

$$\lim_{n\to\infty} \sum_{i=0}^{[\frac{n-1}{p}]} \mathbb{P}(X_0 > u_n, X_p > u_n, X_{2p} > u_n, \ldots, X_{ip} > u_n) = 0. \tag{2.11}$$

Condition (2.10), when $\theta < 1$, imposes some sort of periodicity of period p among the exceedances of high levels u_n, since if at some point the process exceeds the high level u_n, then, regardless of how high u_n is, there is always a strictly positive probability of another exceedance occurring at the (finite) time p. In fact, if the process is generated by a deterministic dynamical system $f : \mathcal{X} \to \mathcal{X}$ and f is continuous then (2.10) implies that z is a periodic point of period p, *i.e.*, $f^p(z) = z$.

The next two conditions concern to the dependence structure of $X_0, X_1, \ldots$ and can be described as being obtained from $D_2(u_n)$ and $D'(u_n)$ by replacing balls by annuli. Let $\mathscr{Q}_n(u_n) := \bigcap_{j=0}^{n-1} f^{-j}(Q(u_n)^c)$. Note that while the occurrence of the event $\{M_n \le u_n\}$ means that no entrance in the ball $\{X_0 > u_n\}$ has occurred up to time n, the occurrence of $\mathscr{Q}_n(u_n)$ means that no entrance in the annulus $Q(u_n)$ has occurred up to time n.

Condition ($D^p(u_n)$)**.** We say that $D^p(u_n)$ holds for the sequence $X_0, X_1, X_2, \ldots$ if for any integers ℓ, t and n

$$|\mathbb{P}\left(Q_{p,0}(u_n) \cap \mathscr{Q}_{p,t,\ell}(u_n)\right) - \mathbb{P}(Q_{p,0}(u_n))\mathbb{P}(\mathscr{Q}_{p,0,\ell}(u_n))| \le \gamma(n,t),$$

where $\gamma(n,t)$ is non increasing in t for each n and $n\gamma(n,t_n) \to 0$ as $n \to \infty$ for some sequence $t_n = o(n)$.

This condition requires some sort of mixing by demanding that an escape at time 0 is an event which gets more and more independent from an event corresponding to no escapes during some period, as the time gap between these two events gets larger and larger. The main advantage of this condition when compared to Leadbetter's $D(u_n)$ (or others of the same sort) is that it follows directly from sufficiently fast decay of correlations as observed in Section 3.3 of [17], on the contrary to $D(u_n)$.

Assuming $D^p(u_n)$ holds let $(k_n)_{n\in\mathbb{N}}$ be a sequence of integers such that

$$k_n \to \infty \quad \text{and} \quad k_n t_n = o(n). \tag{2.12}$$

Condition $(D'_p(u_n))$**.** We say that $D'_p(u_n)$ holds for the sequence $X_0, X_1, X_2, \ldots$ if there exists a sequence $\{k_n\}_{n\in\mathbb{N}}$ satisfying (2.12) and such that

$$\lim_{n\to\infty} n \sum_{j=1}^{[n/k_n]} \mathbb{P}(Q_{p,0}(u_n) \cap Q_{p,j}(u_n)) = 0. \tag{2.13}$$

This last condition is very similar to Leadbetter's $D'(u_n)$ from [23], except that instead of preventing the clustering of exceedances it prevents the clustering of escapes by requiring that they should appear scattered fairly evenly through the time interval from 0 to $n-1$.

In [17, Theorem 1], it was proved that if a stationary stochastic process satisfies conditions $\mathrm{SP}_{p,\theta}(u_n)$, $D^p(u_n)$ and $D'_p(u_n)$ then we have an EVL for M_n with $\bar{H}(\tau) = \mathrm{e}^{-\theta\tau}$.

2.2. **Stochastic processes arising from randomly perturbed systems.** To simplify the exposition we will consider one dimensional maps f on the circle $\mathbb{S}^1$ or on the unit interval I (from now on we will use I to identify both spaces), provided, in the latter case, that the image $f(I)$ is strictly included into I. We will perturb them with additive noise, namely, we introduce the family of maps

$$f_\omega(x) = f(x) + \omega$$

where we have to take the mod-1 operation if the map is considered on the circle. The quantity ω is chosen on the interval $\Omega_\varepsilon = [-\varepsilon, \varepsilon]$ with distribution ϑ_ε, which we will take equivalent to Lebesgue on Ω_ε. Consider now an i.i.d. sequence ω_k, $k \in \mathbb{N}$ taking values on the interval Ω_ε and distributed according to ϑ_ε. We construct the random orbit (or *random transformation*) by the concatenation

$$f^n_{\underline{\omega}}(x) = f_{\omega_n} \circ f_{\omega_{n-1}} \circ \cdots \circ f_{\omega_1}(x). \tag{2.14}$$

The role of the invariant measure is now played by the stationary measure μ_ε which is defined as

$$\iint \phi(f_{\underline{\omega}}(x))\, d\mu_\varepsilon(x)\, d\vartheta_\varepsilon^{\mathbb{N}}(\underline{\omega}) = \int \phi(x)\, d\mu_\varepsilon(x),$$

for every $\phi \in L^\infty$ (L^∞ to be intended with respect to the Lebesgue measure Leb)[1]. The previous equality could also be written as

$$\int \mathcal{U}_\varepsilon \phi \, d\mu_\varepsilon = \int \phi \, d\mu_\varepsilon$$

where the operator $\mathcal{U}_\varepsilon : L^\infty \to L^\infty$, is defined as $(\mathcal{U}_\varepsilon \phi)(x) = \int_{\Omega_\varepsilon} \phi(f_\omega(x)) d\vartheta_\varepsilon$ and it is called the *random evolution operator*.
The decay of correlations of the perturbed systems could be formulated in terms of the random evolution operator. More precisely, we will take two non-zero observables, ϕ and ψ and we will suppose that ϕ is of bounded variation with norm $||\cdot||_{BV}$ and $\psi \in L^1_m$[2] Then the correlation integral is

$$\mathrm{Cor}_\varepsilon(\phi, \psi, n) := \frac{1}{\|\phi\|_{BV} \|\psi\|_{L^1_m}} \left| \int \mathcal{U}^n_\varepsilon \phi \; \psi \, d\mu_\varepsilon - \int \phi \, d\mu_\varepsilon \int \psi \, d\mu_\varepsilon \right|$$

In the following we will use the product measure $\mathbb{P} := \mu_\varepsilon \times \vartheta_\varepsilon^{\mathbb{N}}$.

We are now in condition of defining the time series X_0, X_1, X_2,... arising from our system simply by evaluating a given observable $\varphi : I \to \mathbb{R} \cup \{+\infty\}$ along the random orbits of the system:

$$X_n = \varphi \circ f^n_{\underline{\omega}}, \quad \text{for each } n \in \mathbb{N}, \tag{2.15}$$

We assume that φ achieves a global maximum at $z \in I$; for every $u < \varphi(z)$ but sufficiently close to $\varphi(z)$, the event $\{y \in I : \ \varphi(y) > u\} = \{X_0 > u\}$ corresponds to a topological ball "centred" at z and, for every sequence $(u_n)_{n\in\mathbb{N}}$ such that $u_n \to \varphi(z)$, as $n \to \infty$, the sequence of balls $\{U_n\}_{n\in\mathbb{N}}$ given by

$$U_n := \{X_0 > u_n\} \tag{2.16}$$

is a nested sequence of sets such that

$$\bigcap_{n\in\mathbb{N}} U_n = \{z\}. \tag{2.17}$$

As explained in [14] the type of asymptotic distribution obtained depends on the chosen observables. In the simulations, we will cover the observables such that

$$\varphi(\cdot) = g(\mathrm{dist}(\cdot, z)), \tag{2.18}$$

where dist denotes a certain metric chosen on I and g is of one of the following three types:

(1) $g_1(y) = -\log(y)$ to study the convergence to the Gumbel law.
(2) $g_2(y) = y^{-1/a} \qquad a \in \mathbb{R} \quad a > 0$ for the Fréchet law.
(3) $g_3(y) = C - y^{1/a} \qquad a \in \mathbb{R} \quad a > 0 \quad C \in \mathbb{R}$ for the Weibull distribution.

[1]This choice is dictated by the fact that the stationary measure will be equivalent to Lebesgue in all the examples considered below

[2]Of course we could do other choices, but this two spaces will play a major role in the subsequent theory.

The sequences of real numbers $u_n = u_n(\tau)$, $n = 1, 2, \ldots$, are usually taken to be as one parameter linear families like $u_n = y/a_n + b_n$, where $y \in \mathbb{R}$ and $a_n > 0$, for all $n \in \mathbb{N}$. In fact, in the classical theory, one considers the convergence of probabilities of the form

$$\mathbb{P}(a_n(M_n - b_n) \leq y). \tag{2.19}$$

In this case, the *Extremal Types Theorem* says that, whenever the variables X_i are i.i.d, if for some constants $a_n > 0$, b_n, we have

$$\mathbb{P}(a_n(M_n - b_n) \leq x) \to G(x), \tag{2.20}$$

where the convergence occurs at continuity points of G, and G is non degenerate, then G belongs to one of three extreme values types (see below). Observe that τ depends on y through u_n and, in fact, in the i.i.d. case, depending on the tail of the marginal d.f. F, we have that $\tau = \tau(y)$ is of one of the following three types (for some $\alpha > 0$):

$$\tau_1(y) = \mathrm{e}^{-y} \text{ for } y \in \mathbb{R}, \quad \tau_2(y) = y^{-\alpha} \text{ for } y > 0 \quad \text{and} \quad \tau_3(y) = (-y)^{\alpha} \text{ for } y \leq 0. \tag{2.21}$$

In [24, Theorem 1.6.2], it were given sufficient and necessary conditions on the tail of the d.f. F in order to obtain the respective domain of attraction for maxima. Besides, in [24, Corollary 1.6.3] one can find specific formulas for the normalising constants a_n and b_n so that the respective extreme limit laws apply. We used these formulas to perform the numerical computations in this paper.

Remark 1. We emphasise that, for i.i.d. sequences of random variables, the limiting distribution type of the partial maxima is completely determined by the tail of the d.f. F. For the stationary stochastic processes considered here, if an EI $\theta > 0$ applies, then the same can still be said about the limiting distribution type of the partial maxima: namely, it is completely determined by the tail of the d.f. F. This statement follows from the equivalence between (2.3) and (2.5) and the definition of the EI. However, we also quote [24, Corollary 3.7.3] because we will refer to it later:

If for $X_0, X_1, \ldots$ we have an EI $\theta > 0$, then if we considered an i.i.d. sequence $Z_0, Z_1, \ldots$ so that the d.f. of Z_0 is F, the same as that of X_0, and let $\hat{M}_n = \max\{Z_0, \ldots, Z_{n-1}\}$, then the existence of normalizing sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ for which

$$\lim_{n\to\infty} \mathbb{P}(a_n(\hat{M}_n - b_n) \leq y) = G(y)$$

implies that

$$\lim_{n\to\infty} \mathbb{P}(a_n(M_n - b_n) \leq y) = G^{\theta}(y),$$

and the reciprocal is also true. Moreover, since by [24, Corollary 1.3.2] G^{θ} is of the same type of G, we can actually make a linear adjustment to the normalizing sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ so that the second limit is also G.

Remark 2. From Remark 1, in order to determine the type of extremal distribution G (recall that $G(y) = \mathrm{e}^{\tau(y)}$, where $\tau(y)$ is of one of the three types described in (2.21)) which applies to our stochastic processes $X_0, X_1, \ldots$, one needs to analyse the tail of the d.f. F. The choice of the observables in (2.18) implies that the shape of g determines the type of

extremal distribution we get. In particular, for observables of type g_i we get an extremal law of type e^{τ_i}, for $i = 1, 2, 3$. (See [15, Remark 1] for more details on this correspondence). While the type of the extremal distribution is essentially determined by the shape of the observable, in the cases when types 2 and 3 apply, *i.e.*, the Fréchet and Weibull families of distributions, respectively, the exponent α is also influenced by other quantities such as the EI and the local dimension of the stationary (invariant) measure μ_ϵ (μ). In particular, when such measure is absolutely continuous with respect to Lebesgue and its Radon-Nikodym derivative has a singularity at z, then the order of the singularity also influences the value of α.

3. Numerical experiments

For the numerical computations we will use the approach already described in [9]. It consists in considering unnormalised maxima selected using the block Maxima approach. Once computed the orbit of the dynamical systems, the series of the g_i observables are divided into m blocks of length n observations. Maxima of the observables g_i are selected for each blocks and fitted to a single family of generalised distribution called GEV distribution with d.f.:

$$F_G(x; \nu, \sigma, \kappa) = \exp\left\{ -\left[1 + \kappa\left(\frac{x-\nu}{\sigma}\right)\right]^{-1/\kappa}\right\}; \tag{3.1}$$

which holds for $1 + \kappa(x - \nu)/\sigma > 0$, using $\nu \in \mathbb{R}$ (location parameter), $\sigma > 0$ (scale parameter) and $\kappa \in \mathbb{R}$ is the shape parameter also called the tail index: when $\kappa \to 0$, the distribution corresponds to a Gumbel type (Type 1 distribution). When the index is positive, it corresponds to a Fréchet (Type 2 distribution); when the index is negative, it corresponds to a Weibull (Type 3 distribution). In this fitting procedure, as we have seen in [9], the following relations hold:

$$\text{for type 1,} \quad \kappa = 0, \qquad \nu = b_n, \qquad \sigma = 1/a_n; \tag{3.2}$$

$$\text{for type 2,} \quad \kappa = 1/\alpha, \qquad \nu = b_n + 1/a_n, \qquad \sigma = \kappa/a_n; \tag{3.3}$$

$$\text{for type 3,} \quad \kappa = -1/\alpha, \qquad \nu = b_n - 1/a_n, \qquad \sigma = -\kappa/a_n. \tag{3.4}$$

For the discrete maps like $f_\omega(x)$, we select a value for z and repeat the following operations for various values \$$\epsilon$ of the noise intensity:

(1) We construct an empirical physical measure of the system by performing a long run and saving the obtained long trajectory. The realisations $f_{\omega_{n-1}}, \cdots, f_{\omega_1}$ are constructed by taking the ω_k with the uniform distribution in Ω_ϵ.
(2) We select 500 initial conditions according to the physical measure described above; such initial conditions will be used as initial conditions (x variables) to generate 500 realisations of the stochastic process.

(3) For each of the 500 realisations, we obtain an orbit containing $r = m \cdot n = 10^6, 10^7$ data.

(4) We split the series in $m = 1000$ bins each containing $n = 1000$ (or $n = 10000$) observations. These values are chosen in agreement with the investigations previously carried out in [9], since they provide enough statistics to observe EVL for maps which satisfy the D_2 and D' conditions.

(5) We fit the GEV distribution for the $g_i, i = 1, 2, 3$ observables and the maxima and minima and study the behaviour of selected parameters.

In the forthcoming discussion we will present the results for the shape parameters $\kappa(g_i)$ and for the location parameter $\sigma(g_1)$. This choice is convenient, as explained in [29, 10], since these parameters do not depend on the number of observations n and therefore their asymptotic expected value does not change when r is modified. For the one dimensional maps considered we expect to obtain the following asymptotic parameters for systems satisfying the mixing conditions described in the former sections (see [10]):

$$\kappa(g_1) = 0 \quad \kappa(g_2) = 1/a \quad \kappa(g_3) = -1/a \quad \sigma(g_1) = 1 \tag{3.5}$$

The inference procedure follows [9] where the authors have used a Maximum Likelihood Estimation (MLE) that is well defined when the underlying physical measure is absolutely continuous. In the applications described in [29, 10] an inference via a L-moments procedure has been preferred as it provides reliable values for the parameters even when the d.f. corresponds to fractal or multi fractal measures. However, since the L-moments procedure does not give information about the reliability of the fit[3], hereby we exploit the MLE procedure since it helps to highlight the situations where the fitting does not succeed because of the poor data sampling.

The following numerical analysis is constructed in a strict correspondence with the theoretical setup: the role of the simulations is to follow step-by-step the proofs of the theorems trying to reproduce the analytical results. Whenever it occurs, the failure of the numerical approach is analysed in relation to the length of the sample and the intensity of the noise, just to highlight which parameters are crucial for the convergence procedure.

We focus on simple discrete maps as they already contain interesting features regarding the convergence issues. We start by analysing the rotations on the 1-D torus where the conditions $D_2(u_n)$ and $D'(u_n)$ are not satisfied, but they could be restored under perturbation. In the case of expanding maps of the interval, for which the existence of EVL has been already shown both analytically and numerically, we show that the computation of the EI reveals to be crucial to discriminate whether the system was randomly perturbed or not.

[3]The L-moments inference procedure does not provide any confidence intervals unless these are derived with a bootstrap procedure which is also dependent on the data sample size [6]. The MLE, on the other side, allows for easily compute the confidence intervals with analytical formulas [32].

Eventually, the analysis of the Hénon map, shows how the presence of another basin of attraction (infinity in this case) may affect the EVL regardless the initial condition chosen.

Besides the convergence issues, one of the main problems that we face regarding the estimation of the EI during the realisation of the numerical simulations is the fact that the fitting procedure used in [9], "hides" the EI. This is the case because, as we mentioned in Remark 1, when quoting [24, Corollary 3.7.3], the normalizing constants a_n and b_n can be chosen (after a simple linear rescaling) so that we get exactly the same limit law G of the corresponding i.i.d. process instead of G^θ. This prevents the detection of an EI $0 < \theta < 1$. Basically, our original procedure simply selects the best fitting constants from the sample of maxima available, which means that the EI does not surface at all. In order to overcome this difficulty, instead of applying a blind fitting, we use relations (3.2)–(3.4) (more specifically (3.2)), the knowledge about the stationary measure of the particular systems we considered and the information provided by [15, 17, 2] to compute a priori the normalizing sequences so that we can capture the value of the EI. This adjustment to the procedure revealed also very effective and the results met completely the theoretical predictions.

4. Rotations on the circle

In this section we will study the rotations of the circle and show that the effect of adding noise is enough to create enough randomness in order to make EVLs appear when they do not hold for the original system. This will be the content of Theorem 1 below. Because this nice statistical behaviour appears solely on account of the noise, it is not surprising that the numerical simulations show that for very small noise one needs a large amount of data to the detect the EVLs.

4.1. **Analytical results.** In [2, Theorem C], the following result has been proved essentially for piecewise expanding maps randomly perturbed like above.

To be more precise, let us suppose that the unperturbed map f is continuous on the circle and that [4]:
(i) the correlation integral $\mathrm{Cor}_\varepsilon(\phi, \psi, n)$ decays at least as n^{-2};
(ii) u_n satisfies: $n\mathbb{P}(X_0 > u_n) \to \tau$, as $n \to \infty$;
(iii) $U_n = \{X_0 > u_n\}$ verifies $\bigcap_{n\in\mathbb{N}} U_n = \{z\}$;
(iv) There exists $\eta > 0$ such that $d(f(x), f(y)) \leq \eta\ d(x, y)$, where $d(\cdot, \cdot)$ denotes some metric on I,
then the process $X_0, X_1, \cdots$ satisfies $D_2(u_n)$ and $D'(u_n)$, and this implies that the EVL holds for M_n so that $\bar{H}(\tau) = e^{-\tau}$. The proof strongly relies on the decay of correlations for $L^1(\mathrm{Leb})$ observables.

[4]The result is even more general and applies to multidimensional maps too, but for our concerns, especially for rotations, the 1-D case is enough. We will discuss later on about generalisation to piecewise continuous maps

In this section we show that such a result holds also for rotations (irrational or not), perturbed with additive noise. The key observation is that in the proof of [2, Theorem C], the observables entering the correlation integral are characteristic functions of intervals (see also [2, Remark 3.1]) and for such observables it is possible to prove an exponential decay of correlations for perturbed rotations by using the Fourier series technique.

In what follows we identify $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$.

Theorem 1. *Let $f : \mathbb{S}^1 \to \mathbb{S}^1$ be a rotation of angle $\alpha \in \mathbb{R}$,* i.e., *$f(x) = x + \alpha \ mod\ 1$. We perturb f additively so that the random evolution evolution of an initial state $x \in \mathbb{S}^1$ is given by (2.14), with ϑ_ε denoting the uniform distribution on $[-\varepsilon, \varepsilon]$. Let $X_0, X_1, \dots$ be a stationary stochastic process generated by the random evolution of such f, as in 2.15. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence such that items (ii) and (iii) above hold. Then, the process $X_0, X_1, \dots$ satisfies conditions $D_2(u_n)$ and $D'(u_n)$ which implies that there exists an EVL for M_n with EI equal to* 1, i.e., *$\bar{H}(\tau) = \mathrm{e}^{-\tau}$.*

For ease of exposition, we will change slightly the notation for the random perturbation of f. We write them in this way:

$$f_{\varepsilon\xi}(x) = x + \alpha + \varepsilon\xi \bmod 1 \tag{4.1}$$

where ξ is a random variable uniformly distributed over the interval $[-1, 1]$ and therefore of zero mean.[5] Let us observe that

$$f^j_{\varepsilon\bar{\xi}}(x) = x + j\alpha + \varepsilon(\xi_1 + \cdots + \xi_j) \tag{4.2}$$

In this case, it is straightforward to check (just by using the definition), that the stationary measure coincides in this case with the Lebesgue measure, Leb, on $[0, 1]$ and it is therefore independent of ε.

Let us first establish that the correlation of the right functions decays exponentially fast under the random evolution of the system.

Lemma 4.1. *Under the assumptions of Theorem 1, if $\phi = \chi_A$ and $\psi = \chi_B$, where χ denotes the characteristic function, $B = \cup_{l=1}^{\ell} B_l$, for some $\ell \in \mathbb{N}$ and $A, B_1, \dots, B_\ell \subset \mathbb{S}^1$ are connected intervals, then*

$$C_{j,\varepsilon} := \left| \int_0^1 \mathcal{U}^j_\varepsilon(\psi)\phi dx - \int_0^1 \psi dx \int_0^1 \phi dx \right| \le 4e^{-j\varepsilon^2 \log(2\pi)}, \tag{4.3}$$

as long as $\varepsilon^2 < 1 - \log 2/\log(2\pi)$.

Proof. We begin by writing the modulus of the correlation integral, $C_{j,\varepsilon}$, as follows:

$$C_{j,\varepsilon} = \left| \int_0^1 dx \frac{1}{2^j} \int_{-1}^1 d\xi_1 \cdots \int_{-1}^1 d\xi_j \psi(f^j_{\varepsilon\bar{\xi}}(x))\phi - \int_0^1 \psi dx \int_0^1 \phi dx \right|$$

[5]With respect to the previous notations, we changed Ω_ε into $[-1, 1]$, $\omega = \varepsilon\xi$, with $\xi \in [-1, 1]$ and finally ϑ_ε becomes $d\xi$ over $[-1, 1]$.

We express ϕ and ψ in terms of their respective Fourier series:

$$\psi(x) = \sum_{k\in\mathbb{Z}} \psi_k \, \mathrm{e}^{2\pi i k x} \qquad \phi(x) = \sum_{k\in\mathbb{Z}} \phi_k \mathrm{e}^{2\pi i k x}$$

Note that given $A = [a, b]$ and $B_l = [a_l, b_l]$, $l = 1, \dots, \ell$, we have

$$\phi_k = \int_0^1 \chi_A \, \mathrm{e}^{-2\pi i k x} dx = \frac{\mathrm{e}^{-2\pi i k a} - \mathrm{e}^{-2\pi i k b}}{2\pi i k},$$

and also

$$\psi_k = \int_0^1 \chi_B \mathrm{e}^{-2\pi i k x} dx = \frac{\sum_{l=1}^{\ell} \mathrm{e}^{-2\pi i k a_l} - \mathrm{e}^{-2\pi i k b_l}}{2\pi i k},$$

which implies that

$$|\psi_k|, |\phi_k| \leq 1/|k|. \tag{4.4}$$

Plugging the Fourier series of ϕ, ψ into $C_{j,\varepsilon}$, we obtain

$$C_{j,\varepsilon} = \left| \sum_{k\in\mathbb{Z}/\{0\}} \psi_k \phi_{-k} \frac{\mathrm{e}^{2\pi i k j \alpha}}{2^j} \int_{-1}^1 d\xi_1 \mathrm{e}^{2\pi i k \varepsilon \xi_1} \cdots \int_{-1}^1 d\xi_j \mathrm{e}^{2\pi i k \varepsilon \xi_j} \right| = \left| \sum_{k\in\mathbb{Z}/\{0\}} \psi_k \phi_{-k} S^j(k\varepsilon) \right|$$

where $S(x) = \frac{\sin(2\pi x)}{2\pi x}$. For this quantity we have the bounds: $|S(x)| \leq \mathrm{e}^{-x^2 \log(2\pi)}, \quad |x| < 1$, and $|S(x)| \leq \frac{1}{2\pi|x|}, \quad |x| \geq 1$.
We now continue to estimate $C_{j\varepsilon}$ as

$$C_{j\varepsilon} \leq 2 \sum_{k=1}^{1/\varepsilon} |\psi|_k |\phi_{-k}| \mathrm{e}^{-k^2\varepsilon^2 j \log(2\pi)} + \sum_{k=1/\varepsilon}^{\infty} |\psi|_k |\phi_{-k}| \frac{1}{(2\pi k\varepsilon)^j} \tag{4.5}$$

Using (4.4), we have

$$C_{j\varepsilon} \leq 2\mathrm{e}^{-\varepsilon^2 j \log(2\pi)} \sum_{k=1}^{1/\varepsilon} \frac{1}{k^2} + \frac{2}{(2\pi)^j} \sum_{k=1/\varepsilon}^{\infty} \frac{1}{k^2}$$

Using the inequality $\sum_{k=A}^{B} \frac{1}{k^2} \leq \frac{2}{A} - \frac{1}{B}$, we finally have

$$C_{j,\varepsilon} \leq 2(2-\varepsilon)\mathrm{e}^{-\varepsilon^2 j \log(2\pi)} + 4\varepsilon \mathrm{e}^{-j \log(2\pi)} \leq 4\mathrm{e}^{-j\varepsilon^2 \log(2\pi)}$$

which is surely verified when $\mathrm{e}^{-j\varepsilon^2 \log(2\pi)} > 2\mathrm{e}^{-j \log(2\pi)}$, namely $\varepsilon^2 < 1 - \log 2/\log(2\pi)$. □

Proof of Theorem 1. We now check the conditions $D_2(u_n)$ and $D'(u_n)$ using the previous estimate for the exponential decay of random correlations for characteristic functions. We will use a few results established on the afore mentioned paper [2]. We begin with $D_2(u_n)$ and we consider, with no loss of generality, the observable $g(x) = -\log(d(x, z))$, where

$d(\cdot,\cdot)$ is the distance on the circle and z is a given point on the circle. The event $U_n := \{g > u_n\} = \{X_0 > u_n\}$ will be the ball $B_{\mathrm{e}^{-u_n}}(z)$. We now introduce the observables

$$\phi(x) = \chi_{\{X_0>u_n\}} = \chi_{\{g>u_n\}},$$
$$\psi(x) = \int \chi_{\{g,\, g\circ f_{\varepsilon\xi_1},\, \dots,\, g\circ f^{\ell-1}_{\varepsilon\underline{\xi}} \le u_n\}}\, d\underline{\xi}^{l-1}.$$

It has been shown, in [2, Sect. 4.1], that proving condition $D_2(u_n)$ can be reduced to estimating the following correlation

$$\Big| \int \mu_\varepsilon\big(g > u_n, g\circ f^t_{\varepsilon\underline{\xi}} \le u_n, \dots, g\circ f^{t+\ell-1}_{\varepsilon\underline{\xi}} \le u_n\big)\, d\xi^{\mathbb{N}} - \int \mu_\varepsilon(g > u_n)\, d\xi^{\mathbb{N}} \int \mu_\varepsilon\big(g \le u_n, g\circ f_{\varepsilon\xi_1} \le u_n, \dots, g\circ f^{\ell-1}_{\varepsilon\underline{\xi}} \le u_n\big)\, d\xi^{\mathbb{N}} \Big|$$

Since all the maps $f_{\varepsilon\xi}$ are globally invertible and linear, the events $(g > u_n)$ and $(g\circ f^t_{\varepsilon\underline{\xi}} \le u_n, \dots, g\circ f^{t+\ell-1}_{\varepsilon\underline{\xi}} \le u_n)$ are a connected interval and a finite union of connected intervals, respectively. Therefore we can apply directly formula (4.3) with an exponential decay in t so that any sequence $t_n = n^\kappa$, with $0 < \kappa < 1$, will allow to verify the condition $n\gamma(n, t_n) \to 0$, when $n \to \infty$.
The computation of $D'(u_n)$ is a bit more lengthy. Let us fix a sequence $\alpha_n \to \infty$ so that $\alpha_n = o(k_n)$. Next, we introduce the quantity $R^{\underline{\xi}}(A)$ as the *first return of the set* A into itself under the realisation $\underline{\xi}$. As in [2, Sect. 4.1, just before eq. (4.5)], we have:

$$n \sum_{j=1}^{\lfloor n/k_n\rfloor} \mathbb{P}(U_n \cap f^{-j}_{\underline{\xi}}(U_n)) \le n \sum_{\alpha_n}^{\lfloor n/k_n\rfloor} \mathbb{P}(\{(x,\underline{\xi}) : x \in U_n,\, f^j_{\underline{\xi}}(x) \in U_n, R^{\underline{\xi}}(U_n) > \alpha_n\})$$

$$+ n \sum_{j=1}^{\lfloor n/k_n\rfloor} \mathbb{P}(\{(x,\underline{\xi}) : x \in U_n, R^{\underline{\xi}}(U_n) \le \alpha_n\}) \tag{4.6}$$

It can be shown that the measure of all the realisation such that $R^{\underline{\xi}}(U_n) \le \alpha_n$ is bounded by a constant times $\mathrm{Leb}(U_n)$ times α_n (we use here the fact that η in item (iv) above is 1 in our case). Therefore the second term in (4.6) is bounded by a constant times:

$$\frac{n^2}{k_n}\mathrm{Leb}(U_n)^2\alpha_n.$$

This reduces to $\tau^2\frac{\alpha_n}{k_n}$, whose limit is 0, as $n \to \infty$, given our choices for α_n and k_n. We now consider the first term on the right hand side of the inequality (4.6). We can obtain

the following result:

$$n\sum_{\alpha_n}^{\lfloor n/k_n\rfloor}\mathbb{P}(\{(x,\underline{\xi}):x\in U_n,\, f^j_{\underline{\xi}}(x)\in U_n, R^{\underline{\xi}}(U_n)>\alpha_n\})\le n\left\lfloor\frac{n}{k_n}\right\rfloor\mathrm{Leb}(U_n)^2+n\sum_{j=\alpha_n}^{\lfloor\frac{n}{k_n}\rfloor}|C_{j\varepsilon}| \tag{4.7}$$

where the correlation integral is computed with respect to the characteristic function of U_n. Next we use the decomposition (4.5) for $C_{j\varepsilon}$; we also observe that the product of the modulus of the k-Fourier coefficient of χ_{U_n} for itself gives

$$\frac{1}{2\pi^2k^2}(1-\cos 2\pi k\mathrm{Leb}(U_n)).$$

Using the notation $|U_n|=\mathrm{Leb}(U_n)$, we obtain (note that we discard the first term in (4.7) since it goes to zero because $n^2|U_n|^2\sim\tau^2$):

$$(4.7)\le n\sum_{j=\alpha_n}^{\lfloor\frac{n}{k_n}\rfloor}\Big\{2\sum_{k=1}^{1/\varepsilon}\frac{1}{2\pi^2k^2}(1-\cos 2\pi k|U_n|)e^{-k^2\varepsilon^2j\log(2\pi)} + 2\sum_{1/\varepsilon}^{\infty}\frac{1}{2\pi^2k^2}(1-\cos 2\pi k|U_n|)\frac{1}{(2\pi k\varepsilon)^j}\Big\}.$$

We consider the first term into the bracket: an upper bound can be defined as follows:

$$n|U_n|\sum_{j=\alpha_n}^{\lfloor\frac{n}{k_n}\rfloor}2e^{-\varepsilon^2j\log(2\pi)}\sum_{k=1}^{1/\varepsilon}\frac{1}{2\pi^2k^2}\left(\frac{1-\cos 2\pi k|U_n|}{|U_n|}\right),$$

When $n\to\infty$, $n|U_n|\sim\tau$ and the inner series goes to zero. The second piece in (4.7) could be bounded as before by

$$n|U_n|\sum_{j=\alpha_n}^{\lfloor\frac{n}{k_n}\rfloor}\frac{2}{(2\pi)^j}\sum_{1/\varepsilon}^{\infty}\frac{1}{2\pi^2k^2}\left(\frac{1-\cos 2\pi k|U_n|}{|U_n|}\right)$$

which converges to zero as well. □

4.2. **Numerical Investigation.** In this section we describe the results obtained from the numerical investigation of the stochastically perturbed rotation map following the procedure describe in Sect. 3. The results are displayed in Fig.1 where the green lines correspond to experiments where we have chosen a bin length $n=10^4$, whereas the blue lines refer to experiments where the chosen bin length is $n=10^3$. The red lines indicate the values of the parameters predicted by the theory for a 1-D map satisfying the mixing conditions described above. We set $\epsilon=10^{-p}$ to analyse the role of the perturbations on scales ranging from values smaller than those typical for the numerical noise to $\mathcal{O}(1)$ [11].
The solid lines display the values obtained by averaging over the 500 realisations of the

stochastic process, while the error bars indicate the standard deviation of the sample. Finally, with the dotted lines we indicated the experiments where less than 70% of the 500 realisations produce a statistics of extremes such that the empirical d.f. passes successfully the non-parametric Kolmogorov-Smirnov test [25] when compared to the best GEV fit.

Even though the theory guarantees the existence of EVL for the rotations perturbed with an arbitrarily weak noise (e.g. when compared with the numerical noise), the simulations clearly show that EVLs are obtained when considering small but finite noise amplitudes only when very long trajectories are considered. The quality of the fit improves when larger bins are considered (compare blue and green lines in Fig.1). This is in agreement with the idea that we should get EVL for infinitely small noises in the limit of infinitely long samples. In our case, EVLs are obtained only for $\epsilon > 10^{-4}$, which is still considerably larger than the noise introduced by round-off resulting from double precision, as the round-off procedure is equivalent to the addition to the exact map of a random noise of order 10^{-7} [20, 11].
This suggests that is relatively hard to get rid of the properties of the underlying deterministic dynamics just by adding some noise of unspecified strength and considering generically long time series: the emergence of the smoothing due to the stochastic perturbations is indeed non-trivial when considering very local properties of the invariant measure as we do here. It is interesting to expand the numerical investigation discussed here in order to find empirical laws connecting the strength of the noise perturbations with the length of the time series needed to observe EVLs.

5. Expanding Maps

In [2] it is proved that expanding maps on the circle admit EVLs when they are perturbed with additive noise (see, in the just quoted reference, Corollary 4.1 for smooth maps and Proposition 4.2 for discontinuous maps on I). The proof is completely different of the one we produced for rotations in Theorem 1. The proof there relies on the fact that expanding maps have exponential decay of correlations for BV-functions against L^1 observables. Moreover, as we already anticipated, the proof is not limited to continuous expanding maps, but it also holds for piecewise expanding maps with a finite number of discontinuities, provided that the map is topologically mixing. It is interesting to point out that under the assumptions (items (i)-(iv)) in the beginning of Subsection 4.1, we could prove convergence to the $e^{-\tau}$ law with the observables g_i previously introduced independently of the choice of the point z. We should remind that for deterministic systems and whenever z is a periodic point of prime period p, then the limit law is not $e^{-\tau}$ any more, but rather $e^{-\theta\tau}$, where $0 < \theta < 1$ is the *extremal index* discussed above. Therefore, obtaining experimentally an extremal index with unitary value for an observable $\varphi(\cdot)$ computed with respect to a periodic point z definitely points to the fact that the system is stochastically perturbed.

If a non-periodic point z is considered, it seems relevant to investigate whether stochastic and deterministic systems exhibit any sort of differences when looking in detail into the resulting EVLs. We consider the following map:

$$f_{\varepsilon\xi}(x) = 3x + \varepsilon\xi \bmod 1 \tag{5.1}$$

perturbed with additive noise as in the rotation case discussed above. The stationary measure for such a map is the uniform Lebesgue measure on the unit interval independently of the value of ε. The numerical setup is the same as in the case of the rotations: we begin by taking a non-periodic point as the center of our target set; the results are shown in picture 2. It is clear that the stochastic perturbations do not introduce any changes in the type of statistical behaviour observed for a non-periodic point z and no differences are encountered even if the number of observations in each bin is increased. This is compatible with the idea that the intrinsic chaoticity of the map overcomes the effect of the stochastic perturbations. Summarising, extremes do not help us in this case to distinguish the effect of intrinsic chaos and the effect of adding noise.

It seems more promising the problem of the form of the EVLs for periodic points. As discussed above, we cannot use the usual fitting procedure for the GEV, since it always renormalises in such a way that the extremal index seems to be one (see discussion above). Instead, in order to observe extremal indices different from one, we have to fit the series of minimum distances to the exponential distribution by normalising a priori the data. The inference of the extremal index θ (here already the extremal index) has been obtained via MLE. The normalisation applied consists only in multiplying the distances by the factor $2n$ as we deal with a constant density measure. In Fig. 3 we present the results for the extremal index obtained taking the periodic point $z = 1/2$ of prime period 1 for which $\theta = 2/3$ [6]. The results clearly show that we are able to recognise the perturbed dynamics as the extremal index goes to one when ϵ increases. Interestingly, the rupture with the values expected in the purely deterministic case is observed only for relatively intense noise. Moreover, when longer time series are considered (green experiment), the stochastic nature of the map becomes evident also for weaker perturbations. Finally, it is clear that the numerical noise (corresponding to $\epsilon \simeq 10^{-7}$) is definitely not sufficiently strong for biasing the statistics of the deterministic system.

6. Quadratic map

Following the analysis on the extremal index carried out in the previous section, we study numerically the quadratic map

$$f_{\varepsilon\xi}(x) = 1 - ax^2 + \varepsilon\xi$$

[6]We recall that $\theta = \theta(z) = 1 - |\det D(f^{-p}(z)|$, where z is a periodic point of prime period p, see [17, Theorem 3].

in the two cases $a = 0.014$ and $a = 0.314$. For these values of a, the deterministic dynamics is governed by the existence of an attracting fixed point (whose value depends on a); for both choices of a the fixed point is taken as z in the experiments. The stationary measure of such perturbed quadratic maps is absolutely continuous with respect to Lebesgue [3]. Extremes of this map have been already studied in the unperturbed case in [13], but for chaotic regime (a belonged to the Benedicks-Carleson set of parameters). Here, we proceed exactly as described in the previous section, by fitting the series of minimum distances to an exponential distribution after normalizing the data using the factor $2n$. From the theoretical point of view, the choice of this normalisation comes from assuming that, in the presence of strong noise, the measure becomes smooth in the neighbourhood of the periodic point and therefore it is locally not different with respect to the ternary shift already analysed. We believe that the EI computed with respect to this normalising constant seems to be the relevant one for the simulations in perturbed systems since it does not depend on the intensity of the noise but only on parameters that are known from the set up of the experiment.

The extremal index for the two cases is presented in the Figs 4 and 5 for $a = 0.014$ and $a = 0.314$, respectively. In both cases it is possible to see that θ grows exponentially and the noise increases as ϵ for $\epsilon \geq 10^{-6}$ while for very weak noise we obtain in both cases a result for the EI that depends on the length of the bin. This is probably due to numerical uncertainties intrinsic to operating in double precision. This effect can be clearly recognised by repeating the same experiment in single precision - results not shown here - where θ saturates for values of $\epsilon < 10^{-4}$ because the effect of the single precision numerical noise becomes dominant.

7. Hénon attractor

Finally, we investigate the impact of a stochastic perturbation applied onto a prototypical case of a map possessing a singular physical measure supported on an attractor whose dimension is smaller than the dimension of the phase space, d. For such dynamical systems in [29] we have shown numerically the convergence of maxima for the observables g_i to the three classical EVLs whose parameters (κ, σ, μ) depend on the scaling exponent of the measure of a ball centred around the point z from which the distance is computed. Such a scaling exponent turned out to be the local dimension $d_L(z)$ of the SRB-measure. For this reason, the equations for the asymptotic parameters presented in Eq. 3.5 for the one dimensional and absolutely continuous case, should be modified as follows:

$$\kappa(g_1) = 0 \quad \kappa(g_2) = 1/(ad_L(z)) \quad \kappa(g_3) = -1/(ad_L(z)) \quad \sigma(g_1) = 1/d_L(z). \tag{7.1}$$

In [29] we have also shown that the dimension of the whole attractor can be recovered by averaging over different points z the local dimensions $d_L(z)$ obtained inverting the relations in Eq. (7.1).

Hereby, we choose to investigate the properties of the Hénon map defined as:

$$\begin{aligned} x_{j+1} &= y_j + 1 - ax_j^2 + \epsilon\xi_j \\ y_{j+1} &= bx_j, \end{aligned} \tag{7.2}$$

where we consider the classical parameters $a = 1.4$, $b = 0.3$ [7]. Recall that for the usual Hénon attractor, points outside the basin of attraction escape to infinity. Hence, when the system is perturbed, it is natural to expect that, after waiting enough time, the evolution of every initial condition will escape to infinity since the noise lets the orbit explore the whole d-dimensional phase space. Moreover, as we have already pointed out in the introduction, since the physical measure is indeed smoothed by the action of the noise, so that the resulting attractor has the same dimension d of the phase space. Therefore, the EVLs' parameters depending on the local dimension $d_L(z)$ of the attractor should change value abruptly as soon as the noise is switched on. We have selected 500 different z on the support of the physical measure, and a single realisation has been analysed for each of these z. For the deterministic dynamics, such a set up has been used to reconstruct the attractor's dimension by averaging the local dimension $d_L(z)$ derived for each considered z from the EVLs [29]. Although there are no rigorous results on this system, we will such the numerical evidences and issues emerging from the analysis of Hénon to suggest how to operate in a general case where theoretical results are - for some reasons - unavailable.

In the case of purely deterministic dynamics, the EVLs parameters agree with high precision with the theoretical predictions of [29]. The numerical experiments performed with varying intensity of the noise forcing (see Fig. 6) suggest that, also in this case, the signature of the deterministic dynamics is pretty resilient. We need a rather intense noise to obtain a detectable smearing of the measure, such that our indicators see an absolutely continuous with respect to Lebesgue, invariant measure. Only when $p \simeq 0.1$ we observe the escaping behaviour to infinity which causes a divergence of the EVL from the deterministic one with unreliable parameters, whereas in all the other cases we recover quite well the dependence on the dimension of the usual Hénon attractor. Again, when dealing with finite data samples the correct reconstruction of the phase space and of the physical measure depends on the intensity of the noise. The case of Hénon also calls the attention for the fact that the balls of the perturbed systems keep scaling with the local dimension when the system is weakly perturbed, instead of scaling with the dimension of the phase space. This effect is explained in Collet's paper [7] and is linked to the direction in which the stochastic perturbation operates: we have already seen for the shift map that the addition of noise in chaotic one dimensional maps does not affect much the system's behaviour at

[7] We remind that Benedicks and Carleson [4] proved that there exists a set of positive Lebesgue measure S in the parameter space such that the Hénon map has a strange attractor whenever $a, b \in S$. The value of b is very small and the attractor lives in a small neighbourhood of the x-axis. For those values of a and b, one can prove the existence of the physical measure and of a stationary measure under additive noise, which is supported in the basin of attraction and that converges to the physical measure in the zero noise limit [5]. It is still unknown whether such results could be extended to the "historical" values that we consider here.

generic points. The effect of noise is negligible for the components parallel to the unstable manifold, while the effect is definitely stronger along the stable manifold. Introducing a noise acting as a forcing only along the stable manifold would indeed create a stronger smoothing effect. We will try to investigate this possibility in a future publication. The heuristic take-home message we learn from the case of the Hénon map is that the noise is not an easy fix for the singularity of the invariant measure in the case of dissipative systems, in the sense that its effect becomes noticeable only when its intensity is relatively larger, unless we consider extremely long time series.

As discussed in the introduction, this is a delicate point related to the practical applicability of the FDT in dissipative statistical mechanical systems. While it is clear that FDT does not apply for systems possessing a singular invariant measure [34, 35, 27, 30], some authors [1, 21] have advocated the practical applicability of the FDT thanks to the smoothing effect of noise. It is clear that, while even an arbitrarily small noise indeed smooths the invariant measure, one may need to accumulate an extraordinarily long statistics record in order to observe it, and, therefore, to be able to apply for real the FDT. Obviously, this is a matter worth investigating on its own.

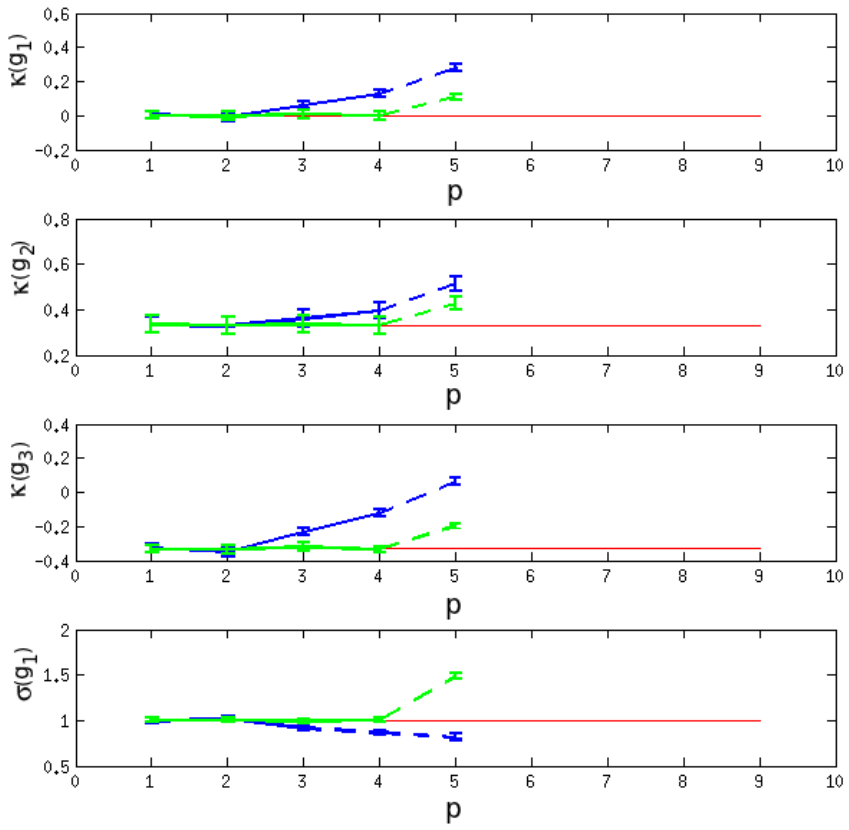


FIGURE 1. GEV parameters VS intensity of the noise $\epsilon = 10^{-p}$ for the circle rotations perturbed map. Blue: $n = 10^3$, $m = 10^3$, Green: $n = 10^4$, $m = 10^3$. Red lines: expected values. $z \simeq 0.7371$. From the top to the bottom: $\kappa(g_1), \kappa(g_2), \kappa(g_3), \sigma(g_1)$.

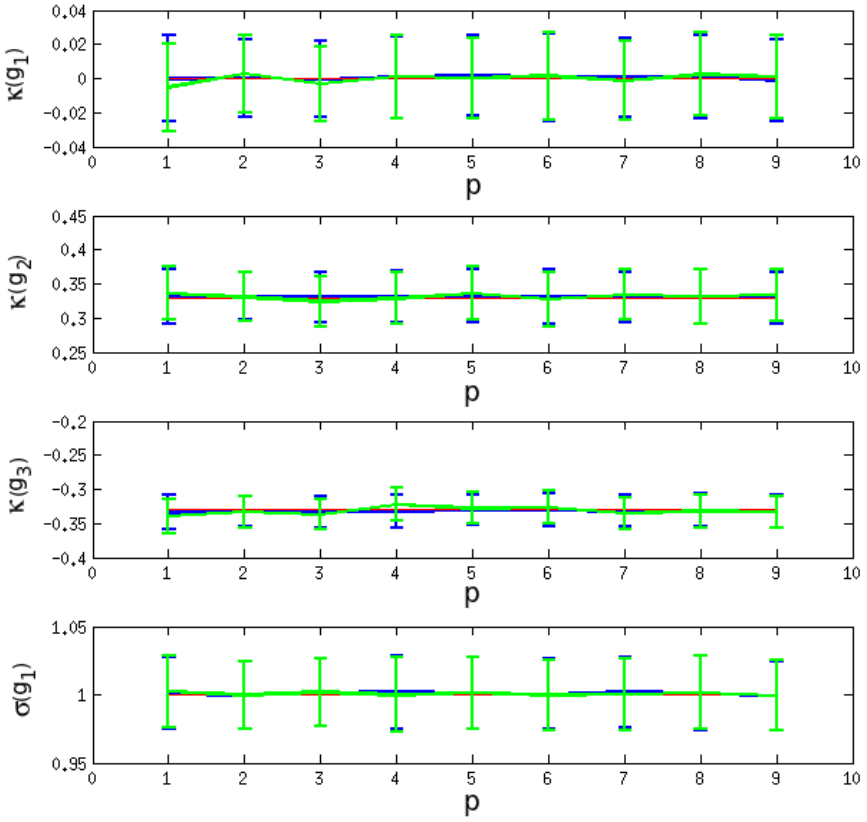


FIGURE 2. GEV parameters VS intensity of the noise $\epsilon = 10^{-p}$ for the ternary shift perturbed map. Blue: $n = 10^3$, $m = 10^3$, Green: $n = 10^4$, $m = 10^3$. Red lines: expected values. $z \simeq 0.7371$. From the top to the bottom: $\kappa(g_1), \kappa(g_2), \kappa(g_3), \sigma(g_1)$.

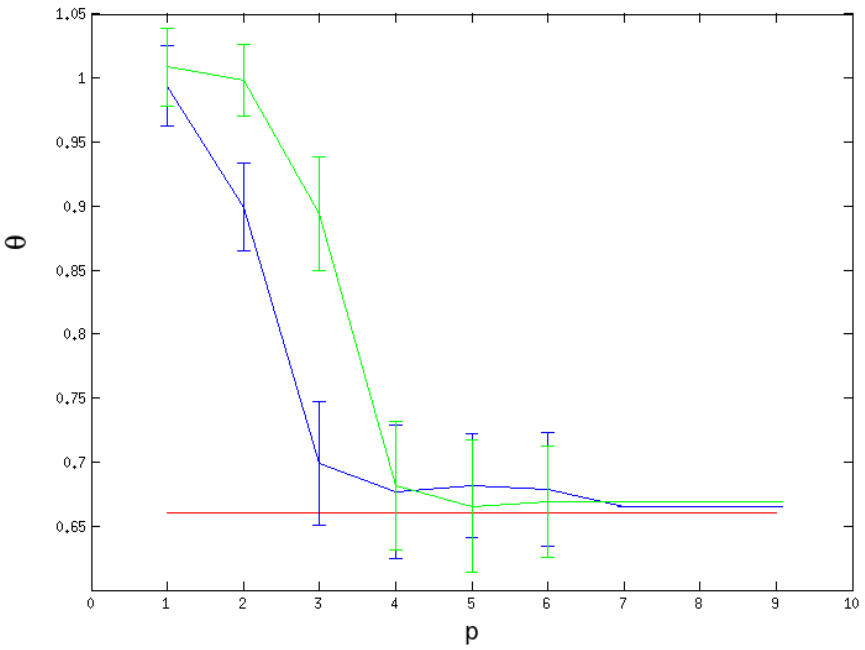


FIGURE 3. Extremal index θ VS intensity of the noise $\epsilon = 10^{-p}$ for the ternary shift perturbed map. Blue: $n = 10^3$, $m = 10^3$, Green: $n = 10^4$, $m = 10^3$. Red line: theoretical θ for $z = 0.5$ of the unperturbed map.

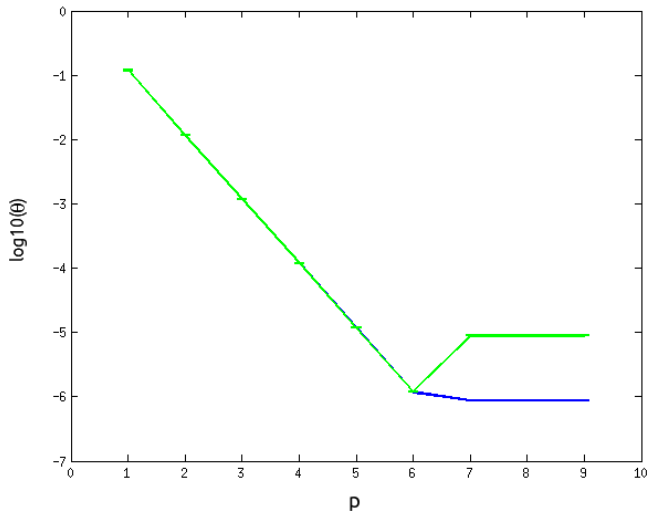


FIGURE 4. Extremal index θ (in log scale) VS intensity of the noise $\epsilon = 10^{-p}$ for the quadratic perturbed map with a=0.314, $z \simeq 0.72$. Blue: $n = 10^3$, $m = 10^3$, Green: $n = 10^4$, $m = 10^3$.

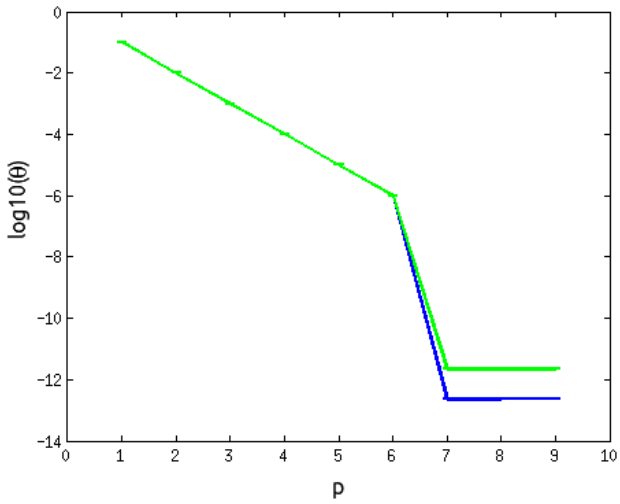


FIGURE 5. Extremal index θ (in log scale) VS intensity of the noise $\epsilon = 10^{-p}$ for the quadratic perturbed map with a=0.014, $z \simeq 0.98$. Blue: $n = 10^3$, $m = 10^3$, Green: $n = 10^4$, $m = 10^3$.

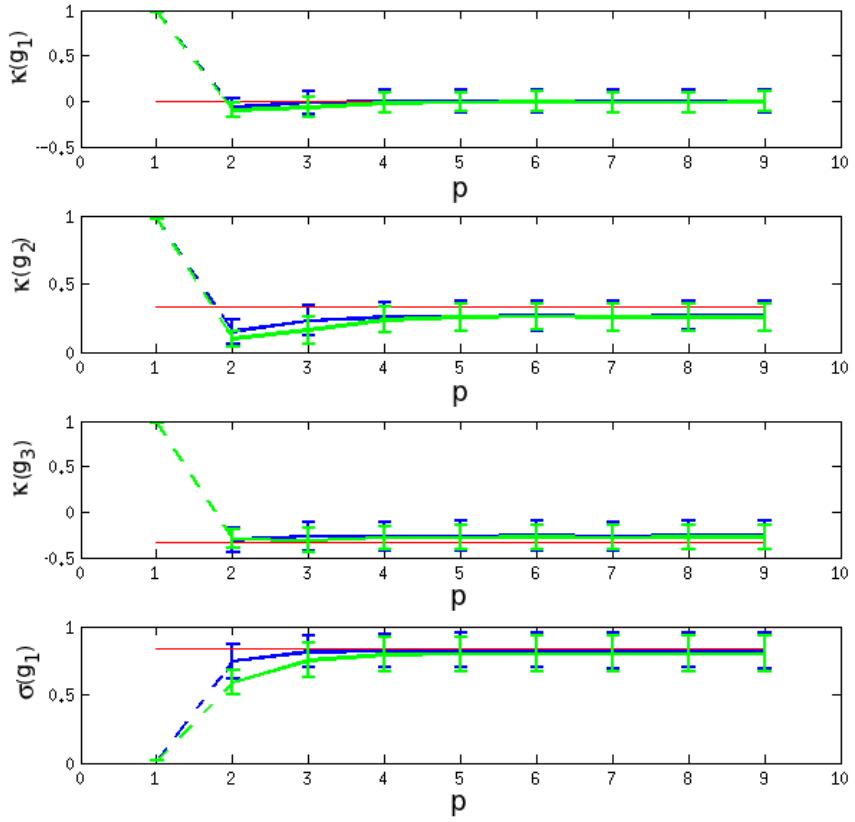


FIGURE 6. GEV parameters VS intensity of the noise $\epsilon = 10^{-p}$ for the Hénon perturbed map. Blue: $n = 10^3$, $m = 10^3$, Green: $n = 10^4$, $m = 10^3$. Red lines: expected values. z is different for each realisation. From the top to the bottom: $\kappa(g_1), \kappa(g_2), \kappa(g_3), \sigma(g_1)$.

REFERENCES

[1] R. Abramov and A. Majda. Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems. *Nonlinearity*, 20:2793, 2007.

[2] H. Aytaç, J. M. Freitas, and S. Vaienti. Laws of rare events for deterministic and random dynamical systems. Preprint arXiv:1207.5188, 2012.

[3] V. Baladi and M. Viana. Strong stochastic stability and rate of mixing for unimodal maps. *Ann. Sci. ' Ecole Norm. Sup*, 29:483–517, 1999.

[4] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Ann. of Math. (2)*, 133(1):73–169, 1991.

[5] M. Benedicks and M. Viana. Random perturbations and statistical properties of Hénon-like maps. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 23(5):713–752, 2006.

[6] D. Burn. The use of resampling for estimating confidence intervals for single site and pooled frequency analysis. *Hydrol Sci. J.*, 48(1):25–38, 2003.
[7] P. Collet. Stochastic perturbations of the invariant measure of some hyperbolic dynamical systems. In *Nonlinear Evolution and Chaotic Phenomena*. Plenum Press, New York, New York London, 1988.
[8] P. Collet. Statistics of closest return for some non-uniformly hyperbolic systems. *Ergodic Theory Dynam. Systems*, 21(2):401–420, 2001.
[9] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti. Numerical convergence of the block-maxima approach to the generalized extreme value distribution. *J. Stat. Phys.*, 145(5):1156–1180, 2011.
[10] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti. Generalized extreme value distribution parameters as dynamical indicators of stability. *to appear in Int. J. of Bifurcat. Chaos*, 2012.
[11] D. Faranda, M. Mestre, and G. Turchetti. Analysis of round off errors with reversibility test as a dynamical indicator. *to appear in Int. J. of Bifurcat. Chaos, Special Issue: NCD 2010, Volume 2*, 22(9), 2012.
[12] A. Ferguson and M. Pollicott. Escape rates for gibbs measures. *Ergod. Theory Dynam. Systems*, 32:961–988, 2012.
[13] A. C. M. Freitas and J. M. Freitas. Extreme values for Benedicks-Carleson quadratic maps. *Ergodic Theory Dynam. Systems*, 28(4):1117–1133, 2008.
[14] 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):1088–1093, 2008.
[15] A. C. M. Freitas, J. M. Freitas, and M. Todd. Hitting time statistics and extreme value theory. *Probab. Theory Related Fields*, 147(3):675–710, 07 2010.
[16] A. C. M. Freitas, J. M. Freitas, and M. Todd. The compound poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics. Preprint arXiv:1204.2304, 2012.
[17] A. C. M. Freitas, J. M. Freitas, and M. Todd. Extremal index, hitting time statistics and periodicity. To appear in Adv. Math., 2012.
[18] M. Hirata. Poisson law for Axiom A diffeomorphisms. *Ergodic Theory Dynam. Systems*, 13(3):533–556, 1993.
[19] G. Keller. Rare events, exponential hitting times and extremal indices via spectral perturbationâ. *Dynamical Systems*, 27(1):11–27, 2012.
[20] D. E. Knuth. *The art of computer programming. Volume 2: Seminumerical Algorithms*. Addison-Wesley, Reading, 1973.
[21] G. Lacorata and A. Vulpiani. Fluctuation-response relation and modeling in systems with fast and slow dynamics. *Nonlinear Proc. Geoph.*, 14(5):681–694, 2007.
[22] M. R. Leadbetter. On extreme values in stationary sequences. *Z. Wahrsch. Verw. Gebiete*, 28:289–303, 1973/74.
[23] M. R. Leadbetter. Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete*, 65(2):291–306, 1983.
[24] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York, 1983.
[25] H. Lilliefors. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318):399–402, 1967.
[26] R. M. Loynes. Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.*, 36:993–999, 1965.
[27] V. Lucarini. Response theory for equilibrium and non-equilibrium statistical mechanics: causality and generalized kramers-kronig relations. *J. Stat. Phys.*, 131(3):543–558, 2008.
[28] V. Lucarini. Stochastic perturbations to dynamical systems: a response theory approach. *J. Stat. Phys.*, pages 1–13, 2012.
[29] V. Lucarini, D. Faranda, G. Turchetti, and S. Vaienti. Extreme value theory for singular measures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(2):023135, 2012.

[30] V. Lucarini and S. Sarno. A statistical mechanical approach for the computation of the climatic response to general forcings. *Nonlinear Proc. Geoph.*, 18(1):7–28, 2011.
[31] T. Palmer and P. Williams. *Stochastic physics and climate modelling*. Cambridge University Press, 2010.
[32] R. Royall. Model robust confidence intervals using maximum likelihood estimators. *International Statistical Review/Revue Internationale de Statistique*, pages 221–226, 1986.
[33] D. Ruelle. Differentiation of srb states. *Comm. Math. Phys.*, 187(1):227–241, 1997.
[34] D. Ruelle. General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium. *Phys. Lett. A*, 245(3-4):220–224, 1998.
[35] D. Ruelle. A review of linear response theory for general differentiable dynamical systems. *Nonlinearity*, 22:855, 2009.

Davide Faranda, Klimacampus, Institute of Meteorology, University of Hamburg, Grindelberg 5, 20144, Hamburg, Germany

E-mail address: davide.faranda@zmaw.de

URL: http://www.mi.uni-hamburg.de/Davide-Faran.6901.0.html?&L=3

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

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

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

Valerio Lucarini, Klimacampus, Institute of Meteorology, University of Hamburg, Grindelberg 5, 20144, Hamburg, Germany. – Department of Mathematics and Statistics, University of Reading, Reading, UK

E-mail address: valerio.lucarini@zmaw.de

URL: http://www.mi.uni-hamburg.de/index.php?id=6870&L=3

Giorgio Turchetti, Department of Physics, University of Bologna.INFN-Bologna, Via Irnerio 46, Bologna, Italy

E-mail address: turchetti@bo.infn.it

URL: http://www.unibo.it/SitoWebDocente/default.htm?UPN=giorgio.turchetti40unibo.it

Sandro Vaienti, (Temporary address) Centro de Modelamiento Matemtico, Av. Blanco Encalada 2120 Piso 7, Santiago de Chile. (On leave from/Permanent address) UMR-7332 Centre de Physique Théorique, CNRS, Université d'Aix-Marseille I, II, Université du Sud, Toulon-Var and FRUMAM, Fédération de Recherche des Unités des Mathématiques de Marseille, CPT Luminy, Case 907, F-13288 Marseille CEDEX 9

E-mail address: vaienti@cpt.univ-mrs.fr

Chapter 8

Article: Bistable systems with stochastic noise: virtues and limits of effective one-dimensional Langevin equations

Nonlin. Processes Geophys., 19, 9–22, 2012
www.nonlin-processes-geophys.net/19/9/2012/
doi:10.5194/npg-19-9-2012

Nonlinear Processes in Geophysics

Bistable systems with stochastic noise: virtues and limits of effective one-dimensional Langevin equations

V. Lucarini[1,2], **D. Faranda**[1], **and M. Willeit**[3]

[1]Klimacampus, Institute of Meteorology, Hamburg Universität, Hamburg, Germany
[2]Department of Mathematics and Statistics, University of Reading, Reading, UK
[3]Potsdam Institute for Climate Impact Research, Potsdam, Germany

Correspondence to: V. Lucarini (valerio.lucarini@zmaw.de)

Received: 5 September 2011 – Revised: 4 December 2011 – Accepted: 8 December 2011 – Published: 3 January 2012

Abstract. The understanding of the statistical properties and of the dynamics of multistable systems is gaining more and more importance in a vast variety of scientific fields. This is especially relevant for the investigation of the tipping points of complex systems. Sometimes, in order to understand the time series of given observables exhibiting bimodal distributions, simple one-dimensional Langevin models are fitted to reproduce the observed statistical properties, and used to investing-ate the projected dynamics of the observable. This is of great relevance for studying potential catastrophic changes in the properties of the underlying system or resonant behaviours like those related to stochastic resonance-like mechanisms. In this paper, we propose a framework for encasing this kind of studies, using simple box models of the oceanic circulation and choosing as observable the strength of the thermohaline circulation. We study the statistical properties of the transitions between the two modes of operation of the thermohaline circulation under symmetric boundary forcings and test their agreement with simplified one-dimensional phenomenological theories. We extend our analysis to include stochastic resonance-like amplification processes. We conclude that fitted one-dimensional Langevin models, when closely scrutinised, may result to be more ad-hoc than they seem, lacking robustness and/or well-posedness. They should be treated with care, more as an empiric descriptive tool than as methodology with predictive power.

1 Introduction

An interesting property of many physical systems with several degrees of freedom is the presence of multiple equilibria (or, more in general, of a disconnected attractor) for a given choice of the parameters. In such a case, the system does not obey ergodicity and its asymptotic state depends on what is the basin of attraction the initial condition belongs to. Among the many interesting properties of multi-stable systems, we may mention their possibility of featuring hysteretic behaviour: starting from an initial equilibrium $x = x_{\rm in}$ realized for a given value of a parameter $P = P_{\rm in}$ and increasing adiabatically the value of P so that the system is always at equilibrium following $x = x'(P)$, we may eventually encounter bifurcations leading the system to a new branch of equilibria $x = x'(P)$ such that, if we revert the direction of variation of P, we may end up to a different final stable state $x_{\rm fin} = x'(P_{\rm in}) \neq x(P_{\rm in}) = x_{\rm in}$. More generally, we can say that the history of the system determines which of the stable states is realized for a given choice of the parameters.

Whereas hysteretic behaviour has first been discussed in the context of magnetism, climate dynamics offers some outstanding examples where multistability is of great relevance, such as the classical problem of the snowball/snowfree Earth (Saltzman 2002; Lucarini et al., 2010; Pierrehumbert et al., 2011). In this context, the problem which has probably attracted the greatest deal of interest in the last two decades is that of the stability properties of the thermohaline circulation (THC). Since paleoclimatic evidences suggest that the large scale circulation of the Atlantic Ocean presents at least two, qualitatively different, stable modes of operation (Boyle et al., 1987; Rahmstorf, 2002), theoretical and modellistic efforts have long been directed to understanding the mathematical properties of the circulation and the physical processes responsible for switching from one to the other stable mode and those responsible for ensuring the stability of either equilibrium.

Interestingly, it has been possible to construct very simple models of the THC (Stommel, 1961; Rooth, 1982) able to feature most of the desired properties, and models of higher degrees of complexity have basically confirmed the robustness of such properties of multistability, from simple

two-dimensional convective equations models (Cessi and Young, 1992; Vellinga, 1996; Lucarini et al., 2005, 2007) to simplified climate models (Stocker and Wright, 1991; Rahmstorf, 1995; Stocker and Schmittner, 1997). Whereas climate models of intermediate complexity now consistently represent the THC as a multistable system (Rahmstorf et al., 2005), results are not conclusive when full 3-D climate models are considered (Stouffer and Manabe, 2003; Scott et al., 2008). Nonetheless, recent simulations performed by Hawkins et al. (2011) with a full 3-D climate model have successfully reproduced the kind of bistability properties shown by Rahmstorf et al. (2005). In the case of THC, the strength of the hydrological cycle plays the role of dominant parameter, whose variation can lead the system through bifurcations (Sijp and England, 2006, 2011; Sijp et al., 2011). A detailed account of these analyses can be found in Rahmstorf (1995), Scott et al. (1999), and Titz et al. (2001, 2002). The matter is of great relevance for understanding climate variability and climate change because, if the system is close to a bifurcation point, small changes in the parameters value could have virtually irreversible effects, driving the climate system a qualitatively different steady state. As evidenced in many studies (see, e.g. Kuhlbrodt et al., 2007), a transition from the present state of the THC to a state featuring weaker meridional circulation would have very relevant climatic effects at regional and global scale, as the northward ocean heat transport in the Atlantic would be greatly reduced.

The potential of shut-off of the THC is considered a high impact climate risk – even if its likelihood for the present climate is considered very low (Rahmstorf, 2006) – and the conditions under which such a transition can occur are probably the best example of a "climate tipping point" (Lenton et al., 2008).

Since we are dealing in principle with a 3-D fluid with complex thermodynamical and dynamical properties, a lot of efforts have been directing at finding, using suitable scaling and simplified theoretical setting, an approximate one-dimensional ordinary differential equation equation of the form $\dot{q} = F(q, P)$, where q is the intensity of the THC and P is a set of parameters of the system. Such an equation would be able to represent at least in a semi-quantitative way the evolution of the THC strength as a function of the strength and some parameters only and, by solving $F(q, P) = 0$, would provide the (in general) multiple equilibria corresponding to a specific choice of the set of parameters P. An excellent account of this methodology can be found in Dijkstra (2005). Note that, very recently, a related surrogate one-dimensional dynamics for a salinity indicator has been proposed to fit the output of a comprehensive climate model (Sijp et al., 2011).

The dynamics of multistable systems becomes rather interesting when stochastic forcing is considered. In the most basic case, such a forcing is represented in the form of additive white noise. Noise introduces on one side small scale variability around each of the stable equilibria of the system, and, on the other side, allows for jumps (large scale variability) driving the system across the boundaries separating the basin of attraction of the fixed points. See Friedlin and Wentzel (1998) for a detailed mathematical treatment of these problems. Since the landmark Hasselmann's (1976) contribution, it has become clearer and clearer in the climate science community that stochastic forcing components can be treated as quite reliable surrogates for high frequency processes not captured by the variables included in the climate model under consideration (Fraedrich, 1978; Saltzman, 2002). This has raised the interest in exploring whether transitions between stable modes of operation of the THC far from the actual tipping points could be triggered by noise, representing high-frequency (with respect to the ocean's time scale) atmospheric forcings, of sufficient amplitude (Cessi, 1994; Monahan, 2002). Along these lines, it has become especially tempting to interpret the dynamics of the THC in presence of noise as resulting from an effective Langevin equation of the form $dq = F(q, p)dt + \varepsilon dW$ – where dW is the increment of a Wiener process – as this opens the way to approaching the problem in terms of one-dimensional Fokker-Planck equation. Ditlevsen (1999) suggested the possibility of considering more general stochastic processes for accommodating the statistical properties of observational data.

The Langevin equation approach has also led various researchers to study whether the process of stochastic resonance (Gammaitoni et al., 1998) – basically noise-enhanced response amplification to periodic forcing – could explain the strength of the climate response (in terms of actual THC strength) in spite of the relative weakness of the Milankovic forcing. Note that stochastic resonance, which has enjoyed great success in fields ranging from microscopic physics to neurobiology and perception, was first proposed in a climatic context (Benzi et al., 1982; Nicolis, 1982; Benzi, 2010). Velez-Belchi et al. (2001) provided the first example of a simple THC box-model featuring stochastic resonance, and later Ganopolski and Rahmstorf (2002) observed a similar mechanism in action in a much more realistic climate model.

In this work we would like to examine critically the effectiveness and robustness of using one-dimensional Langevin equations to represent the dynamical and statistical properties of the THC strength resulting from models which feature more than one degree of freedom. This is methodologically relevant, in the context of recent efforts directed at understanding whether the transitions between different steady states associated to the tipping points can be highlighted by early indicators (Scheffer et al., 2008). In Sect. 2 we show how to construct, from data, the form of the effective driving force and the intensity of noise, and propose tests for investigating the robustness of the approach. In Sect. 3 we describe the main properties of the simple box models of the oceanic circulation analysed in this paper. In Sect. 4 we test our methodology by studying, in various cases, whether it is possible to represent consistently the statistical properties

of the THC strength resulting from the stochastic forcing of the models introduced in Sect. 3 using an one-dimensional Langevin model. In Sect. 5 we further expand our analysis by testing whether the matching conditions for observing stochastic resonance are obeyed. In Sect. 6 we present our conclusions.

2 Theoretical background

2.1 Modelling a bistable system

Let's consider a one dimensional Langevin equation for the variable x of the form:

$$dx = F(x)dt + \varepsilon dW, \tag{1}$$

where $F(x)$ is a smooth function of x giving the drift term, W is a standard Wiener process and dW is its infinitesimal increment, so that ε parameterises the strength of the stochastic forcing. As well known, the invariant probability density function (pdf) $\pi(x)$ can be written as:

$$\pi(x) = Ce^{-2\frac{V(x)}{\varepsilon^2}} \tag{2}$$

where $V(x)$ is the effective potential such that $dV(x)/dx = -F(x)$ and C is the normalisation constant. The local extrema of the potential correspond to the fixed points of the deterministic system obtained when $\varepsilon = 0$, and in particular its local minima (maxima) giving the stable (unstable) equilibria. Quite intuitively, in the stochastic case, the peaks of the invariant probability distribution correspond to the minima of the potential.

In the prototypical situation of a confining double well potential, where we refer to the position of the right and left minimum of $V(x)$ as x_+, x$_-$, and to local maximum as x_0, the two peaks of the $\pi(x)$ are separated by a dip corresponding to the local maximum of the potential, while for large positive and negative values of x the $\pi(x)$ approaches zero as $V(x)$ diverges for $|x| \to \infty$. The average rate of transition $r(+ \to -)$ from the basin of attraction of x_+ to that of x_- can be approximated using the Kramer's formula:

$$r(+ \to -) = \frac{1}{2\pi}\sqrt{V''(x_+)V''(x_0)}e^{-2\frac{V(x_0)-V(x_+)}{\varepsilon^2}} \tag{3}$$

under the condition that the absolute value of the exponent is larger than one, so that $V(x_0) - V(x_+) \geq \varepsilon^2/2$. This corresponds to the physical condition that the noise is moderate with respect to the depth of the potential well. The average rate of transition $r(- \to +)$ in the opposite direction can be obtained by exchanging the sign plus with the sign minus in the previous expression. The Kramer's formula basically expresses the general fact that at stationary state a detailed balance conditions applies.

Let's now consider the less idealised case where we observe a scalar output signal $y(\boldsymbol{x})$ generated by a stochastic or chaotic deterministic flow of the system $\boldsymbol{x}$ living, in general, in an N-dimensional phase space, and let's assume that the empirical pdf $\pi(y)$ is bimodal, so that two peaks are found at $y = y_+$, $y = y_-$, separated by a local minimum at $y = y_0$. We wish to test the possibility of constructing a Langevin equation for the scalar y:

$$dy = F_{\text{eff}}(y)dt + \varepsilon_{\text{eff}}dW \tag{4}$$

so that the statistics generated by Eq. (4) closely resembles the one deriving from the full system. Such a representation would bypass the details of the full dynamics of the system and its construction can be approached by imposing constraints based upon the populations of the basin of attractions of the two modes and upon the transition probability between such basins. Basically, this amounts to defining an effective projected dynamics.

The observation time of the variable y must be long enough to allow for a robust estimate of the pdf and for observing many transitions between the two modes. From the empirical pdf of the considered observable $\pi(y)$ we derive the function $(y) = -\ln(\pi(y))$. In order to achieve compatibility with Eq. (2), we define

$$2V_{\text{eff}}(y)/\varepsilon^2_{\text{eff}} = U(y) + \text{const.} \tag{5}$$

The function $U(y)$ contains information on both the effective potential of the system and on the effective intensity of the noise. Substituting the expression of $V_{\text{eff}}(y)$ in Eq. (3), we can write, e.g. the average rate $r(+ \to -)$ as follows:

$$r(+ \to -) \approx \frac{\varepsilon^2_{\text{eff}}}{4\pi}\sqrt{U''(y_+)U''(y_0)}e^{-(U(y_0)-U(y_+))} \tag{6}$$

By comparing this expression with the observed transition rate, we can finally find the actual value of ε_{eff}, because all the other terms can be computed from the time series of the y observable. Assuming that in the region between the two minima and the local maximum the function $U(y)$ is smooth, we obtain the following expression, which is numerically more robust as the second derivatives disappear:

$$\begin{aligned} r(+ \to -) &\approx \frac{\sqrt{2}\varepsilon^2_{\text{eff}}}{\pi}\frac{(U(y_0)-U(y_+))}{(y_+-y_0)^2}e^{-(U(y_0)-U(y_+))} \\ &= \frac{2\sqrt{2}}{\pi}\frac{(V(y_0)-V(y_+))}{(y_+-y_0)^2}e^{-\left(2\frac{V(y_0)-V(y_+)}{\varepsilon^2_{\text{eff}}}\right)} \end{aligned} \tag{7}$$

where $G(+ \to -) = \frac{\sqrt{2}}{\pi}\frac{(U(y_0)-U(y_+))}{(y_+-y_0)^2}e^{-(U(y_0)-U(y_+))}$ is a factor depending only on the observed probability. Note that Eqs. (6)–(7) are valid under the condition that $U(y_0) - U(y_+) \gtrsim 1$ or, equivalently, that $\pi(y_+)/\pi(y_0) \gtrsim e$. By plugging the obtained value of ε_{eff} into Eq. (5) we derive $V_{\text{eff}}(y)$. We can then reconstruct the effective Langevin equation in the form given in Eq. (4). It is crucial that the same value for

ε_{eff} is obtained when using as benchmark the average rate $r(-\rightarrow +)$, because otherwise we have that the reconstructed dynamics does not obey detailed balance, or, equivalently, there is a mismatch between time-scales of the transitions and steady state populations of the two basins. Such an issue is not relevant in the special case when the pdf of the system is symmetric with respect to y_0.

2.2 Robustness

Let's now consider an N-dimensional Langevin equation of the form:

$$dx_i = F_i(x_1,...,x_N)dt + \varepsilon_{ij}dW_j \tag{8}$$

where the dw_j terms indicate increments of independent Wiener processes and the $F_i(x_1,...,x_N)$ are the (generally nonlinear) drift terms. We can write the Langevin equation for the observable y constructed as linear combination of the system variables $y = c_i x_i$ as follows:

$$\begin{aligned} dy &= c_i F_i(x_1,...,x_N)dr + c_i\varepsilon_{ij}dW_j \\ &= c_i F_i(x_1,...,x_N)dt + \bar{\varepsilon}dW \end{aligned} \tag{9}$$

where, exploiting the independence of the Wiener processes, we have $\bar{\varepsilon} = \sqrt{\Sigma_{j=1}^{N}(c_i\varepsilon_{ij})^2}$. If the deterministic part of system (9) features two stable equilibria $\boldsymbol{x}_+$ and $\boldsymbol{x}_-$, when stochastic noise is added, we will see hopping between the basins of attraction of these two points. When looking at the variable $y = c_i x_i$ as output of the system, we will see a bimodal distribution where the two peaks are centred at $y_+ = c_i x_{i,+}$ and $y_- = c_i x_{i,-}$, respectively, with a local minimum in-between situated at y_0.

The true dynamics of the observable y is indeed given in Eq. (9), but, if we are provided only with the time series of y, the best we can do in to try to derive the "best" approximate equation – the one-dimensional Langevin model – of the form given in Eq. (4), using the heuristic procedure described above. Comparing Eqs. (4) and (9), we understand that the difference in the drift terms $c_i F_i(x_1,...,x_N) - F_{\text{eff}}(y)$ describes the deterministic dynamics of the y variable which cannot be parameterised in terms of the y variable alone. We may expect that if y is a slow variable, which retains the long term memory of the system, such a difference is small, and the dynamics of y is truly quasi-one-dimensional In there is time-scale separation, one may expect that the impact of the faster variables on its evolution can be effectively expressed as noise, thus determining the value of $\bar{\varepsilon}$ (Saltzman, 2002). Therefore, the ratio $\bar{\varepsilon}/\varepsilon_{\text{eff}}$ gives an indirect measure of how strong is this effect, being close to 1 if the dominant contribution for comes from the direct stochastic forcing into the system.

Since constructing an ad-hoc one-dimensional model from a time-series is typically possible by following the lines described above, we need to introduce some criteria to test the robustness of the approach we have undertaken. This is needed in order to check whether our methodology is solely descriptive of the statistics of the chosen observable system for a given choice of parameters, or, instead, has predictive power, in terms of allowing one to understand how the statistical properties of the observable change when the parameters of the full system are altered. Obviously, as mentioned above, choosing a suitable variable y will be crucial in ensuring the effectiveness of this one-dimensional parameterisation:

- As a first condition of robustness, we may ask that if we multiply the noise intensity of the true system by a factor α, so that $\varepsilon_{ij} \rightarrow \alpha\varepsilon_{ij}$ in Eq. (8), we obtain that correspondingly $\varepsilon_{\text{eff}} \rightarrow \alpha\varepsilon_{\text{eff}}$ and the effective potential $V_{\text{eff}}(y)$ is not altered in Eq. (4).
- Another condition of robustness is that if in Eq. (9) we alter the noise matrix ε_{ij} in such a way that $\bar{\varepsilon}$ is not altered in Eq. (9), we would like that, correspondingly, ε_{eff} and V_{eff} are not altered in Eq. (4).

These conditions basically require that the reconstructed deterministic drift term is independent of the intensity of the noise – so that an underlying deterministic dynamics is well defined – and the reconstructed noise intensity scales linearly with the actual noise applied to the system, so that we can construct a relationship such as $\varepsilon_{\text{eff}}(\bar{\varepsilon}) \approx \gamma\bar{\varepsilon}$

3 Simple box models of the thermohaline circulation

3.1 Full model

We consider the simple deterministic three-box model of the deep circulation of the Atlantic Ocean introduced by Rooth (1982) and thoroughly discussed in Scott et al. (1999) and Lucarini and Stone (2005a, b). The model consists of a northern high-latitude box (box 1), a tropical box (box 2), and a southern high-latitude box (box 3). The volume of the two high-latitude boxes is the same and is $1/V$ times the volume of the tropical box, where V is chosen to be equal to 2. The physical state of the box i is described by its temperature T_i and its salinity S_i; the box i receives from the atmosphere the net flux of heat H_i and the net flux of freshwater F_i; the freshwater fluxes globally sum up to 0, so that the average oceanic salinity is a conserved quantity of the system. The box i is subjected to the oceanic advection of heat and salt from the upstream box through the THC, whose strength is q. The dynamics of the system is described by the evolution equation for the temperature and the salinity of each box. After a suitable procedure of non-dimensionalisation (Lucarini and Stone, 2005a), we obtain the following final form for the temperature and salinity tendency equations for the three boxes:

$$\dot{T}_1 = \begin{cases} q(T_2 - T_1) + H_1 \\ |q|(T_3 - T_1) + H_1 \end{cases} \tag{10a}$$

$$\dot{T}_2 = \begin{cases} \frac{q}{V}(T_3 - T_2) + H_2 \\ |\frac{q}{V}|(T_1 - T_2) + H_2 \end{cases} \quad (10b)$$

$$\dot{T}_3 = \begin{cases} q(T_1 - T_3) + H_3, & q > 0 \\ |q|(T_2 - T_3) + H_3, & q \leq 0 \end{cases} \quad (10c)$$

$$\dot{S}_1 = \begin{cases} q(S_2 - S_1) - F_1 \\ |q|(S_3 - S_1) - F_1 \end{cases} \quad (10d)$$

$$\dot{S}_2 = -\frac{1}{V}(\dot{S}_1 + \dot{S}_3) \quad (10e)$$

$$\dot{S}_3 = \begin{cases} q(S_1 - S_3) - F_3, & q > 0 \\ |q|(S_2 - S_3) - F_3, & q \leq 0 \end{cases} \quad (10f)$$

where $q = k(\rho_1 - \rho_3)$ and $\rho_i = \rho_0(1 - \alpha T_i + \beta S_i)$, so that $q = k(\beta(S_1 - S_3) - \alpha(T_1 - T_3))$, α and β are the usual thermal and haline expansion coefficients, ρ_0 is a baseline density, and k is the hydraulic constant controlling the water transport. Such a parameterisation was first introduced by Stommel (1961) for a hemispheric box model, whereas the approximate linear relationship between the density difference between the two high-latitude regions and the THC strength has been confirmed by simplified yet realistic GCM simulations (Rahmstorf, 1995; Scott et al., 2008).

The freshwater fluxes F_i are considered given constants, with $F_2 = -(F_1 + F_3)/V$, so that $S_1 + V S_2 + S3 = (V+2)S_0$ at all times, where $S_0 = 35$ psu is a baseline salinity. Instead, the heat flux $H_i = \lambda(\bar{T}_i - T_i)$ is such that the box temperature is relaxed to a fixed target temperature $\bar{T}_i$ with the time constant λ^{-1}. Such a representation mimics the combined effect of radiative heat transfer and of a meridional heat transport. This also implies that the spatial average of ocean temperature $\widetilde{T} = (T_1 + V T_2 + T_3)/(2 + V)$ obeys the evolution equation $\dot{\widetilde{T}} = \lambda(\widetilde{\bar{T}} - \widetilde{T})$, where $\widetilde{\bar{T}}$ is the spatial average of the target temperature, so that asymptotically (and, in practise, after few units of λ^{-1})$\widetilde{T}$ is a conserved quantity. Therefore, we practically have $\dot{T}_2 \approx -(\dot{T}_1 + \dot{T}_3)/V$. We usually have that the internal time scale of the system q^{-1} is much larger than λ^{-1}, thus implying that the thermal relaxation is fast.

The water of the high latitude ocean box where downwelling occurs is warmer and more saline that than situated on the opposite side of the planet, since it receives advection from the warm and saline equatorial box. Since the haline contribution is stronger, the downwelling box is denser than the upwelling box. The sign of q is positive if downwelling occurs in box 1, and negative if it occurs in box 3. The buoyancy fluxes in the upwelling and downwelling boxes serve different purposes in determining the dynamics of the system. In fact, the strength of the circulation at equilibrium q_{ref} depends only on the strength of the buoyancy fluxes in the upwelling box $H_{\mathrm{u,eq}}$ and F_{u}:

$$|q_{\mathrm{ref}}| = \sqrt{k(\alpha H_{\mathrm{u,eq}} + \beta F_{\mathrm{u}})},$$

where the sign of q_{eq} is positive if $u = 1$ and negative if $u = 1$. Instead, for a given value of F_{u} the realised pattern of circulation is stable as long as $F_{\mathrm{d}} \lesssim 3F_{\mathrm{u}}$, which implies that $F_{\mathrm{d}} = 3F_{\mathrm{u}}$ a bifurcation leading to an instability of the system is found. Such an instability exchanges the role of the upwelling and downwelling boxes. Therefore, if, e.g. in the initial state $u = 3$ and $d = 1$, and F_3 is kept fixed, the system features bistability for the following range of values of $F_1 : 1/3F_3 \lesssim F_1 \lesssim 3F_3$. These approximate relations become exact in the limit of infinitely fast thermal relaxation.

Following Scott et al. (1999) and Lucarini and Stone (2005a), we select for the constants of the system the values $k = 1.5 \times 10^{-6}\,\mathrm{s}^{-1}$, $\alpha = 1.5 \times 10^{-4}$ K^{-1}, $\beta = 8.0 \times 10^{-4}\,(\mathrm{psu})^{-1}$, $\lambda = 1.3 \times 10^{-9}\,\mathrm{s}^{-1}$, $V = 2$. When symmetric boundary conditions are considered with $\bar{T}_1 = \bar{T}_3 = 0\,^{\circ}\mathrm{C}$, $\bar{T}_3 = 30\,^{\circ}\mathrm{C}$, $F_1 = F_3 = 9 \times 10^{-11}\,\mathrm{psu\,s}^{-1}$, we obtain at steady state $|q_{\mathrm{ref}}| = 1.47 \times 10^{-11}\,\mathrm{s}^{-1}$. The sign of q depends uniquely on the initial conditions of the integration: we have 50 % probability of ending up in either the northern or the southern downwelling state if random initial conditions are chosen. Since the internal time scale $|q_{\mathrm{ref}}|^{-1} \approx 215y$ is much larger than the thermal time scale $\lambda^{-1} \approx 25y$, we conclude that the thermal relaxation is a fast process.

The physical value $\widetilde{q}$ of the strength of the thermohaline circulation can be found from the normalised value above as $\widetilde{q} = q V_{\mathrm{box,1}}$, where $V_{\mathrm{box,1}} = V_{\mathrm{box,3}} = V_{\mathrm{box,2}}/V = 1.1 \times 10^{17}\,\mathrm{m}^3$ is the volume of either high-latitude box. Instead, the physical value of the net freshwater flux $\widetilde{F}_i$ into box i is obtained as $\widetilde{F}_i = F_I V_{\mathrm{box,i}}/S_0$; its intensive value per unit surface results to be $\widetilde{F}_i = F_i V_{\mathrm{box,i}}/A_{\mathrm{box,i}} S_0 = F_i D_{\mathrm{box}}/S_0$ where $D_{\mathrm{box}} = 5000\,\mathrm{m}$ is the common depth of the three oceanic boxes. Therefore, our base state features reasonable values for the net poleward transport of freshwater flux – about $2.8 \times 10^5\,\mathrm{m}^3\,\mathrm{s}^{-1} = 0.28\,\mathrm{Sv}$, and for the THC strength – about $1.55 \times 10^7\,\mathrm{m}^3\,\mathrm{s}^{-1} = 15.5\,\mathrm{Sv}$.

3.2 Simplified model

A simplified version of the model given in Eq. (10a–f) can be derived by assuming that the thermal restoring constant $\lambda \to \infty$ so that the time scale of the feedback $\lambda^{-1} \to 0$. Thus, the temperatures of the three boxes are such that at all times $T_i = \bar{T}_i$, so that we obtain a reduced dynamical system with only 2 degrees of freedom (d.o.f.):

$$\dot{S}_1 = \begin{cases} q(2S_0 - 3/2S_1 - 1/2S_3) - F_1 & q > 0 \\ |q|(S_3 - S_1) - F_1 & q \leq 0 \end{cases} \quad (11a)$$

$$\dot{S}_3 = \begin{cases} q(S_1 - S_3) - F_3 & q > 0 \\ |q|(2S_0 - 3/2S_3 - 1/2S_1) - F_3 & q \leq 0 \end{cases} \quad (11b)$$

where the THC strength can be written as $q = k\beta(S_3 - S_1)$. Note that, since the system (11a–b) has been obtained by performing a singular perturbation to (10a–f), we need to renormalize the value of the hydraulic constant k in order to obtain $q_{\mathrm{ref}} = 1.47 \times 10^{-11}\,\mathrm{s}^{-1}$ at steady state when choosing $F_1 = F_3 = 9 \times 10^{-11}\,\mathrm{psu\,s}^{-1}$ as above. The resulting value is $k = 3.0 \times 10^{-7}\,\mathrm{s}^{-1}$, with $|q_{\mathrm{ref}}| = \sqrt{k\beta F_{\mathrm{u}}} = \sqrt{k\beta F_1}$. The same physical scalings described above apply here. Such a simplified model retains the most relevant elements of the dynamics of the full model, even if the thermal dynamical feedback (Scott et al., 1999) are missing.

4 Numerical experiments: fitting the dynamics from the population and the transition rates

4.1 Symmetric forcing to the simplified model

We now modify the dynamical system (11a–b) by including additive noise in both the evolution equations for both S_3 and S_1 so that we obtain the following system of stochastic differential equations in the Ito form:

$$dS_3 = \begin{cases} q(2S_0 - 3/2S_1 - 1/2S_3)dt - F_1 dt + \varepsilon_1 dW_1, & q > 0 \\ |q|(S_3 - S_1)dt - F_1 dt + \varepsilon_1 dW_1 & q \leq 0 \end{cases} \quad (12a)$$

$$dS_3 = \begin{cases} q(S_1 - S_3)dt - F_3 dt + \varepsilon_3 dW_3, & q > 0 \\ |q|(2S_0 - 3/2S_3 - 1/2S_1)dt - F_3 dt + \varepsilon_3 dW_3 & q \leq 0 \end{cases} \quad (12b)$$

where $dW_{1,3}$ are the increments of two independent Wiener processes. Note that Ditlevsen (1999) proposed the possibility of considering more general noise processes to explain the THC dynamics. Hereby, we stick to the more usual white noise case.

We then perform a set of experiments by integrating the stochastic differential Eq. (12a–b) using the numerical scheme proposed in Mannella and Palleschi (1989) for values of $\varepsilon_1 = \varepsilon_3 = \varepsilon$ ranging from $3.6 \times 10^{-10}\,\mathrm{psu\,s}^{-1/2}$ to $6.2 \times 10^{-10}\,\mathrm{psu\,s}^{-1/2}$. This corresponds to a range of noise strength for the physical freshwater flux of $1.12 \times 10^6\,\mathrm{m}^3\,\mathrm{s}^{-1/2}$ to $1.96 \times 10^6\,\mathrm{m}^3\,\mathrm{s}^{-1/2}$. In more concrete terms, we are exploring stochastic perturbations to the freshwater flux whose variability (standard deviation), over the characteristic internal time scale $|q_{\mathrm{ref}}|^{-1} \approx 215\,\mathrm{yr}$, range between 27 % and 47 % of the baseline value $F_1 = F_3$. Results are presented for sets of 100 ensemble members for each value of ε, with each integration lasting $10^6\,\mathrm{yr}$. The chosen time step is 1 yr.

We wish to study the possibility of defining up to a good degree of precision a consistent stochastic dynamics for the THC strength q involving only q itself and noise. Following the procedure outlined in Sect. 2, for each value of $\varepsilon_1 = \varepsilon_3 = \varepsilon$, where we have on purpose kept the system's parameters invariant with respect to exchanging the box 1 and the box 3, we attempt the derivation of the deterministic drift term and the stochastic noise defining the effective Langevin equation for the THC strength:

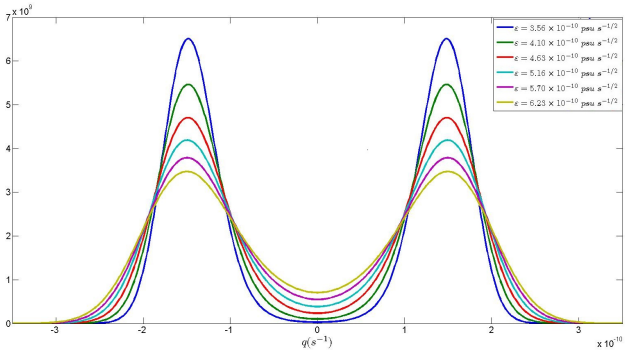


Fig. 1. Empirical probability distribution function for the THC strength in the 2 d.o.f. model for selected values of symmetrically applied noise.

$$dq = F_{\mathrm{eff}}(q,\varepsilon)dt + \varepsilon_{\mathrm{eff}}(\varepsilon)dW \quad (13)$$

where our notation accommodates for a noise-dependent effective drift term, which corresponds to an efficient potential $V_{\mathrm{eff}}(q,\varepsilon)$, such that $F_{\mathrm{eff}}(q,\varepsilon) = -dV_{\mathrm{eff}}(q,\varepsilon)/dq$. The pdfs $\pi(q)$ feature a very strong dependence on the intensity of the noise, with, as expected, higher noise intensity associated to flatter distributions (Fig. 1). We then derive the normalised potential $U(q,\varepsilon) = -\ln(\pi(q,\varepsilon))$ (Fig. 2). By matching the observed hopping rate $r(+\to -)$ (Fig. 3a) with the right hand side of the formula given in Eq. (7) – in Fig. 3b we present the values of the factor $r(+\to -)$ – we derive for each value of ε the corresponding value of $\varepsilon_{\mathrm{eff}}$. Note that for each value of the noise we use only the observed difference between the value of $U(q,\varepsilon)$ evaluated in $q = 0$ and in $q = |q_{\mathrm{ref}}|$ and the value of $|q_{\mathrm{ref}}|$, and the upper bound of ε has been chosen so that $U(q_0,\varepsilon) - U(q_+,\varepsilon) = U(q_0,\varepsilon) - U(q_-,\varepsilon) \gtrsim 1.5$. As shown in Fig. 3c, we obtain that up to a high degree of precision $\varepsilon_{\mathrm{eff}} \approx \gamma\sqrt{2}k\beta\varepsilon = \gamma\bar{\varepsilon}$, where $\gamma \approx 1$ for all values of ε. Moreover, also is in agreement with our expectations given at the end of Sect. 2, we have that $V_{\mathrm{eff}}(q,\varepsilon) \approx V_{\mathrm{eff}}(q)$, so that the effect of adding noise does not impact the deterministic drift term, or, in other terms, a deterministic dynamics is well defined. Figure 4 shows that for all values of noise the obtained effective potentials collapse into a single universal function, apart from an additive constant of no physical significance.

Our experimental procedure has shown quite convincingly that we can reduce the time evolution of the THC strength to a one-dimensional Langevin equation. We wish now to investigate how to derive analytically the drift and the noise term in Eq. (13) and an expression for the hopping rate

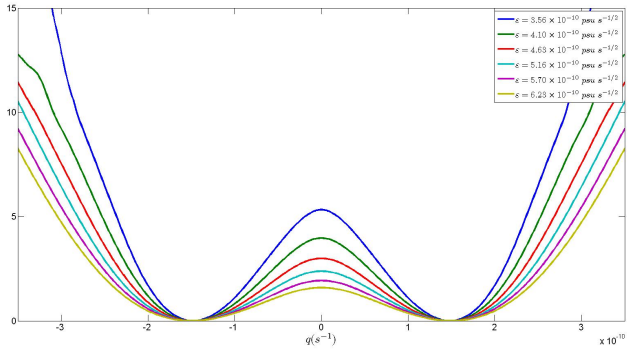


Fig. 2. One-dimensional adimensional potential $U(q)$ obtained as minus the logarithm of the pdf given in Fig. 1.

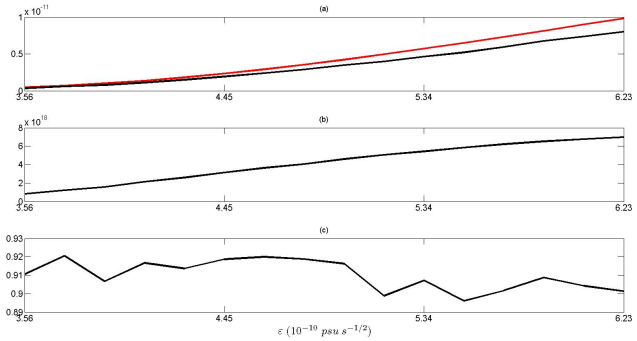


Fig. 3. Goodness of the one-dimensional approach for the reduced 2 d.o.f. model. **(a)** Average rate of hopping between the northern and southern sinking equilibria $r(+\to -) = r(-\to +)$ (black line). The theoretical value is shown with the red line. Data are in units of s^{-1}. **(b)** Geometrical factor $G(+\to -) = G(-\to +)$ of the hopping rate computed from Eq. (7). Adimensional quantity. **(c)** Value of the parameter γ giving the ratio between ε_{eff} (obtained as **a** divided by **b**) and the theoretically derived expression $\sqrt{2}k\beta\varepsilon$. Adimensional quantity.

$r(+\to -)$. Rewriting the system (10a–b) with respect to the new variables $q = k\beta(S_1 - S_3) = k(\rho_1 - \rho_3)$ and $Q = k(\rho_1 + \rho_3)$, we obtain the following coupled evolution equations:

$$dq = q(2k\widetilde{\rho} - 3/2q^2/|q| - Q)dt + \varepsilon_q dW_q \tag{14a}$$

$$dQ = q(2k\widetilde{\rho} + 1/2q^2/|q| - Q)dt - k\beta(F_1 + F_3)dt + \varepsilon_Q dW_Q \tag{14b}$$

where $\widetilde{\rho} = \rho_0(1 - \alpha(\widetilde{T}_1 + 2\widetilde{T}_2 + \widetilde{T}_3)/4 + \beta S_0)$ is the average density of the system, we have that, following Eq. (7), $\varepsilon_q = \varepsilon_Q = k\beta\sqrt{\varepsilon_1^2 + \varepsilon_3^2}$, and dW_q, dW_Q are increments of Wiener processes. Note that the drift terms in both Eq. (11) is odd with respect to the $q \to -q$ transformation and the SDE for the THC strength is in the form of Eq. (9), with $\varepsilon_q - \bar{\varepsilon}$. We assume that the system spends most of its time near the two deterministic equilibria with $q = \mp|q_{\text{ref}}|$, and that over the timescales of our interest the deterministic drift term of the variable Q vanishes. Assuming that the random forcing on Q has little impact on q, we derive the following approximate Langevin equation for the evolution of the THC strength:

$$dq \approx -2q(q^2/|q| - |q_{\text{ref}}|)dt + \varepsilon_q dW_q, \tag{15}$$

where, the drift term is odd with respect to parity and is independent of the noise strength. The corresponding effective potential $V_{\text{eff}}(q)$ is independent of ε and can be written as $V_{\text{eff}}(q) = -q^2|q_{\text{ref}}| + 2/3q^2|q| + \text{const}$. Note that this is a not a quartic symmetric potential but has the same parity properties and is twice differentiable everywhere. As shown in Fig. 4, this functional form closely approximates the experimental findings previously described, with discrepancies where the probability density is exponentially vanishing and small deviations also for $q \approx 0$ (where the density is also low). Moreover, if $\varepsilon_1 = \varepsilon_3$, we obtain that $\varepsilon_q = \bar{\varepsilon} = \sqrt{2}k\beta\varepsilon_1 \approx \varepsilon_{\text{eff}}$, so that the agreement between our experimental and theoretical findings for both the deterministic and stochastic part of the dynamics is quite satisfactory. This suggests that in the experimental setting of Eq. (12a–b) it is possible to project very efficiently the dynamics of the system on the variable q alone, which seems to capture well the slow manifold (Saltzman, 2002) of the system. Using Eq. (15), we obtain the following approximate expression for the average rate of transition $r(+\to -)$ between the northern sinking and the southern sinking state,

$$r(+\to -) \approx \frac{2\sqrt{2}}{3\pi}|q_{\text{ref}}|e^{-\frac{|q_{\text{ref}}|^2}{a(k\beta\varepsilon)^2}}$$

$$= \frac{2\sqrt{2}}{3\pi}\sqrt{k\beta F_1}e^{-\frac{(F_1)^{3/2}}{a\sqrt{k\beta\varepsilon^2}}} = r(-\to +), \tag{16}$$

where the last identity is due to the symmetry of the potential. This formula provides rates in excellent agreement with the outputs of the numerical simulations, as can be seen by comparing the red and the black line in Fig. 3a.

4.2 Asymmetric forcing to the simplified model

The obtained results suggest that the simplified model of the THC with only 2 degrees of freedom allows for a robust treatment of the one-dimensional stochastic dynamics of the THC strength. Nonetheless, in the previous set of experiments we have only verified the first condition for the robustness of the one-dimensional representation (well-posedness for linear scaling on the forcings). In this section, we wish to test how the system behaves when, following Eq. (9), we change the noise matrix ε_{ij} in such a way that the term noise strength $\bar{\varepsilon}$ for the considered observable is not altered. Therefore, we perform a new set of experiments, where the stochastic forcing is exerted only in one of the two boxes, e.g. box 1, so that in Eq. (12a–b) we set $\varepsilon_3 = 0$ and $\varepsilon_1 = \sqrt{2}\varepsilon$ for each value of ε considered in the previous set of experiments. This choice

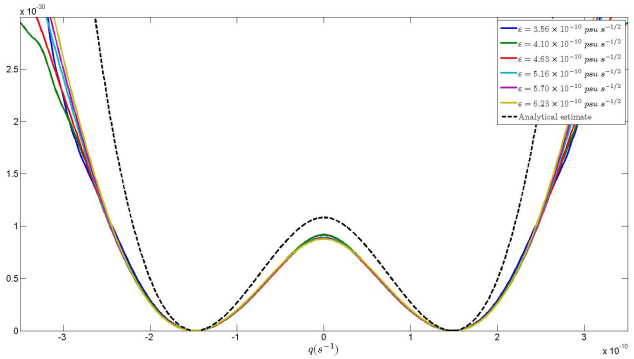


Fig. 4. One-dimensional efficient potential $V_{\text{eff}}(q)$ controlling the evolution of the THC strength in the 2 d.o.f. model with symmetric noise. The results of some numerical experiments are shown together with the analytical estimate.

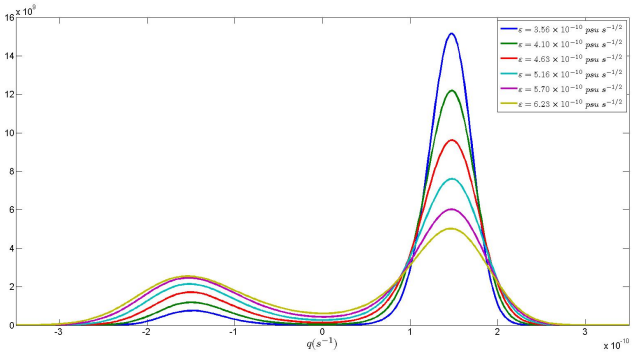


Fig. 5. Empirical probability distribution function for the THC strength in the 2 d.o.f. model with noise applied only in box 1.

guarantees that we have exact correspondence for the values of $\varepsilon_q = \bar{\varepsilon} = k\beta\sqrt{\varepsilon_1^2+\varepsilon_3^2} = \sqrt{2}k\beta\varepsilon$, so that Eq. (9) looks exactly the same as in the previous set of experiments. Unfortunately, as shown in Fig. 5, following the procedure described in Sect. 2 we obtain for all values of ε an asymmetric probability distribution function, with the northern sinking equilibrium being the most probable state. The prominence of the $q > 0$ conditions become stronger as we consider weaker intensities for the noise. Therefore, the proposed stochastic modelling is not as robust as one could have guessed.

At a second thought the presence of asymmetry in this case becomes clearer. In this case the two sinking states undergo different forcing, because when $q > 0$ the stochastic forcing is exerted only in the box where downwelling occurs, whereas when $q \leq 0$ the stochastic forcing impacts only the box where upwelling occurs. Since, as explained at the beginning of Sect. 3 and discussed in Lucarini and Stone (2005a), the impact of changing freshwater fluxes is different in terms of destabilising the system depending on whether the forcing is applied in the box where downwelling or upwelling occurs, the two states are not equivalent

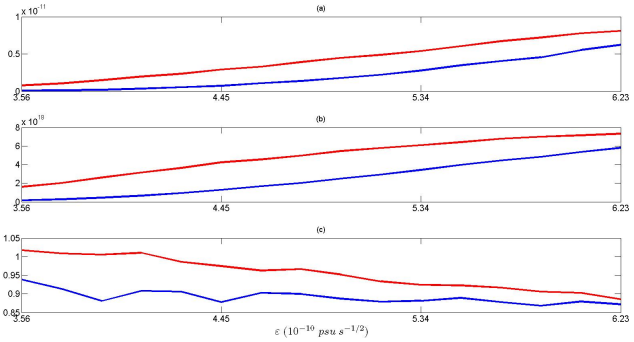


Fig. 6. Goodness of the one-dimensional approach for the non-symmetric case. **(a)** Average hopping rates $r(+\rightarrow -)$ (blue line) and $r(-\rightarrow +)$ (red line). Data are in units of s^{-1}. **(b)** Geometrical factors $G(+\rightarrow -)$ (blue line) and $G(-\rightarrow +)$ (red line) computed from Eq. (7). Adimensional quantities. **(c)** Value of the parameter $r\gamma$ giving the ratio between ε_{eff} (obtained as **a** divided by **b**) and the theoretically derived expression $\sqrt{2}k\beta\varepsilon$. Adimensional quantities. The values obtained using the $+\rightarrow -$ (blue) and $-\rightarrow +$ (red) processes are not compatible.

anymore. In the previous set of experiments, as opposed to that, both boxes were equally (in a statistical sense) stochastically forced, and so that the northern and southern sinking states had equivalent forcings at all times. The more formal mathematical reason why the statistical properties of q are different in the two sets of experiments even if Eq. (9) is apparently the same can be traced to the differences in the correlation between the stochastic forcings to q and Q – compare Eq. (14a–b). In the case of symmetric stochastic forcing in the freshwater fluxes into the two boxes with $\varepsilon_1 = \varepsilon_3$, it is easy to see that the increments to the Wiener processes dW_q and dW_Q are not correlated, whereas when $\varepsilon_3 = 0$ the two quantities dW_q and dW_Q are identical so that their correlation is unitary.

We wish now to test whether, in such an asymmetric setting of forcings, the pdfs of the THC strength scale with the intensity of the noise in such a way to allow the possibility of defining consistently an effective potential $V_{\text{eff}}(q,\varepsilon)$ driving the deterministic part of the one-dimensional stochastic evolution of the THC strength. Given the asymmetry of the pdfs, such an effective potential would, unavoidably, be different from the one derived in the previous set of experiments, so that in no way we can be satisfied in terms of robustness of the one-dimensional Langevin approach. But, if we are able to define such an effective potential, we can deduce that each choice of the correlation matrix for the noise in the full system determines a specific projected effective deterministic dynamics, which is a weaker, but maybe still useful, result.

We follow the procedure described in Sect. 2, and for all the chosen values of ε we have that $U(q_0,\varepsilon) - U(q_-,\varepsilon) \gtrsim 1.5$ and $U(q_0,\varepsilon) - U(q_=,\varepsilon) \gtrsim 1.5$. In Fig. 6a we present the hopping rates $r(+\rightarrow -)$ (blue) and $r(-\rightarrow +)$ (red). We see that both values increase monotonically with ε, as stronger

noise favours transitions. Note that, quite unexpectedly, for $\varepsilon > 7 \times 10^{-10}$ psu s$^{-1/2}$ (beyond the range where our full analysis is performed, not shown), $r(+ \to -)$ becomes bigger. By simple population algebra, this implies that the fraction of states with $q < 0$ is larger than 1/2, even if the most probable state is given, in all cases, by $q = q_{\rm ref} > 0$. In Fig. 6b we plot the factors $G(+ \to -)$ and $G(- \to +)$. By matching the rate of transition with the corresponding G factor, one obtains for each value of ε the effective noise intensity εeff such that the probability distribution of the states and the transitions statistics are compatible. As mentioned before, in order to have a consistent picture, we need to obtain the same value of $\varepsilon_{\rm eff}$ by either using the $+ \to -$ or the $- \to +$ path. In Fig. 6c (compare with Fig. 3c) we show that $\varepsilon_{\rm eff}$ does not scale linearly with ε (or, equivalently, γ is not a constant) as found in the symmetric case, and, much more seriously, that there is no consistency between its value as obtained using the statistics of the $+ \to -$ and $- \to +$ transitions. Given the mismatch between the time-scale of the transitions and the populations in the two basins, we cannot reconstruct a well-defined effective potential $V_{\rm eff}(q, \varepsilon)$, so that a consistent one-dimensional Langevin representation of the dynamics and statistics of the THC strength as proposed in Eq. (13) is not possible here. This can be the case if noise can effectively activate non-trivial dynamical processes allowing for a transition between the neighbourhoods of the two steady states with $q = \pm|q_{\rm eff}|$, where by non-trivial we mean that they cannot be represented even approximately as a function of q only.

4.3 Symmetric forcing to the full system

We now revert to the full system described at the beginning of Sect. 3 and consider the case of symmetric boundary conditions. By adding stochastic perturbations to the freshwater fluxes in a similar fashion as in Eq. (11a–b), so that $F_j \to F_j = \varepsilon_j dW_j/dt$ with $j = 1.3$ in Eq. (10a–f) we obtain the following Langevin equations for the variables $q = k(\rho_1 - \rho_3)$ and $Q = k(\rho_1 + \rho_3)$:

$$dq = q(2\widetilde{q} - 3/2q^2/|q| - Q)dt + k\alpha\lambda(T_1 - T_3)dt + \varepsilon_q dW_q, \quad (17a)$$

$$dQ = q(2k\widetilde{\rho} + 1/2q^2/|q| - Q)dt - \beta(F_1 + F_3)dt + k\alpha\lambda(\widetilde{T_1} - T_1 + \widetilde{T_3} - T_3)dt + \varepsilon_Q dW_Q, \quad (17b)$$

where the same notation as in Eq. (14a–b) has been used. Note that, as opposed to Eq. (11), the deterministic drift term is not odd with respect to the $q \to -q$ transformation, since in this case explicit temperature dependent terms are present, so that a negative parity is realised only when the signs of both $T_1 - T_3$ and $S_1 - S_3$ are changed. Note that, following the same derivation as in the case of the system with 2 d.o.f. and assuming that over time-scales of interest the drift of Q is negligible, we end up writing the same approximate autonomous Eq. (15) for the THC strength under the hypothesis that the system spends most time near the two deterministic fixed points $q = \pm|q_{\rm eq}|$. This suggests that also in this case we might empirically derive a well-defined effective potential $V_{\rm eff}(q)$ analogous to the one obtained for the 2 d.o.f. model if considering stochastic forcing acting on both boxes 1 and 3.

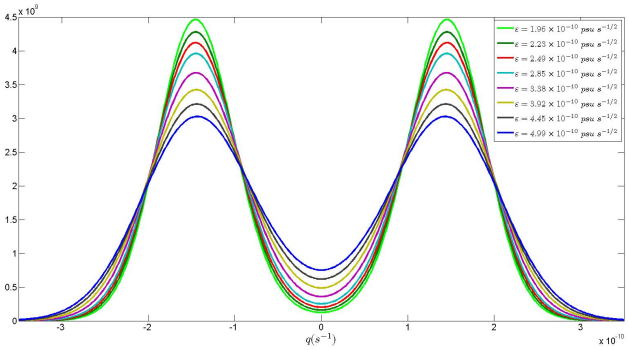


Fig. 7. Empirical probability distribution function for the THC strength in the full model for some selected values of symmetrically applied noise.

Therefore, we follow the analysis of the previous subsection, and concentrate to stochastic perturbations to the freshwater flux having the same strength $\varepsilon_1 = \varepsilon_3 = \varepsilon$ in both hemispheres. We first observe that, as anticipated, for a given value of ε the distribution of the THC strength is flatter than in the case of the 2 d.o.f. model (Fig. 7). In order to obtain a pdf analogous to what obtained in the 2 d.o.f. case, in the full model we need to consider a stochastic forcing smaller by about 25 %, which suggests that the full model is, in some sense, less stable. We will discuss this below. As before, we select values of ε such that $U(q_0, \varepsilon) - U(q_=, \varepsilon) = U(q_o, \varepsilon) - U(q_-, \varepsilon) \gtrsim 1.5$ in order to be able to use Kramers' formula as a constraint to check the consistency of our data with the one-dimensional Langevin model. In Fig. 8a we report the hopping rates as a function of the intensity of the noise. The behaviour is qualitatively analogous to what shown in Fig. 3a for the 2 d.o.f. model, but, in agreement with the discussion above, the hopping rate between the $q > 0$ and the $q < 0$ states are consistently higher for the full model when the same stochastic forcing is considered. In Fig. 8b we present the factor $G(+ \to -) = G(- \to +)$ introduced in Eq. (7), which depends uniquely on the ratio between the probability density at the two maxima and at the local minimum for $q = 0$. Figure 8c, similarly to Fig. 3c, shows the proportionality constant γ between the value of $\varepsilon_{\rm eff}$ and the value of $\bar{\varepsilon} = \sqrt{2k\beta}\varepsilon$. The parameter γ should be close to unity in the case the projection of the dynamics on q is "trivial" (as in the case of the 2 d.o.f. model with symmetric stochastic forcing), and, more importantly, γ should be approximately independent of ε. In the case analysed here both

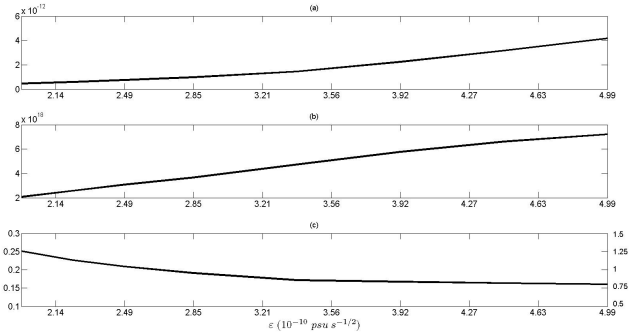


Fig. 8. Goodness of the one-dimensional approach for the full model. **(a)** Average rate of hopping between the northern and southern sinking equilibria $r(+\to-)=r(-\to+)$ (black line). The theoretical value is shown with the red line. Data are in units of s^{-1}. **(b)** Geometrical factor $G(+\to-)=G(-\to+)$ of the hopping rate computed from Eq. (7). Adimensional quantity. **(c)** Value of the parameter γ giving the ratio between $\varepsilon_{\rm eff}$ (obtained as **a** divided by **b**) and the theoretically derived expression $\sqrt{2}k\beta\varepsilon$. The value of γ obtained if considering the renormalized value (referred to the 2 d.o.f. model) for k can be read on the right hand side scale. Adimensional quantities.

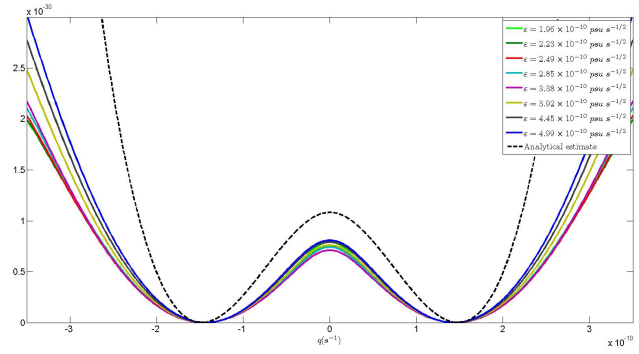


Fig. 9. One-dimensional efficient potential $V_{\rm eff}(q)$ controlling the evolution of the THC strength in the 5 d.o.f. model with symmetric noise. The results of some numerical experiments are shown together with the analytical estimate. The results are rather similar to what shown in Fig. 4.

conditions are not satisfied. The first issue points to the fact that we need to renormalize the constants when developing a lower dimensional projected dynamics (which is exactly what we did when constructing the 2 d.o.f. model from the full model). In fact, if we consider as effective k the one considered for the 2 d.o.f. model, the values of γ are relatively close to 1 (check the scale on the right hand side of Fig. 8c). More critical is the presence of a nonlinear relation between $\varepsilon_{\rm eff}$ and ε, which is caused by the fact that in the full model faster modes are excited by noise and impact in a nontrivial way the effective surrogate noise acting on the q variable taken as independent.

When reconstructing from $U(q,\varepsilon)$ the actual effective potential $V_{\rm eff}(q,\varepsilon)$ using Eq. (5), we obtain that the effective potential is a function of q only, so that $V_{\rm eff}(q,\varepsilon)=V_{\rm eff}(q)$ (see Fig. 9) so that it is possible to disentangle completely the drift term from the stochastic forcing. Moreover, such potential is very similar to what obtained in the reduced model with 2 d.o.f., as can be seen by comparing Figs. 4 and 9. The height of the potential barrier between the two minima is slightly lower in the case of the full model analysed here, in agreement with the argument of the destabilising feedback due to the thermal restoring process presented by Lucarini and Stone (2005a). This can be explained as follows: since perturbations in the value of $q-|q_{\rm eq}|$ are positively correlated to perturbations T_1-T_3 thanks to advection, the contribution $k\alpha\lambda(T_1-T_3)$ in Eq. (17a) weakens the force driving the system towards the nearby deterministic fixed point, thus enhancing its instability. Overall, the good agreement between Figs. 4 and 9 implies that the deterministic dynamics of the THC strength is robustly consistent between the full and reduced model.

5 Numerical experiments: stochastic resonance

Stochastic resonance (Benzi et al., 1982; Nicolis 1982; Gammaitoni et al., 1998) is an exceedingly interesting process whereby noise amplifies the response of the system at the same frequency of a periodic forcing. Typically, it is realised when we consider a Langevin equation of the form:

$$dx = F(x)dt + A\sin(\omega t+\phi)dt + \varepsilon dW \tag{18}$$

where the drift term $F(x)=-dV(x)/dx$ derives from a (symmetric, but necessarily so) potential $V(x)$ with a double well-structure like those considered in this study. We can associate the time dependent drift $G(x,t)=F(x)+A\sin(\omega t+\phi)$ to a time dependent potential $W(x,t)=-\int dx F(x)-A\sin(\omega t+\phi)=V(x)-Ax\sin(\omega t+\phi)$. The periodic forcing modulates the bistable system, so that one stable state corresponding to one of the two minima results to be less stable than the other every half a period of the forcing, so that every period the populations of the neighbourhoods of these points are, alternatively, exponentially increased and suppressed by a factor $\exp(\pm 2A(|x_{=}-x_0|)/\varepsilon^2)$. As accurately discussed in Gammaitoni et al. (1998), when we tune the noise intensity ε so that the inverse of the hopping rate given in Eq. (7) for the unperturbed system is approximately equal to half of the period π/ω, the system is in a condition where there is maximum probability for leaving the less stable state into the more stable one, before, randomly, the system switches back. In this case, the system attains a high degree of synchronisation with the input periodic signal, so that the output is basically a square wave with constant phase difference with the sinusoidal forcing.

Interestingly, even if the stochastic resonance is apparently a highly nonlinear process, it can be accurately described using linear response theory, whereby one studies the amplitude of the output signal at the same frequency of the periodic forcing for various values of the noise. Such amplitude,

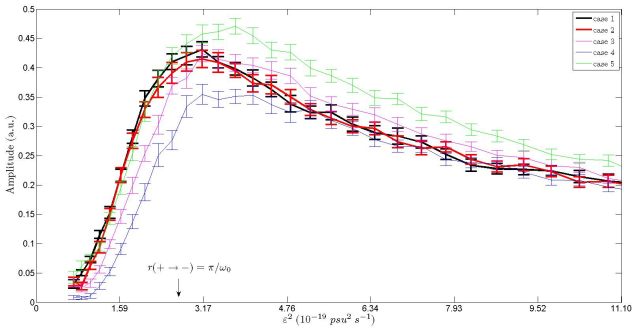


Fig. 10. Response to linear periodic perturbation as a function of the background noise for five experimental settings. Case 1 and 2 are indicative of stochastic resonance. The value of noise realising the approximate matching condition $r(+\to-)=r(-\to+)=\pi/\omega_0$ is indicated.

in the weak field limit, is proportional to the amplitude of the periodic forcing A (Gammaitoni et al., 1998).

Along these lines, we consider the 2 d.o.f. model, and take into account a periodic modulation to the freshwater fluxes, so that $F_1 \to F_2 + \phi \Delta F \sin(\omega t)$ and $F_3 \to F_3 + \psi \Delta F \sin(\omega t)$, with symmetric background forcing $F_1 = F_3$. Assuming that the stochastic forcing acts with equal strength $\varepsilon_1 = \varepsilon_3 = \varepsilon$ on both boxes 1 and 3, as analysed in Sect. 4.1, we obtain that Eq. (15), which satisfactorily describes the one-dimensional stochastic dynamics of q, is modified as follows:

$$dq \approx -2q(q^2/|q| - |q_{\mathrm{ref}}|)dt + k\beta(\phi - \psi)\Delta F \sin(\omega t) + \varepsilon_q dW_q, \quad (19)$$

which is exactly in the form of Eq. (15). As we see, for a given value of ΔF, the strength of the periodic forcing to q depends only on the absolute value of the difference $(\phi - \psi)$, and not separately on the values of ϕ and ψ. If the dynamics of q is accurately described by a one-dimensional Langevin equation, we expect to be able to observe the process of stochastic resonance when ω and ε_q are suitably matched. In order to test this, we set the period $\omega = \omega_0 = 2\pi/19\,000$ yr – where the period of 19 000 yr has been chosen because it is long compared to the internal time scale $|q_{\mathrm{ref}}|^{-1} \approx 215$ yr and has paleoclimatic relevance in conjunction to Milankovitch theory (Velez-Bechi et al., 2001; Saltzman, 2002) – choose a moderate value for the amplitude of the sinuisodal modulation $\Delta F = 9 \times 10^{-12}\,\mathrm{s}^{-2}$ and study the amplitude of the ω_0 frequency component of the times series of q as a function of ε, and create an ensemble of 100 members for each considered setting. We can state that the phenomenology of stochastic resonance is well reproduced if (a) we find the characteristic peak for the response in the vicinity of a value of ε such that the corresponding hopping rate for the unperturbed system given in Fig. 3a is close to π/ω_0, and (b) such response depends, for all values of ε, on $(\phi - \psi)$ only. We refer to the scenario where $\phi = -\psi = 1/2$ as case 1, and the scenario where $\phi = 1, \psi = 0$ as case 2. Note that the case $\phi = 0, \psi = -1$ is identical to case 2 by symmetry. The results are reported in Fig. 10, with the black line corresponding to case 1 and the red line corresponding to case 2. The obtained curves for the amplitude of the response agree very accurately, especially considering the rather small uncertainty, and feature exactly the right shape as presented in Gammaitoni et al. (1998). We observe a relatively broad resonance for values of noise comparable to those inducing in the unperturbed system transitions with average rate similar to the semiperiod of the forcing. Finding quite accurately the signature of stochastic resonance is a further proof that in the special setting of the unperturbed system considered here the dynamics of q is indeed quasi-one dimensional.

We want to contrast this positive outcome with what one obtains by adding periodic perturbations of the same form as above to an "unperturbed" state featuring stochastic forcing acting on box 1 only, described in Sect. 4.2. We choose exactly the same forcing parameters as above and repeat the experiments using the same ensemble size. We refer to the scenario where $\phi = -\psi = 1/2$ as case 3, and the scenario where $\phi = 1, \psi = 0$ as case 4, and to the scenario where $\phi = 0, \psi = -1$ as case 5. Note that here case 4 and case 5 are not equivalent. We obtain (see Fig. 10) that in the three cases, whereas we obtain qualitatively and quantitatively analogous results for the normalised amplitude of the response of the output at the same frequency of the forcing, the curves are distinct with high statistical significance and differ also from what obtained for cases 1 and 2. Such discrepancy would not be possible if the dynamics of q were accurately described with a one-dimensional effective potential plus stochastic noise plus periodic forcing. The fact that the three curves 3, 4, and 5 are not superimposed (and disagree with 1 and 2) further supports the fact that the dynamics of q is not quasi-one dimensional. Nonetheless, an obvious process of resonance (not strictly one-dimensional) is obviously still in place.

6 Conclusions

In this work we have re-examined the classic problem of trying to reconstruct the effective stochastic dynamics of an observable from its time series in the special case of a clearly bimodal empirical probability density function. This issue is especially relevant in climatic and paleoclimatic research, where it is very tempting to try to deduce the large scale, qualitative properties of the climate system, their modulation with time, the potential presence of tipping points through the observations of long time series of proxy data (Livina and Lenton, 2007; Livina et al., 2009). Furthermore, since amplifying mechanisms such as stochastic resonance have been proposed to explain enhanced low frequency variability of the oceanic circulation as a result of slow modulations of some parameters (Ganopolski and Rahmstorf, 2002), such empirically reconstructed statistical/dynamical

properties may be interpreted as starting points to deduce special "sensitivities" of the climate system. Therefore, an important question is to understand how accurate and robust these procedures of reconstruction are.

From this work it is apparent that from the statistical properties of the time series of an observable featuring a symmetric bimodal pdf it seems relatively easy to construct the corresponding one-dimensional Langevin equation by determining the drift term and the intensity of the white noise by basically imposing a self-consistent population dynamics. Assuming that the observable is a function of the phase space variables of a stochastic dynamical systems, it is an obvious temptation to interpret the obtained equation as the description of the projected dynamics for the observable, where the impact of the other (in general, many) neglected degrees of freedom of the system contributes to defining the effective deterministic dynamics and to creating a surrogate white noise term. Nonetheless, if the pdf of the observable is not symmetric, the possibility of constructing a meaningful surrogate stochastic dynamics relies on the fact that one should be able to describe consistently the hopping process and between the two attraction basins and their steady state populations.

In this work we have considered two very simple box models of the oceanic circulation (Rooth, 1982; Scott et al., 1999; Lucarini and Stone, 2005a, b), comprising two high latitude and a low latitude boxes with time-dependent temperature and salinity as testbeds for these methodologies. These models are able to reproduce the bistability properties of the thermohaline circulation, by featuring two possible asymmetric circulations (one mirror image of the other) in presence of symmetric external heat and freshwater forcings. In both models the circulation strength is parameterised as proportional to the difference between the densities of the two high-latitude boxes. The simpler 2 d.o.f. model is suitably derived from the full, 5 d.o.f. model by imposing a fixed temperature for the boxes.

We first impose stochastic forcing of the same intensity on the freshwater forcings to the two high-latitude boxes and observe that the resulting pdf of the thermohaline circulation strength is bimodal and symmetric. More importantly, for both models the dynamics of q can be accurately described with a Langevin equation with a drift term derived from a one-dimensional effective potential plus stochastic noise. An excellent approximation to the true dynamics (as well as to the hopping rates) can be obtained in an explicit form by imposing that the sum of the densities of the two high-latitude boxes is a slow variable. The main difference between the two is that in the full model the nonlinear feedbacks acting on the variable we are neglecting alter in a nontrivial, nonlinear way the effective surrogate noise acting on the q variable. In other terms, in the case of the full model a careful tuning of the noise allows for taking care very accurately – in a statistical sense – of the effect of all the variables we are neglecting. Our results are obtained for a specific value of the background freshwater forcing $F_1 = F_3$, but the following scaling allow to extensively generalise our findings: $q \sim (F_1)^{1/2}, V_{\text{eff}}(q) \sim (F_1)^{3/2}, \varepsilon_{\text{eff}} \sim (F_1)^{3/4}$. Things change dramatically when considering the case of stochastic forcing acting only on one of the two high-latitude boxes. We tune our experiments in such a way that, apparently, the Langevin equation for the q variable is not altered with respect to the previous case. Not only we obtain a non-symmetric pdf, but, moreover, it is not possible to reconstruct an approximate but consistent stochastic dynamics for the q variable alone. Therefore, there is no ground for achieving a satisfactory projection of the dynamics, and a one-dimensional Langevin equation cannot be derived.

Finally, we test the possibility of observing the mechanism of stochastic resonance in the simplified 2 d.o.f. model by superimposing a slow periodic modulation on the freshwater fluxes in the two high-latitude boxes to the acting stochastic forcings. Whereas in the scenario where the noise acts with equal strength in both boxes we obtain numerically outputs in close agreement with the theory stochastic resonance for one dimensional systems, thus supporting the idea that a projected dynamics is indeed a good approximation when attempting a description of the properties of q, the opposite holds in the scenario where the noise acts only on one box. This further supports the fact that in this case the dynamics of q is not trivially quasi-one dimensional, and transitions occur through processes that cannot be written precisely as a function of q only.

Our results support the idea that deducing the approximate stochastic dynamics for an observable of a multidimensional dynamical system from its time series is definitely a non-trivial operation. The reconstructed drift term and the noise forcing depend, in general, in a non-trivial way on the intensity and correlation properties of the white noise of the true system. This implies that a "true" projected dynamics cannot be defined. Therefore, in practical applications, it seems tentative to assume that from the pdf of a bimodal observable obtained with a given level of noise it is possible to understand how the bistability property of the full system will change when the level of the input noise is altered. In particular, it seems difficult to be confident on obtaining information on how the rate of transition between the two equilibria will change and on the characteristics of tipping points. This also suggests that recent claims of the possibility of detecting robustly early warning signals for critical transitions at tipping points from time series as proposed in Scheffer (2008) must be carefully checked.

A better understanding of the properties of multistable models can be reached only by going beyond a simplified description of the statistical properties of the observables we are mostly interested into. In order to address these points, we will attempt the kind of critical analysis proposed in this paper on more complex models but still idealised models of the thermohaline circulation, such as that considered in Lucarini et al. (2005, 2007). We will test to what extent a

simple one-dimensional Langevin model for the evolution of the THC strength can be defined and in which range of the parameters determining the boundary conditions, and how the drift and noise terms of the Langevin model – if it can be defined – are related to the internal parameters of the full system. At a more theoretical level, we will try to propose approximate ways for dealing with the transitions between high occupations regimes in multidimensional gradient flows, attempting to derive a markovian description of transitions between discrete states.

Acknowledgements. The authors wish to acknowledge the exchanges with J. Imbers and T. Kuna. VL and DF acknowledge the financial support of the EU-ERC project "Thermodynamics of the Climate System" – NAMASTE. This work stems from scientific interactions coming from lectures and seminars held within the Bachelor and Master Degree course in Atmospheric Physics and Meteorology at the University of Bologna. We regret that Her Excellency the then Italian Minister of Education, University, and Scientific Research, M. S. Gelmini, made this course redundant in 2009–2010 because of its apparent cultural irrelevance.

Edited by: J. Duan
Reviewed by: L. Wan and another anonymous referee

References

Benzi, R., Parisi G., Sutera A., and Vulpiani A.: Stochastic resonance in climatic change, Tellus 34, 10–16, 1982.

Benzi, R.: Stochastic resonance: from climate to biology, Nonlin. Processes Geophys., 17, 431–441, doi:10.5194/npg-17-431-2010, 2010.

Boyle, E. and Kegwin, L.: North Atlantic thermohaline circulation during the past 20,000 years linked to high-latitude surface temperature, Nature, 330, 35–40, 1987.

Cessi, P.: A simple box-model of Stochastically forced Thermohaline flow, J. Phys. Ocean., 24, 1911–1920, 1994.

Cessi, P. and Young, W. R.: Multiple equilibria in two-dimensional thermohaline circulation, J. Fluid Mech., 241, 291–309, 1992.

Dijkstra, H.: Nonlinear Physical Oceanography, Springer, Berlin, 2005.

Ditlevsen, P.: Observation of alpha-stable noise and a bistable climate potential in an ice-core record, Geophys. Res. Lett., 26, 1441–1444, 1999.

Freidlin, M. I. and Wentzell, A. D.: Random perturbations of dynamical systems, Springer, Berlin, 1998.

Fraedrich, K.: Structural and stochastic analysis of a zero-dimensional climate system, Q. J. R. Meteorol. Soc., 104, 461–474, 1978.

Gammaitoni, L., Hänggi, P., Jung, P., and Marchesoni, F.: Stochastic Resonance, Rev. Mod. Phys., 70, 223–287, 1998.

Ganopolski, A. and Rahmstorf, S.: Abrupt Glacial Climate Change due to Stochastic Resonance, Phys. Rev. Lett., 88, 038501, doi:10.1103/PhysRevLett.88.038501, 2002.

Hasselmann, K.: Stochastic climate models, Part 1: Theory, Tellus, 28, 473–485, 1976.

Hawkins, E., Smith, R. S., Allison, L. C., Gregory, J. M., Woollings, T. J., Pohlmann, H., and de Cuevas, B.: Bistability of the Atlantic overturning circulation in a global climate model and links to ocean freshwater transport, Geophys. Res. Lett., 38, L10605, doi:10.1029/2011GL047208, 2011.

Kuhlbrodt, T., Griesel, A., Montoya, M., Levermann, A., Hofmann, M., and Rahmstorf, S.: On the driving processes of the Atlantic meridional overturning circulation, Rev. Geophys., 45, RG2001, doi:10.1029/2004RG000166, 2007.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, Proc. Natl. Acad. Sci. USA, 105, 1786–1793, 2008.

Livina, V. N. and Lenton, T. M.: A modified method for detecting incipient bifurcations in a dynamical system, Geophys. Res. Lett., 34, L03712, doi:10.1029/2006GL028672, 2007.

Livina, V. N., Kwasniok, F., and Lenton, T. M.: Potential analysis reveals changing number of climate states during the last 60 kyr, Clim. Past, 6, 77–82, doi:10.5194/cp-6-77-2010, 2010.

Lucarini, V. and Stone, P. H.: Thermohaline circulation stability: a box model study, Part I: uncoupled model, J. Climate, 18, 501–513, 2005a.

Lucarini, V. and Stone, P. H.: Thermohaline circulation stability: a box model study, Part II: coupled atmosphere-ocean model, J. Climate, 18, 514–529, 2005b.

Lucarini, V., Calmanti, S., and Artale, V.: Destabilization of the thermohaline circulation by transient perturbations to the hydrological cycle, Clim. Dynam., 24, 253–262, 2005.

Lucarini, V., Calmanti, S., and Artale, V.: Experimental mathematics: Dependence of the stability properties of a two-dimensional model of the Atlantic ocean circulation on the boundary conditions, Russ. J. Math. Phys., 14, 224–231, 2007.

Lucarini, V., Fraedrich, K., and Lunkeit, F.: Thermodynamic Analysis of snowball Earth hysteresis experiment: Efficiency, Entropy Production, and irreversibility, Q. J. Royal Met. Soc., 136, 2–11, 2010.

Mannella, R. and Palleschi, V.: Fast and precise algorithm for computer simulation of stochastic differential equations, Phys. Rev. A, 40, 3381–3386, 1989.

Monaham A. H.: Stabilization of Climate Regimes by Noise in a Simple Model of the Thermohaline Circulation, J. Phys. Ocean., 32, 2072–2085, 2002.

Nicolis, C.: Stochastic aspects of climatic transitions-response to a periodic forcing, Tellus, 34, 1–9, 1982.

Pierrehumbert, R. T, Abbot, D., Voight, A., and Koll, D.: Neoproterozoic Climate, Ann. Rev. Earth. Plan. Sci. 39, 417–460, 2011.

Rahmstorf, S.: Bifurcation of the Atlantic thermohaline circulation in response to changes in the hydrological cycle, Nature, 378, 145–149, 1995.

Rahmstorf, S.: Ocean Circulation and Climate during the last 120.000 years, Nature, 419, 207–214, 2002.

Rahmstorf, S.: Timing of abrupt climate change: a precise clock, Geophys. Res. Lett., 30, 1510, doi:10.1029/2003GL017115, 2003.

Rahmstorf, S., Crucifix, M., Ganopolski, A., Goosse, H., Kamenkovich, I., Knutti, R., Lohmann, G., Marsh, R., Mysak, L. A., Wang, Z., and Weaver, A. J.: Thermohaline circulation hysteresis: A model intercomparison, Geophys. Res. Lett., 32, L23605, doi:10.1029/2005GL023655, 2005.

Rooth, C.: Hydrology and ocean circulation, Progr. Oceanogr., 11, 131–149, 1982.

Saltzman, B.: Dynamical Paleoclimatology: Generalized Theory of Global Climate Change, Academic Press, New York, 2002.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., and Sugihara, G.: Early warning signals for critical transitions, Nature, 461, 53–59, 2009.

Scott, J. R., Marotzke, J., and Stone, P. H.: Interemispheric Thermohaline circulation in a coupled box-model, J. Phys. Ocean., 29, 351–365, 1999.

Scott, J. R., Sokolov, A. P., Stone, P. H., and Webster, M. D.: Relative roles of climate sensitivity and forcing in defining the ocean circulation response to climate change, Clim. Dynam., 30, 441–454, 2008.

Sijp, W. P. and England, M. H.: Sensitivity of the Atlantic thermohaline circulation to basin-scale variations in vertical mixing and its stability to fresh water perturbations, J. Climate, 19, 5467–5478, 2006.

Sijp, W. P. and England, M. H.: Precise calculations of the existence of multiple AMOC equilibria in a coupled climate model Part II: transient behaviour, J. Climate, in press, doi:10.1175/2011JCLI4246.1, 2011.

Sijp, W. P., England, M. H., and Gregory, J. M.: Precise calculations of the existence of multiple AMOC equilibria in coupled climate models Part I: equilibrium states, J. Climate, in press, doi:10.1175/2011JCLI4245.1, 2011.

Stocker, T. F. and Schmittner, A.: Influence of CO_2 emission rates on the stability of the thermohaline circulation, Nature, 388, 862–865, 1997.

Stocker, T. F. and Wright, D. G.: Rapid transitions of the ocean's deep circulation induced by changes in the surface water fluxes, Nature, 351, 729–732, 1991.

Stouffer, R. J. and Manabe, S.: Equilibrium Response of Thermohaine Circulation to large Changes in Atmospheric CO_2 concentration, Clim. Dynam. 20, 759–777, 2003.

Stommel, H.: Thermohaline convection with two stable regimes of flow, Tellus, 13, 224–230, 1961.

Titz, S., Kuhlbrodt, T., and Feudel, U.: Homoclinic bifurcation in an ocean circulation box model, Int. J. Bif. Chaos, 12, 869–875, 2001.

Titz, S., Kuhlbrodt, T., Rahmstorf, S., and Feudel, U.: On freshwater-dependent bifurcations in box models of the interhemispheric thermohaline circulation, Tellus, 54A, 89–98, 2002.

Velez-Belchi, P., Alvarez, A., Colet, P., Tintor, J., and Haney, R.L.: Stochastic Resonance in the thermohaline circulation, Geophys. Res. Lett., 28, 2053–2057, 2001.

Vellinga, M.: Instability of two-dimensional thermohaline circulation, J. Phys. Oceanogr., 26, 305–319, 1996.

Chapter 9

Article: Extreme Value Theory provides early warnings for critical transitions

Extreme value Theory provides early warnings for critical transitions

Davide Faranda, Valerio Lucarini, and Jeroen Wouters
Klimacampus, Universitaet Hamburg
*Grindelberg 5, 20144, Hamburg, Germany**

Early warning of critical transitions have been extensively used to detect abrupt changes of dynamical regimes. In this paper we introduce new indicators based on the extreme value statistics of observables of the system. The indicators are based on quantitative modifications in the properties of extremes for chaotic systems which possess extreme value laws when far from bifurcation points. By measuring the deviation from these expected laws is possible to detect the approaching critical transitions. Moreover, relations and connections between traditional and extreme value indicators are explained and commented in detail. Numerical experiments have been performed on a stochastic differential equation describing the motion of a particle in an asymmetric double well potential under the effect of noise. The results meet the theoretical expectations and the example provide a gateway for using operatively the method described in other systems or data series analysis.

I. INTRODUCTION

An astonishing variety of complex systems ranging from finance to climate can experience abrupt changes at which a sudden shift of dynamical regime occurs. These so called "critical transitions" or "tipping points" can be formally seen as bifurcations in the dynamical systems terminology [1]. There is currently a great interest in understanding the dynamical behaviour in the proximity of a tipping point mainly for its importance in analysing and forecasting events of extreme social and economic relevance [2, 3]. To accomplish this task in the past decades a growing interest in this topic has emerged and many indicators of criticality have been developed to identify early warnings of abrupt transitions to different dynamical states: the most used indicators are based on the modifications of the auto-correlation properties of particular observables when the system is pushed towards a transition, usually accompanied by an increase of the variance and the skewness of the distribution of the observable analysed [2–4]. Although for specific systems such techniques successfully highlight approaching tipping points, several difficulties prevent to extend these results to general cases. In particular, for high dimensional and complex systems featuring oscillations and intricate bifurcation patterns, early warning signals may be misleading. The main issue for these systems concerns the accessibility of dynamical and geometrical properties of the physical measure (i.e. the attractor) in the proximity of tipping points and the difficulties in recognizing the nature of the bifurcation involved. Moreover, since all the early warnings indicators rely on the knowledge of asymptotical statistical properties of the systems, so called "false alarms" - wrongly identified bifurcations - may arise when considering finite dataset and model outputs [3].
In this context it is highly desirable to build early warnings indicators which provide themselves some estimations of dynamical and geometrical properties of the system in the asymptotic regime so that false alarms can be discriminated from real early warnings. The main achievement of this paper is to introduce a new indicator of criticality based on the extreme value analysis which features the possibility to explore the asymptotic properties of the system and connect the fluctuations of maxima and minima of physical observables to the ongoing critical transition. Hereby we exploit the theoretical results obtained so far for extreme values of dynamical systems to study the behaviour of a prototypical stochastic model in the proximity of a tipping point, showing analytically and numerically the capabilities of the new indicators.
Although our aim is to provide a general tool to analyse tipping points, the motivation for the research presented here originates mainly from the analysis of the data series of turbulent energy in the so-called plane Couette flow, i.e. the flow of a viscous fluid between two parallel plates, one of which is moving relative to the other [5]. In this system the control parameter is the Reynolds number whose value determines either turbulent or laminar flow regimes. We observed that, when the Reynolds number approaches the threshold at which the turbulence decays, the probability of observing a very low minimum of turbulent energy increases. As a consequence, maxima and minima exhibit an asymmetric behaviour that can be quantified by applying the tool provided by the extreme value theory. This experimental result has been the starting point to devise a heuristic approach for more general scenarios, which is the subject of this paper. The results on the ongoing analysis of the breakdown of turbulence in the plane Couette flow will be reported in forthcoming papers.

The distribution of the largest or smallest values of certain observables has been widely studied as it is of great interest in many practical applications (e.g. the analysis of environmental extreme events). In the past years, this interest has led to a fast development of a so-called extreme events theory [6–9]. The traditional approach

*davide.faranda@zmaw.de

introduced by Gnedenko [6] is based on the analysis of extremes obtained dividing the considered dataset into bins of fixed length and choosing the maximum for each bin. There, statistical inference is performed on such a reduced dataset by considering as model the Generalized Extreme Value (GEV) distribution family, which includes, as members, the Weibull, Gumbel and Frechet distributions, which greatly differ in terms of mathematical properties. Besides the block maxima approach, the so called peak over threshold is also widely used to tackle the problem of extremes. The extreme values are selected as exceedances over a certain threshold, fitted to the Generalized Pareto (GP) distribution. The asymptotic convergence to the GP model is then guaranteed by the Pickands-Balkema-de Haan theorem [7, 10].

Whereas the theory has been originally designed for the study of extremes for series of independent and identically distributed (i.i.d.) variables, in the last decade the existence of asymptotic laws has been proven for maxima of observable computed on the orbits of dynamical systems . From the first rigourous paper by Collet [11], in the period of a few years relevant new results have been obtained [12–17]. The starting point of all these investigations has been associating to the stationary stochastic process given by the dynamical system a new stationary independent sequence distributed according to the GEV model, and then pulling back the obtained statistical laws to the original sequence extracted from the dynamical system. The assumptions necessary to observe a GEV distribution in dynamical systems rely on the choice of suitable observables and the fulfilment of particular mixing conditions. These results can also be used to study extremes of stochastically perturbed dynamical system as recently shown in [18, 19].

In order to construct robust indicators of criticality, able to discern when the underlying system is close to a bifurcation, from the extreme value analysis we will take two complementary points of view:

1. The classical approach of looking at the autocorrelation properties of the time series of the considered variable is replaced with the investigation of the appropriateness of GEV for the extremes of the time series: the presence of long time-correlations in in antithesis with the fulfilment of the mixing conditions mentioned above.

2. The search for changes in the skewness of the bulk statistics is substituted with the analysis of whether parametric modulations to the system lead to changes in the qualitative properties of the extremes, i.e. whether we observe transitions among the three kinds of distributions included in the GEV family.

The capabilities of these new indicators will be analysed for probably the simplest conceivable dynamical model featuring multi-stability, i.e. the overdamped 1D motion of a stochastically forced particle in a double well potential. The choice of this model naturally relies on the fact that, even for complex and chaotic dynamical systems which feature tipping points, a reduction to the Langevin model is often attempted as it opens the way to approaching the problem in terms of the one-dimensional Fokker-Planck equation [20], even if one must handle with care the construction of a surrogate, stochastic dynamics, as discussed in [21] regarding a relevant example of geophysical relevance. Moreover this model can be easily explored numerically and analytically. In future papers we will test the indicators in complex systems arising from fluid dynamics.

This work is organised as follows: in Section 2 we present some basic notions of extreme value theory for independent and identically distributed variables and dynamical systems. In Section 3 we introduce the new indicators and describe their properties. Section 3 is devoted to explaining the theoretical and numerical results for the motion of a particle in an asymmetric double well potential under the effect of noise. Finally in Section 4 we review the results obtained for the system analysed and propose an algorithm to extend the analysis to other dynamical systems or time series drawing our conclusions and more suggestions for further work.

II. ELEMENTS OF EXTREME VALUE THEORY FOR DYNAMICAL SYSTEMS

A. Traditional Extreme Value Theory

Gnedenko [6] studied the convergence of maxima of i.i.d. variables

$$X_0, X_1, ..X_{m-1}$$

with cumulative distribution (cdf) $F(x)$ of the form:

$$F(x) = P\{a_m(M_m - b_m) \leq x\}$$

Where a_m and b_m are normalizing sequences and $M_m = \max\{X_0, X_1, ..., X_{m-1}\}$. Under general hypothesis on the nature of the parent distribution of data, it has been shown that the distribution of maxima, up to an affine change of variables, converges in the limit for $m \to \infty$ to a single family of generalized distribution called GEV distribution with probability density function:

$$h_G(x; \mu, \alpha, \kappa) = \frac{1}{\alpha}\, t(x)^{\kappa+1} e^{-t(x)} \tag{1}$$

where

$$t(x) = \begin{cases} \left(1 + \kappa(\frac{x-\mu}{\alpha})\right)^{-1/\kappa} & \text{if } \kappa \neq 0 \\ e^{-(x-\mu)/\alpha} & \text{if } \kappa = 0 \end{cases} \tag{2}$$

which holds for $1+\kappa(x-\mu)/\alpha > 0$, using $\mu \in \mathbb{R}$ (location parameter) and $\alpha > 0$ (scale parameter) as scaling constants in place of b_m, and a_m [7, 19]. $\kappa \in \mathbb{R}$ is the shape parameter also called the tail index and discriminate among the classical EVLs: when $\kappa \to 0$, the distribution corresponds to a Gumbel type (Type 1 distribution). When the index is positive, it corresponds to a Fréchet (Type 2 distribution); when the index is negative, it corresponds to a Weibull (Type 3 distribution). The type of distribution observed is very important as it discriminates the kind of tail decay of the parent distribution. A Gumbel distribution is typically observed when the parent distribution features an exponential tail, whereas Fréchet and Weibull laws are typical associated to power law tails: a Weibull is usually observed when the tail is limited above by a certain threshold whereas a Fréchet when the tail is limited below. In order to compare the properties of maxima and minima distributions, they should both be threaten as they correspond to right or left tails of the parent distribution. In the forthcoming analysis, we will always change sign to the minima distribution [22].
Gnedenko's results are related to a precise way of selecting extremes, the so called block maxima approach: it consists of dividing the data series of length k of some observable into n bins each containing the same number m of observations, and selecting the maximum (or the minimum) value in each of them [22]. The extremes are then fitted to the GEV distribution. In Section 4.A we will give a detailed description of the inference method used to extract the parameters of the EVLs.

B. Extremes of Dynamical Observables

Let us consider a dynamical system $(\Omega, \mathcal{B}, \nu, f)$, where Ω is the invariant set in some manifold, usually $\mathbb{R}^d$, $\mathcal{B}$ is the Borel σ-algebra, $f : \Omega \to \Omega$ is a measurable map and ν a probability f-invariant Borel measure.
In order to adapt the extreme value theory to dynamical systems, we will consider the stationary stochastic process $X_0, X_1, ...$ given by:

$$X_m(x) = g(\mathrm{dist}(f^m(x), \zeta)) \qquad \forall m \in \mathbb{N} \tag{3}$$

where 'dist' is a distance on the ambient space Ω, ζ is a given point and g is an observable function. As we said above, we will use three kinds of observables $g_i, i = 1, 2, 3$, defined as:

$$g_1(x) = -\log(r) \quad g_2(x) = r^{-\beta} \quad g_3(x) = C - r^{\beta} \tag{4}$$

where $r = \mathrm{dist}(x, \zeta)$, C is a constant and $\beta > 0 \in \mathbb{R}$.
Using these observables we can obtain convergence to the type 1,2,3 distribution if one can prove two sufficient conditions called D_2 and D' that the dynamical system obeys. These conditions require the presence of a sufficiently fast decays of correlation for the stochastic dynamical sequence and limit the possibility of having clustered extremes. Another way to approach the problem of extremes for the g_i observables relies on studying the statistics of first return and hitting times, which provide information on how fast the point starting from a certain initial conditions return in its neighbourhood, see the papers by [13] and [23]. They showed in particular that for dynamical systems preserving an absolutely continuous invariant measure or a singular continuous invariant measure ν, the existence of an exponential hitting time statistics on balls around ν almost any point ζ implies the existence of extreme value laws for one of the observables of type $g_i, i = 1, 2, 3$ described above. The converse is also true, namely if we have an extreme value law which applies to the observables of type $g_i, i = 1, 2, 3$ achieving a maximum at ζ, then we have exponential hitting time statistics to balls with center ζ. Recently these results have been generalized to local returns around balls centred at periodic points [15].
Since it is difficult to check the exponential decay of hitting time statistics as recurrences are hard to tackle analytically or numerically in more than three dimensions, it is highly desirable to connect the theory explained above to a more straightforward dynamical indicator computable for a wide class of observables as well as in high dimensional dynamical systems. A key observation is that condition D_2, introduced in its actual form by Freitas-Freitas [12], could be checked directly by estimating the rate of decay of correlations for Hölder observables. Starting from this observation, Aytaç et al. proved the existence of EVLs for dynamical systems whose correlations decay is at least summable [18]. In the following discussion, instead of checking on the mixing conditions D_2 and D', we will refer exclusively to the results described in [18] since correlations or autocorrelations functions are well known tools in the study of dynamical systems properties and they are extensively used to model critical transitions.
In [24] and [25] we have verified the convergence to the classical EVL and the relations wit the GEV family from both an analytical and numerical point of view in a wide class of mixing low dimensional maps. Using g_i observable functions we have shown that GEV parameters in mixing maps can be written in terms of m (or equivalently n) and the local dimension of the attractor (ζ):
For g_1 type observable:

$$\alpha = \frac{1}{\delta(\zeta)} \qquad \mu \sim \frac{1}{\delta(\zeta)} \ln(m) \qquad \kappa = 0 \tag{5}$$

For g_2 type observable:

$$\alpha \sim n^{-\frac{\beta}{\delta(\zeta)}} \qquad \mu \sim n^{-\frac{\beta}{\delta(\zeta)}} \qquad \kappa = \frac{\beta}{\delta(\zeta)} \tag{6}$$

For g_3 type observable:

$$\alpha \sim n^{\frac{\beta}{\delta(\zeta)}} \qquad \mu = C \qquad \kappa = -\frac{\beta}{\delta(\zeta)} \tag{7}$$

In [26] we have clearly shown that a GEV distribution can be fitted only if the mixing conditions are fulfilled whereas other kind of distributions not belonging to the GEV family are observed for quasi-periodic and periodic motions. The theory has been extended to stochastically perturbed dynamical systems in [18]: in particular, systems whose deterministic dynamics is governed by a finite number of fixed points possess EVLs when perturbed with additive noise. In these cases, the extreme value parameters depend on the phase space dimension instead on the local dimension as the noise help the perturbed system to explore the neighbour of a fixed point in a ball which scale exactly with the dimension of the phase space. In [19], the authors define some guidelines to effectively observe such EVLs in numerical simulations, pointing out the range of noise to be used. These results will be particularly useful for describing the behaviour of the stochastic differential equation whose deterministic part is driven by a double well potential dynamics which will be extensively analysed in Section 4.
In the next section we will describe what we expect to happen in a general case. The notations will refer to a deterministic dynamical systems which feature at least two disjoint attractors but - under the conditions previously described - the conditions are valid also in the stochastically perturbed case.

III. TIPPING POINTS AND EXTREMES

Let us suppose that $f : \Omega \to \Omega$ is a measurable map with ν (f-invariant Borel measure) its physical measure. In fact, since a dynamical systems may have several different invariant measures, we will always refer to the physical measure as the one arising naturally in any numerical simulations. We will also assume that for a given initial condition x_0, the physical measure $\nu(x_0)$ is unique but it is not the same for all the initial conditions in the phase space, i.e. there are at least two disjoint attractors. Let us perturb the dynamics introducing a parameter λ such that $f = f(x_0, \lambda)$ and $\nu = \nu(x_0, \lambda)$. We define λ_c as the critical value of λ such that if $0 \leq \lambda < \lambda_c$ varies smoothly, the same happens also to the physical measure $\nu(x_0, \lambda)$ but whenever $\lambda \geq \lambda_c$ the system undergoes a bifurcation such that the corresponding physical measure cannot be smoothly obtained by $\nu(x_0, \lambda < \lambda_c)$ and corresponds to another disjoint attractor of the phase space. For $f(\lambda < \lambda_c)$, the system has a summable decay of correlations and therefore the block maxima of the observables of the form $g_i(\cdot)$ asymptotically obeys one of the Extreme value laws described above. We have already mentioned in the introduction that the behavior of the system changes in the limit $\lambda \to \lambda_c$ as the correlations in the system increases [2, 27]. Therefore, if the decay of correlations gets slower and slower the order of extreme statistics - in practise, the minimum length of our experimental bin from which we extract the maximum - needed to observe convergence to the predicted EVLs becomes higher. In the limit, the bin length needed to decorrelate the data tends to infinity and this prevents from obtaining EVLs. At finite time, this effect gives rise to increasing deviations from the theoretical behaviour that can be explored numerically whenever the asymptotic expected values for μ, α and κ are known. This condition is easily met when dealing with systems whose physical measure is absolutely continuous with respect to Lebesgue. In this case the EVL parameters given in Eqs. 5-7 only depend on the phase space dimension, since the physical measure of the system is absolutely continuous. For stochastically perturbed dynamical systems this is also true provided that we choose a suitable noise amplitude as explained in [19]. Moreover, once found the bin length needed to obtain the asymptotic results for λ such that in the perturbed system we do not observe any critical transition, we can fix it to study the convergence to EVLs when approaching λ_c.

In high dimensional systems, the theoretical framework described above is still valid even though it may be difficult to perform numerical simulations and control the variables in the phase space. In this case, critical transitions are highlighted by specific physical observables that experience abrupt changes when the system crosses a tipping point. These observables undergo greater amplitude deviations in the direction of the state they are destined to shift to, than in the opposite direction, often showing an increase in the skewness of the distribution. Recently, the skewness increase near a tipping point has been proposed as an early warning indicator by Guttal [28]. However, the method does not allow for having a quantitative estimation of the threshold value to be recognized as λ_c. An extreme value analysis on the maxima and on the minima of the distribution is capable to overcome this problem and highlight the critical transition by providing a quantitative way to determine the value of λ_c. Let us call the physical observable ϕ, for $\lambda < \lambda_c$, the extreme fluctuations of the observable are bounded in the direction of the minima and of the maxima giving rise to a Weibull distribution typically observed when fluctuations are bounded by an upper threshold. Instead, when $\lambda \to \lambda_c$ the fluctuations feel the presence of the other attractor in the direction of the state it is destined to shift to and this is reflected by a change in the extreme distributions from a Weibull one to a Gumbel one. This mechanism also allow to understand in which direction the transition will take place if the goal is to use the early warnings in a predicting way.

The critical indicators introduced so far using the EVT results can be divided into two classes: *dynamical indicators* based on the divergence of the classical

EVLs in the proximity of tipping points for the maxima of $g_i \quad i = 1, 2, 3$ and *physical indicators* based on the changes in the distributions of maxima and "reversed minima" of a certain relevant physical observable ϕ near the critical transitions. In the next section we will show an example in which the behaviour for the critical indicators introduced in this section can be derived analytically and observed numerically, defining guideline for the application in other systems in Section 5.

IV. A CASE STUDY

The model considered in the following analysis is widely used for modelling several natural phenomena featuring a bistable behaviour. It consists on the following stochastic differential equation (SDE):

$$dx = -V'(x)dt + \epsilon dW, \tag{8}$$

under the effect of the potential $V(x) = \frac{1}{4}x^4 - ax^2 + cx$ with $a, c > 0$. The potential is represented for some specific values of the parameters a and c in Fig. 1. The stochastic forcing consists of a Wiener process W whose amplitude is modulated by the parameter $\epsilon > 0$. By letting $\epsilon = 0$ the system is completely deterministic and features two stable fixed points ($\bar{x_1} < 0$ and $\bar{x_2} > 0$) and one unstable fixed point. It is trivial to check that extremes of any observables extracted from the deterministic dynamics do not obey any of the classical EVLs. Instead, as soon as noise is switched on by setting a non-zero stochastic forcing, the physical measure becomes absolutely continuous with respect to Lebesgue so that, starting from any initial conditions x_0, asymptotic EVL exists at any point ζ. In fact, when c is small, the probability distribution for the variable x can be computed by solving the Fokker Planck equation and considering a steady solution [29]. The result is obeyed in general and it is known as the Boltzman factor:

$$\rho^{\epsilon}(x; x_0 \simeq \bar{x_2}) = C^{\epsilon} \exp\left\{-\frac{2V(x)}{\epsilon}\right\} \tag{9}$$

When also the effect of the linear term is considered, the behaviour of Eq. 8 can be understood by computing the mean exit time for the basin of attraction of $\bar{x_2} > 0$ as in [30]:

$$\langle \tau(\bar{x_2}) \rangle = \frac{\pi}{2^{3/2}a} \exp\left(\frac{2a^2}{\epsilon^2}\left(1 - \frac{4c}{(2a)^{3/2}}\right)\right). \tag{10}$$

while the mean exit time of the basin of attraction of x_1 is obtained by changing c to $-c$.
For $c = 0$ the mean exit time from the negative and positive basin are the same. Instead for $c > 0$, fixing the initial condition in $x_0 = \bar{x_2}$, we notice immediately that the mean exit time decrease sensibly when $\frac{4c}{(2a)^{3/2}} \to 1$. Once fixed $0 < \epsilon \ll a$, the average exit time can be controlled by modulating c.

Let us now analyse the behaviour of the *dynamical indicators* described in the previous section when c is changed. When Eq. 9 holds, the indicators converge towards the EVL as described in Eqs. 5-7 with $\delta(\zeta) = 1$ for all ζ since it has been proven for the double well symmetric potential that the correlation decay is a stretched exponential (see Eqs. 47,48 in [31]) when noise intensity is low and a local equilibrium approximation is used. As pointed out in the previous section, such a decay of correlations is enough to prove that g_i extremes have classical EVLs as proven by [18]. Instead, when $\langle \tau(\bar{x_2}) \rangle \to 0$, the correlations decay becomes slower and slower (compare Fig. 4 and Fig. 5 in [31]). Extremes may still posses a classical EVLs but it becomes virtually impossible to observe in a finite time series: this points at the importance of looking not only at the asymptotic statistics of processes, but also at the pace at which the asymptotic behaviour is obtained.

As *physical indicator* we simply examine the maxima and reversed minima distribution of $\phi(x) = x$ which here represents the position of the particle in the right well of the potential $V(X)$ being the initial condition $x_0 = \bar{x_2}$. We remind that in a general case the analysis can be carried out analysing the series of any relevant observables of the systems whose expectation value is different in the two basins of attraction. When $c \to 0, \epsilon \ll a$, an harmonic expansion of the potential around $\bar{x_2}$ holds and the particle is confined in both the direction of the minima and the maxima. In terms of extremes distributions we observe in both cases a Weibull type for the reasons described in the previous section. For increasing c the harmonic expansion of the potential does not hold any more and we have to consider the full expression of $V(x)$: the maxima are still bounded by the quartic term but, towards the minima, the particle "feels" the presence of the other well experiencing a change of the distribution from Weibull to a Gumbel type.

A. Numerical experiments

The numerical experiments, performed on the one dimensional SDE described in Eq. 8, are devised to follow exactly the theoretical set-up described in the previous sections. An ensemble of 100 realisations of the systems is produced starting from $x_0 = \bar{x_2}(c = 0)$, once set the noise amplitude $\epsilon = 0.5$, $dt = 0.1$ under the potential $V(x) = 1/2x^4 + 1/2x^2 + cx$ represented in Fig. 1. For each realisation, $n = 1000$ extremes are selected with two different bin length: $m = 1000, m = 2000$ - for a total number of iterations $s = 10^6$ and $s = 2 \cdot 10^6$

respectively - in order to show that the method work independently of the choice of m provided that is large enough to guarantee convergence when $c = 0$. Approaching the critical transition, whenever the particle falls in the left well we interrupt the realisation reinitialising the initial condition x_0 counting, among all the realisations, the number of transitions experienced which provides itself an indication on how likely a transition will be experienced during the time interval considered. Even if we expect our indicators to be effective as soon as transitions are observed, the reinitialisation helps in studying the behaviour of the indicators in the interval of c values for which the jump to the other basin of attraction is probable but not certain. This may help whenever only a realisation of the system is considered or when dealing with experimental data set.

The inference procedure follows [32, 33] where the authors have used a L-moments estimation as it provides reliable values for the GEV parameters even when the cdf corresponds to fractal or multi fractal measures [34]. In fact, since the nature of the extremes distributions is unknown when approaching the tipping points, this procedure is more reliable than the Maximum Likelihood Estimation (MLE), which is well defined only if the underlying physical measure is absolutely continuous, as described in [24]. Nonetheless, we tested both the inference methods founding comparable results. The L-moments inference procedure does not provide any confidence intervals unless these are derived with a bootstrap procedure which is also dependent on the data sample size [35]. The MLE, on the other side, allows for easily compute the confidence intervals with analytical formulas [36] so it may be preferred if only a realisation of the system is available.

Let us first analyse what happen for the *dynamical indicators* for which results are shown in Fig. 2 where the shape parameters κ for the g_i observable are displayed against the values of c. The left plots correspond to the case $m = 10^3$, the right ones to the case $m = 10^4$, $\beta = 1/3$. Substituting the value of the parameters chosen for the simulation in Eq.10, the exponential term vanishes when $c \to 0.25 = c_{crit}$. Nonetheless, the value of c for which the system experiences at least one transition from one basin of attraction to another and a reinitialisation from the initial condition is needed is about $c_{crit} \simeq 0.21$ in both the experiments. The difference between the theoretical and experimental value of c_{crit} is due to the fact that the noise has a finite size. The divergence is far more evident for the observable g_1 and g_3 whereas it seems slower for the g_2. This is due to the different weight assigned by g_2 to the minimum distances: in fact, with respect to the observables g_1 and g_3, g_2 weights more the points which come closer to ζ. Since spurious extremes are located relatively far away from ζ, they do not contribute sensibly to the divergence of the shape parameter of g_2. This asymmetry with respect to the other observables may be removed by choosing a smaller β so that farther extremes are weighted more. Repeating the experiment setting $\beta = 1/10$ for g_2 we obtain an estimation of c_crit consistent with the other observables. Repeating the experiment setting $\beta = 1/10$ for g_2 we obtain an estimation of c_crit consistent with the other observables. a way As expected from the theory, the indicators for the three observable behave in a similar way and they diverge approaching c_{crit}. The divergent behaviour appears at lower values of c in the case $m = 1000$ with respect to the other case. In fact, by stretching the bin length the problem of lower decay of correlations appears at higher c than in the case $m = 1000$ as we improve the convergence by selecting more "authentic" extremes. Encouragingly, also in the $m = 2000$ experiment, the critical transition can be highlighted well before the probability of jumping to the other attracting states is higher than the 20%. From this example it is clear that is impossible to univocally define a threshold that indicates the tipping point. This is an intrinsic problem of the critical transitions in systems driven by the noise and it is directly related to the time scales on which we observe our system. Therefore the divergence obtained by the use of different g_i should be critically analysed to detect the approaching tipping point. Instead, in a modelling framework, one can test the system's behaviour with changing levels of noise. As discussed in [21] when looking at average transitions rates, in a non 1D dimensional system, one should expect the same scaling properties of c_{crit} and ϵ is the same as in the simple 1D model analysed here.

The results for the *physical indicators* are shown in Fig.3. The upper plots represent the shape parameter for the maxima of $\phi(x) = x$, the middle plot the same but for the reversed minima whereas the lower plot is the same of Fig. 1 repeated here for convenience. The set-up is exactly the same as for the *dynamical indicators*. In this case, the MLE procedure always succeed and what we observe is a change in the type of distribution only in the direction of the minima. Remarkably, the transition happens at different c values for the two different cases. This is explainable by observing that an increase the bin length, the probability of observing a very low minimum within a single bin increases and this reflects in the overall distribution. In the case $m = 2000$ we observe even a Fréchet law but for the values of the system for which the system has experienced a critical shift at least in the 20% of the realisations. In this case it seems indeed possible to define c_{crit} as the value for which the GEV distribution of the minima or of the maxima changes sign. For more CPU demanding experiments or for experimental datasets an ensemble of realisations is usually unavailable. In these cases the switching between different types of distribution for only the minima - or the maxima - may be still interpreted as a signal of an approaching tipping point using confidence intervals

as uncertainty for the parameters instead of the standard deviation of the sample. This is exactly what we did for analysing the data of the turbulent energy in the plane Couette flow and will be reported elsewhere.

V. FINAL REMARKS

This paper has addressed the problem of using extremes statistics to define robust indicators for approaching tipping points of dynamical systems. The indicators have been grouped into two categories: *dynamical indicators* and *physical indicators*. For the formers, by knowing some properties of the physical measure it has been possible to identify the asymptotic EVLs when the system is well away from bifurcation points and thereby detect approaching tipping points by the divergence from the expected EVLs. Similarly, for the *physical indicators* built up using extremes of relevant observables of the system, a comparison between maxima and reversed minima asymptotic EVL parameters provide a straightforward way to identify critical transitions. From the numerical results, described in the previous section, we suggest some guidelines to implement the indicators presented in an algorithmic way as described in Fig.4:

- The *dynamical indicators* are mainly devised for applications in low dimensional systems as the trajectories must be explicitly computed. For deterministic dynamical systems the local dimension should be known at the point ζ whereas for stochastic dynamical systems the asymptotic EVLs depend only on the phase space dimension. The algorithm, represented schematically in the upper panel of Fig. 4, begins by setting $\lambda = 0$ and the computation of the length of the series s needed to obtain theoretical parameters consistent with the asymptotic EVLs. Once s is determined, λ is increased and the fitting procedure repeated until the GEV parameters diverge from the theoretical expected one. The critical λ may be recognized when the experimental parameters are not consistent with the theoretical ones.

- The *physical indicators* algorithm is described in the bottom panel of figure 4. It can be applied to a series of observables $\phi(x)$ originating from dynamical systems or to an experimental dataset. The first step, as in the previous case, is to chose a suitable s in order to obtain a Weibull law for maxima and reversed minima of $\phi(x)$ in the unperturbed system. Once the length of the data series s is fixed, the critical λ is the one for which the shape parameter of the distribution either of the minima or of the maxima changes sign - provided that the other remains negative.

The extreme values indicators present some noticeable advantages with respect to the indicators commonly used to highlight critical transitions: the knowledge of the asymptotic EVLs provide itself a robust indication of the order statistics needed to profit from an extreme value analysis. In other terms, we have the exact statistical model to conform to (under given conditions on the dynamical system's properties), and we can correctly infer that if the statistics does not obey the GEV model, then the conditions are not obeyed. This gives a much stronger mathematical framework to our analysis than in most previous investigations. In many applications this information is not available and this cause the identification of early warning signal as false tipping points. For the *physical indicators*, the interesting intuition that the skewness of the distribution changes when approaching the transition [28] can be related to modifications in the EVLs quantifiable with a change of sign of the shape parameters corresponding to a different type of extreme value distribution. Clearly, results will improve choosing observables whose values change more in percentage crossing the tipping point. The strict connection between the *dynamical indicators* and the recurrences in a point suggest that we are able to highlight critical transitions only looking at the behaviour of the systems in the neighbour of a specific point ζ which can be located in any point of the attractor. This could help in all the situations for which the dynamic may be better represented or analysed in a sub-domain of the phase space. Moreover these indicators themselves provide interesting information on events which happen with very small probability but that are usually relevant in ecosystems, climate and financial models. Certainly the methods based on extreme value statistics require a great availability of good quality data and/or the possibility to perform time consuming simulations capable to extract authentic extremes. This problem is especially relevant for applications in climate science and in finance since the models involved are very demanding in terms of CPU time and long-term data availability is poor. However, such data will be more and more available in the next future and not only for simplified models.

The numerical tests, carried out on a SDE representing a material point in a double well potential, exploit the possibility of operatively using the method described in practical applications. Although the results meet the theoretical set-up and provide a net description of the tipping point, it is indeed evident that other tests should be carried out to assess the general validity of this approach. Since the statistical tools used to study extreme values are commonly distributed with scientific software, the algorithm can be easily checked and compared to other methods. Our aim is to test the indicator on complex systems arising from fluid dynamical studies which feature tipping points whose nature remain not well understood and on other theoretical low dimensional models. One must bear in mind that, as discussed in [21],

great caution must be paid when trying to extend considerations valid for 1D models onto higher dimensional systems: the operation of defining a 1D effective projected dynamics is far from being trivial.

ACKNOWLEDGMENTS

We wish to acknowledge various useful exchanges with J. Freitas, A.M. Freitas, S. Vaienti and P. Manneville.The research leading to these results has received funding from the European Research Council under the European Community Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement No. 257106. We regret that on October 22nd 2012 an Italian court has sentenced seven scientists - Barberi, Boschi, Calvi, Dolce, De Bernardinis, Eva and Selvaggi - to six years in prison for non having been able to predict the L'Aquila earthquake in 2009. This sentence seems to have no basis whatsoever on the current scientific knowledge. Instead, we wish to acknowledge them for their efforts and contributions in the assessment of risk factors linked to geophysical phenomena.

[1] C. Kuehn, Physica D: Nonlinear Phenomena(2011).
[2] M. Scheffer, J. Bascompte, W. Brock, V. Brovkin, S. Carpenter, V. Dakos, H. Held, E. Van Nes, M. Rietkerk, and G. Sugihara, Nature **461**, 53 (2009).
[3] T. Lenton, Nature Climate Change **1**, 201 (2011).
[4] L. Dai, D. Vorselen, K. Korolev, and J. Gore, Science **336**, 1175 (2012).
[5] P. Manneville, Pramana **70**, 1009 (2008).
[6] B. Gnedenko, The Annals of Mathematics **44**, 423 (1943).
[7] J. Pickands III, the Annals of Statistics, 119(1975), ISSN 0090-5364.
[8] M. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes and related properties of random sequences and processes.* (Springer, New York, 1983).
[9] M. Ghil, P. Yiou, S. Hallegatte, B. Malamud, P. Naveau, A. Soloviev, P. Friederichs, V. Keilis-Borok, D. Kondrashov, V. Kossobokov, *et al.*, Nonlin. Processes Geophys **18**, 295 (2011).
[10] A. Balkema and L. De Haan, The Annals of Probability, 792(1974).
[11] P. Collet, Ergodic Theory and Dynamical Systems **21**, 401 (2001), ISSN 0143-3857.
[12] A. Freitas and J. Freitas, Statistics & Probability Letters **78**, 1088 (2008), ISSN 0167-7152.
[13] A. Freitas, J. Freitas, and M. Todd, Probability Theory and Related Fields, 1(2009).
[14] C. Gupta, M. Holland, and M. Nicol, Lozi-like maps, and Lorenz-like maps, preprint(2009).
[15] A. C. M. Freitas, J. M. Freitas, and M. Todd, "Extremal index, hitting time statistics and periodicity," To appear in Adv. Math. (2012), `http://arxiv.org/abs/1008.1350`.
[16] M. Holland, R. Vitolo, P. Rabassa, A. Sterk, and H. Broer, Physica D: Nonlinear Phenomena(2011).
[17] V. Lucarini, D. Faranda, and J. Wouters, Arxiv preprint arXiv:1110.0176(2011).
[18] H. Aytaç, J. Freitas, and S. Vaienti, Arxiv preprint arXiv:1207.5188(2012).
[19] D. Faranda, J. Freitas, V. Lucarini, G. Turchetti, and S. Vaienti, Arxiv preprint arXiv:1208.5582(2012).
[20] T. Palmer and P. Williams, *Stochastic physics and climate modelling* (Cambridge University Press, 2010).
[21] V. Lucarini, D. Faranda, and M. Willeit, Nonlin. Processes Geophys **19**, 9 (2012).
[22] S. Coles, J. Heffernan, and J. Tawn, Extremes **2**, 339 (1999), ISSN 1386-1999.
[23] A. Freitas, J. Freitas, and M. Todd, Journal of Statistical Physics, 1(2011).
[24] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti, J. Stat. Phys. **145**, 1156 (2011).
[25] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti, Arxiv preprint arXiv:1106.2299(2011).
[26] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti, To appear: Int. Jou. bif. Chaos(2011).
[27] C. Wissel, Oecologia **65**, 101 (1984), ISSN 0029-8549, 10.1007/BF00384470, `http://dx.doi.org/10.1007/BF00384470`.
[28] V. Guttal and C. Jayaprakash, Ecology Letters **11**, 450 (2008).
[29] K. M. Rattray and A. J. McKane, Journal of Physics A: Mathematical and General **24**, 4375 (1991), `http://stacks.iop.org/0305-4470/24/i=18/a=023`.
[30] R. Benzi, A. Sutera, and A. Vulpiani, Journal of Physics A: mathematical and general **14**, L453 (1981).
[31] J. J. Brey, J. M. Casado, and M. Morillo, Phys. Rev. A **32**, 2893 (Nov 1985), `http://link.aps.org/doi/10.1103/PhysRevA.32.2893`.
[32] V. Lucarini, D. Faranda, G. Turchetti, and S. Vaienti, Chaos: An Interdisciplinary Journal of Nonlinear Science **22**, 023135 (2012), `http://link.aip.org/link/?CHA/22/023135/1`.
[33] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti, to appear in Int. J. of Bifurcat. Chaos(2012).
[34] J. Hosking, Journal of the Royal Statistical Society. Series B (Methodological) **52**, 105 (1990), ISSN 0035-9246.
[35] D. Burn, Hydrol Sci. J. **48**, 25 (2003).
[36] R. Royall, International Statistical Review/Revue Internationale de Statistique, 221(1986).

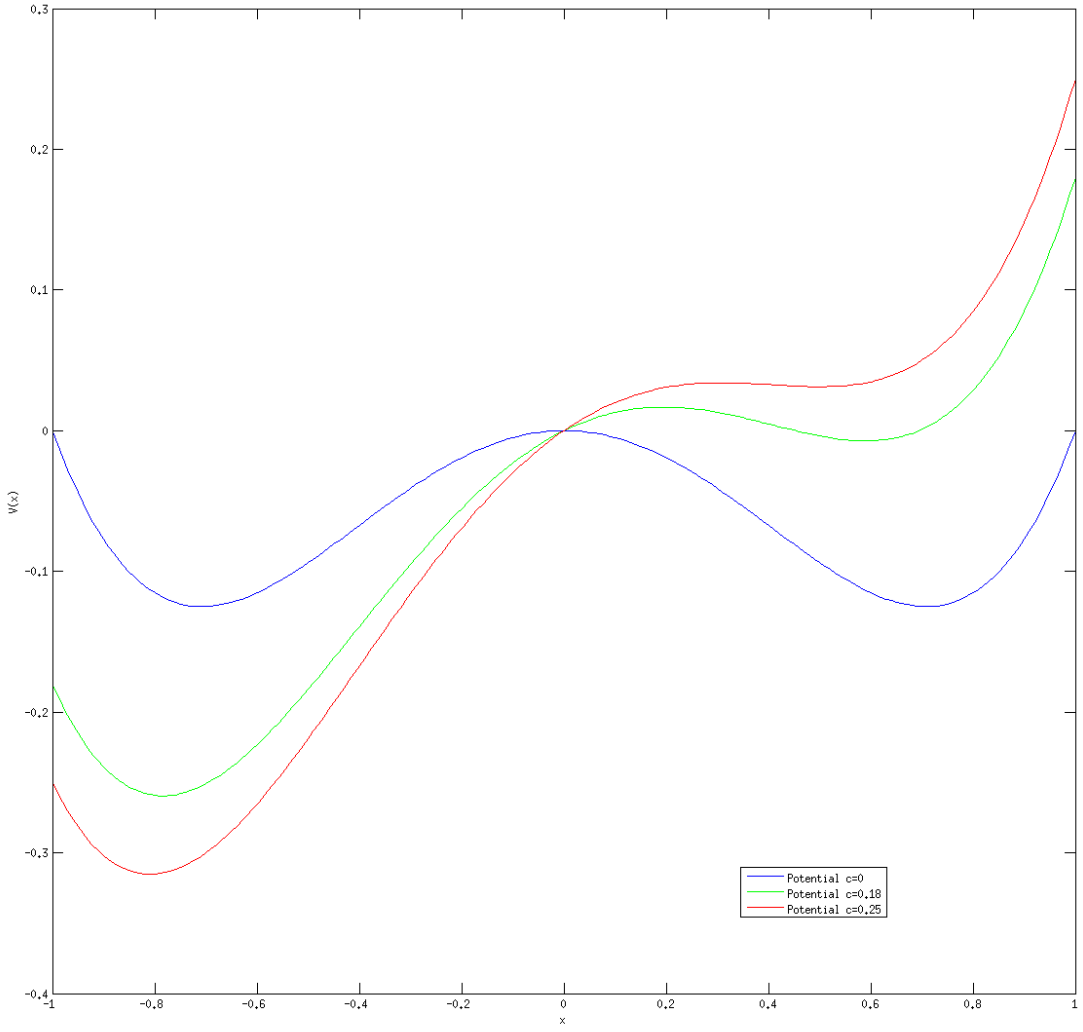


FIG. 1. Potential $V(x) = 1/2x^4 + 1/2x^2 + cx$ for different values of parameter c.

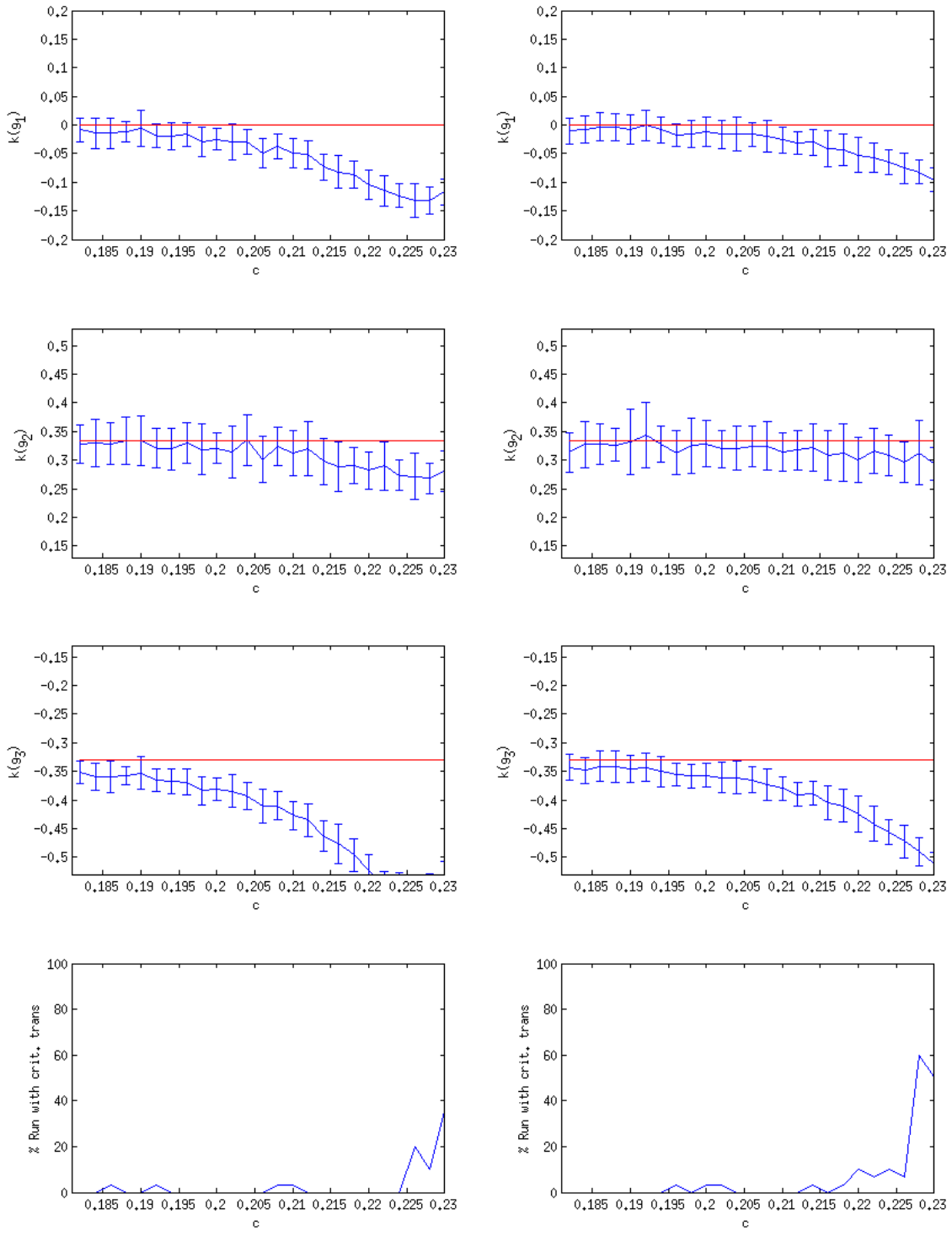


FIG. 2. Extreme value shape parameter κ vs the potential asymmetry parameter c for an ensemble of 30 trajectories starting in $x_0 = \bar{x_2}$ of the system described in Eq. 8. **(a)**: g_1 observable, **(b)**: g_2 observable with $\beta = 1/3$, **(c)**: g_3 observable with $\beta = 1/3, C = 1$, **(d)**: number of runs for which the particle jump at least once in the left% well. **Left**: $n = 1000$, $m = 1000$. **Right**: $n = 1000$, $m = 2000$. The errorbars represent a standard deviation of the sample, the red bars represent theoretical expected parameters for $c = 0$ case.

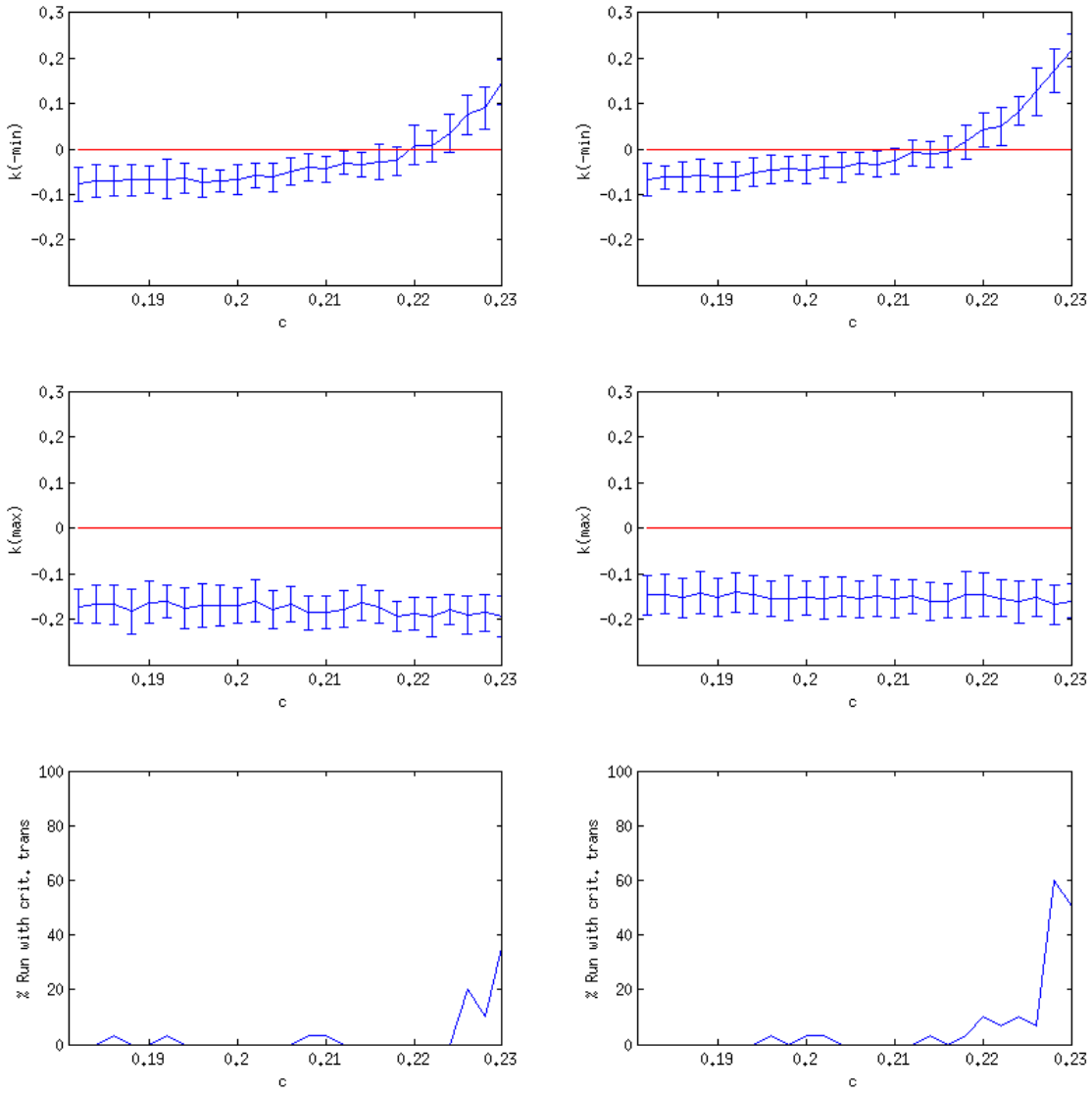


FIG. 3. Extreme value shape parameter κ vs the potential asymmetry parameter c for an ensemble of 30 trajectories starting in $x_0 = \bar{x_2}$ of the system described in Eq. 8. **(a)**: maxima for the observable $\phi(x) = x$ **(b)**: reversed minima for the observable $\phi(x) = x$; **(c)**: number of runs for which the particle jump at least once in the left% well. **Left**: $n = 1000$, $m = 1000$. **Right**: $n = 1000$, $m = 2000$. The errorbars represent a standard deviation of the sample, the red bars represent the Gumbel distribution ($\kappa = 0$).

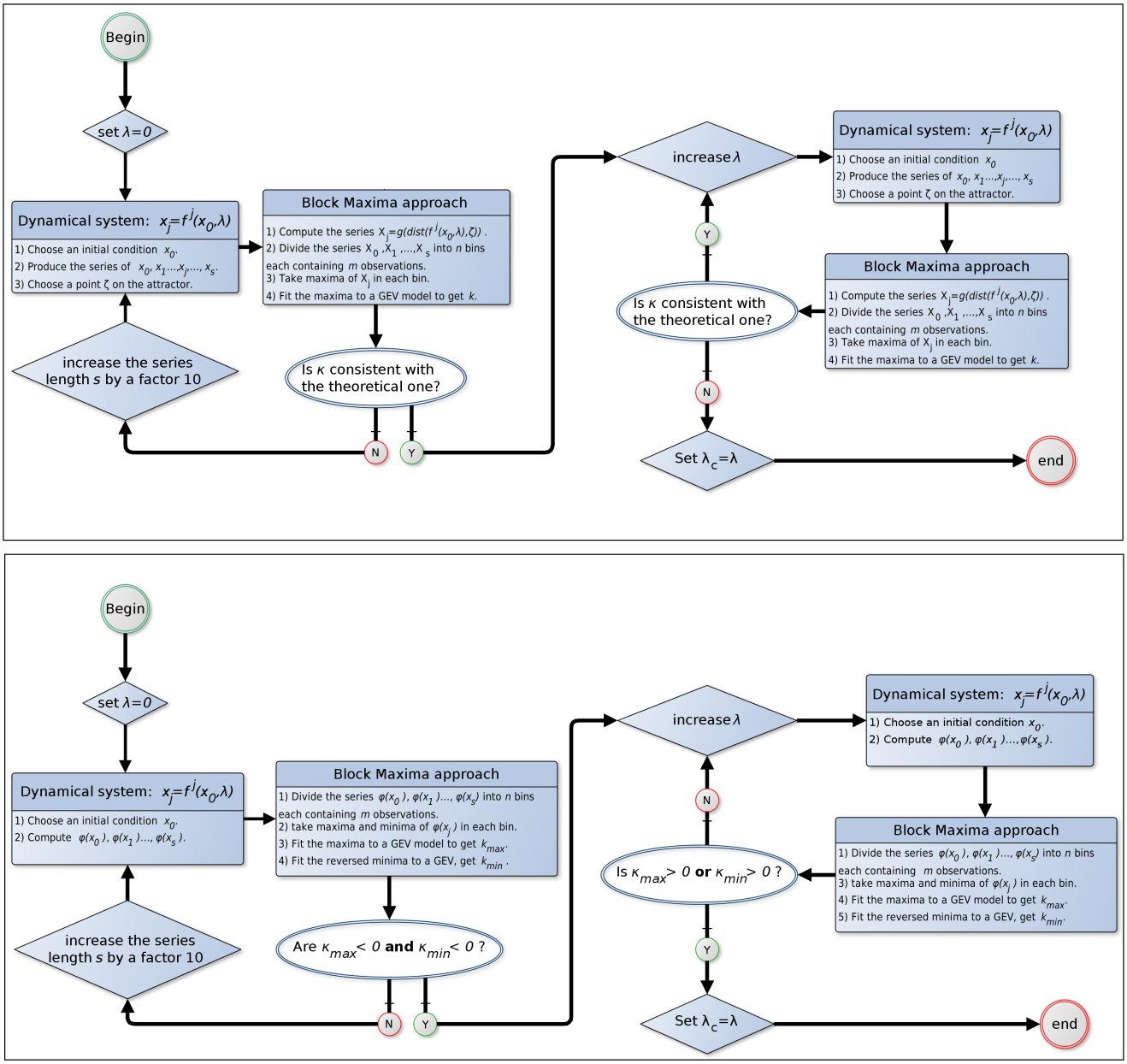


FIG. 4. Schematic representation of the numerical procedure to be used to infer the critical transition happening at $\lambda = \lambda_c$. Upper panel: dynamical indicators. Lower panel: physical indicators

Conclusions

The scientific material presented in this work of thesis highlights the possibility of using the Extreme Value Theory as a powerful tool to understand geometrical and dynamical properties of low dimensional dynamical systems. By observing for the first time the emergence of EVLs for dynamical systems, we have practically shown which order of statistics is required to observe the asymptotic laws. The investigations carried out have always supported theoretical results by the direct verifications of them through numerical experiments devised to exactly follow the proofs of the theorems. The simulations have supported conjectures that produced new theoretical advancements hard to achieve starting from scratch with a pure theoretical set-up. In particular, the possibility of extending results initially devised for absolutely continuous measures to singular continuous measures is the gateway for applying the theory in geophysically relevant systems. The material presented for the extreme value theory allows for considering the parameters related to the EVLs as reliable indicators for dynamical systems. It is clear that an intimate connection exist among all the indicators introduced through the papers. They somehow resemble musical instruments: even if any object that produces sound can serve as a musical instrument, only few of them are used to reproduce a particular orchestral partition. In a similar way, several dynamical indicators can be constructed using the information provided by the knowledge of the trajectory but only few of them provide real insight into the dynamics of the system analysed.

It is evident that the results presented are not exhaustive and have now to be tested for complex dynamical systems such as the ones currently used in the study of climate or meteorology. In this case the amount of CPU time needed for a meaningful statistics of extremes is enormous since a large number of extrema, each of them representing a proper extreme value, is required. Whereas for low dimensional systems the computational cost of producing such long orbits is affordable with the capabilities of a standard PC, when high dimensional models are used, having the computer time needed to provide sufficient data becomes a crucial issue. Appropriately choosing the physical system to analyse becomes relevant since it has to feature an interesting dynamic behaviour, coexistence of stable and unstable trajectories, but it has to produce time series long enough to permit a robust inference of the extreme value statistics. In this respect, our attention has now been focussed on the study of fluid systems that display laminar and turbulent regimes with a coexistence of both regimes under certain particular conditions controlled by the Reynolds number and more specifically on plane Couette flow [52]. We are currently carrying on experiments in this direction and we hope that the results contained in this thesis will arouse more and more interest in application to geophysical systems. In fact, even if extreme value theory is extensively used in the analysis of climate-related dataset, due to the intrinsic computational difficulties, there is still an evident lack of estimation of robust extremes from the output of climate models.

The scientific results presented in this dissertation will be part of the book "Extremes and Recurrence in Dynamical Systems", edited by Wiley and written together with Valerio Lucarini, Jeroen Wouters (U. Hamburg, Germany); Jorge Freitas, Ana Moreira Freitas (U. Oporto, Portugal); Mark Holland (U. Exeter, UK); Matthew Nicol (U. Houston, USA); Mike Todd (U. St. Andrews, UK); S. Vaienti (U. Toulon, France). The book will present a common perspective on recent advances in the theory of extreme values, arising from the investigation of dynamical systems. Hopefully it will support

further investigations in this field and in geophysically relevant systems.

Abstracts

Abstract (english)

Nowadays, it is becoming clearer and clearer that, to advance in the understanding of complex natural phenomena, a deeper understanding of the dynamical properties of complex systems is needed. The results contained in this dissertation try to give a methodological and practical way for understanding relevant properties of dynamical systems by applying and devising new dynamical indicators. These indicators mainly rely on the results obtained for the so called Extreme Value Theory that can be used to deduce general and local properties of the physical measures associated to the trajectories in the phase space.

Abstract (german)

Heutzutage wird es immer deutlicher, dass nur ein tieferes Verständnis der dynamischen Eigenschaften komplexer Systeme, das Verständnis komplexer natürlicher Phänomene verbessern kann. Die Ergebnisse dieser Disseratation geben einen methodischen und praktischen Zugang mittels der Ausarbeitung und Anwendung neuer dynamischer Indikatoren, um relevanten Eigenschaften dynamischer Systeme zu verstehen. Diese Indikatoren beziehen sich hauptsächlich auf Erkenntnisse der sogenannten Extremwerttheorie, welche benutzt werden können um allgemeine und lokale Eigenschaften physikalischer Maße der Bahnen im Phasenraum abzuleiten.

List of Publications

1. Faranda, D., Freitas, J.M., Lucarini, V., Turchetti, G. Vaienti, S. [2012]: Extreme value statistics for dynamical systems with noise. Preprint: arXiv:1208.5582

2. Faranda, D., Lucarini, V., Manneville, P. [2012] Suppression of turbulence in a plane Couette flow. Can extreme fluctuations be used to understand critical transitions? Abstract EGU2012-5458.

3. Faranda, D., Lucarini, V., Turchetti, G. Vaienti, S. [2012]: Generalized Extreme Value distribution parameters as dynamical indicator of stability. To appear: Int. J. Bif. Chaos.

4. Faranda, D., Lucarini, V., Turchetti, G. Vaienti, S. [2011]: Numerical convergence of the block-maxima approach to the Generalized Extreme Value distribution, J. Stat. Phys., 145(5),1156-1180.

5. Faranda, D., Mestre, M. & Turchetti, G. [2012]: Analysis of round off errors with reversibility test as a dynamical indicator. To appear: Int. J. Bif. Chaos.

6. Lucarini, V., Faranda, D., Turchetti, G. Vaienti, S. [2012]: Extreme Value Distribution for singular measure. Chaos 22, 023135.

7. Lucarini, V., Faranda, D., and Willeit, M. [2011]: Bistable systems with stochastic noise: virtues and limits of effective one-dimensional Langevin equations, Nonlin. Processes Geophys., 19, 9-22

8. Lucarini, V., Faranda, D., Wouters, J. [2012]: Universal behavior of extreme value statistics for selected observables of dynamical systems. J. Stat. Phys., 147(1), 63-73.

9. Lucarini, V., Kuna, T., Wouters, J., Faranda, D. [2012]: Relevance of sampling schemes in light of Ruelle's linear response theory. Nonlinearity., 25(5), 1311.

Bibliography

[1] D. Faranda, M.F. Mestre, and G. Turchetti. Analysis of round off errors with reversibility test as a dynamical indicator. *Int. J. of Bifurcat. Chaos*, 22(11), 2012.

[2] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti. Numerical convergence of the block-maxima approach to the generalized extreme value distribution. *J. Stat. Phys.*, 145(5):1156–1180, 2011.

[3] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti. Generalized extreme value distribution parameters as dynamical indicators of stability. *Int. J. of Bifurcat. Chaos*, 2012.

[4] Valerio Lucarini, Davide Faranda, Giorgio Turchetti, and Sandro Vaienti. Extreme value theory for singular measures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(2):023135, 2012.

[5] Valerio Lucarini, Davide Faranda, and Jeroen Wouters. Universal behaviour of extreme value statistics for selected observables of dynamical systems. *Journal of Statistical Physics*, 147:63–73, 2012.

[6] D. Faranda, J.M. Freitas, V. Lucarini, G. Turchetti, and S. Vaienti. Extreme value statistics for dynamical systems with noise. *Arxiv preprint arXiv:1208.5582*, 2012.

[7] V. Lucarini, D. Faranda, and M. Willeit. Bistable systems with stochastic noise: virtues and limits of effective one-dimensional langevin equations. *Nonlin. Processes Geophys*, 19:9–22, 2012.

[8] D. Faranda, V. Lucarini, and J. Wouters. Extreme value theory provides early warnings for critical transitions. *preprint*, 2012.

[9] MR Leadbetter, G. Lindgren, and H. Rootzen. *Extremes and related properties of random sequences and processes.* Springer, New York, 1983.

[10] A.C.M. Freitas, J.M. Freitas, and M. Todd. Hitting time statistics and extreme value theory. *Probability Theory and Related Fields*, pages 1–36, 2009.

[11] M. Felici, V. Lucarini, A. Speranza, and R. Vitolo. Extreme value statistics of the total energy in an intermediate complexity model of the mid-latitude atmospheric jet. part i: stationary case. *arXiv preprint physics/0601081*, 2006.

[12] C. Nicolis, V. Balakrishnan, and G. Nicolis. Extreme events in deterministic dynamical systems. *Physical review letters*, 97(21):210602, 2006.

[13] M. Scheffer, J. Bascompte, W.A. Brock, V. Brovkin, S.R. Carpenter, V. Dakos, H. Held, E.H. Van Nes, M. Rietkerk, and G. Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.

[14] C. Kuehn. A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena*, 2011.

[15] O.G.B. Sveinsson and D.C. Boes. Regional frequency analysis of extreme precipitation in northeastern colorado and fort collins flood of 1997. *Journal of Hydrologic Engineering*, 7:49, 2002.

[16] E. Brodin and C. Kluppelberg. Extreme Value Theory in Finance. *Submitted for publication: Center for Mathematical Sciences, Munich University of Technology*, 2006.

[17] M. Ghil, P. Yiou, S. Hallegatte, BD Malamud, P. Naveau, A. Soloviev, P. Friederichs, V. Keilis-Borok, D. Kondrashov, V. Kossobokov, et al.

Extreme events: dynamics, statistics and prediction. *Nonlin. Processes Geophys*, 18:295–350, 2011.

[18] F. Pauli and S. Coles. Penalized likelihood inference in extreme value analyses. *Journal of Applied Statistics*, 28(5):547–560, 2001.

[19] RA Fisher and LHC Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Proceedings of the Cambridge philosophical society*, volume 24, page 180, 1928.

[20] B. Gnedenko. Sur la distribution limite du terme maximum d'une série aléatoire. *The Annals of Mathematics*, 44(3):423–453, 1943.

[21] L. De Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Verlag, 2006.

[22] J.R.M. Hosking. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):105–124, 1990.

[23] H.W. Lilliefors. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318):399–402, 1967.

[24] J. Pickands III. Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131, 1975.

[25] A.A. Balkema and L. De Haan. Residual life time at great age. *The Annals of Probability*, pages 792–804, 1974.

[26] P. Embrechts, S.I. Resnick, and G. Samorodnitsky. Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3:30–41, 1999.

[27] P. Embrechts, CP Kluppelberg, and T. Mikosh. Modelling Extremal Events, 645 pp, 1997.

[28] E.G. Altmann, S. Hallerberg, and H. Kantz. Reactions to extreme events: Moving threshold model. *Physica A: Statistical Mechanics and its Applications*, 364:435–444, 2006.

[29] L. Rossi, G. Turchetti, and S. Vaienti. Poincaré recurrences as a tool to investigate the statistical properties of dynamical systems with integrable and mixing components. In *Journal of Physics: Conference Series*, volume 7, page 94. IOP Publishing, 2005.

[30] H. Hu, A. Rampioni, L. Rossi, G. Turchetti, and S. Vaienti. Statistics of Poincaré recurrences for maps with integrable and ergodic components. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 14:160, 2004.

[31] R. Artuso. Correlation decay and return time statistics. *Physica D: Nonlinear Phenomena*, 131(1-4):68–77, 1999.

[32] H. Furstenberg. Poincaré recurrence and number theory. *American Mathematical Society*, 5(3), 1981.

[33] J. Barrow-Green. *Poincaré and the three body problem*. Amer Mathematical Society, 1997.

[34] Z. Coelho. Asymptotic laws for symbolic dynamical systems. *Topics in Symbolic Dynamics and Applications*, pages 123–165, 2000.

[35] M. Pollicott and M. Yuri. *Dynamical systems and ergodic theory*. Cambridge Univ Pr, 1998.

[36] P. Marie. Propriétés statistiques des systèmes dynamiques déterministes et aléatoires. 2009.

[37] M. Abadi and A. Galves. Inequalities for the occurrence times of rare events in mixing processes. The state of the art. *Markov Process. Related Fields*, 7(1):97–112, 2001.

[38] V.I. Arnold and A. Avez. *Ergodic problems of classical mechanics*. Benjamin New York, 1968.

[39] M. Kupsa and Y. Lacroix. Asymptotics for hitting times. *Annals of probability*, pages 610–619, 2005.

[40] M. Hirata, B. Saussol, and S. Vaienti. Statistics of Return Times: A General Framework and New Applications. *Communications in Mathematical Physics*, 206(1):33–55, 1999.

[41] A.C.M. Freitas and J.M. Freitas. On the link between dependence and independence in extreme value theory for dynamical systems. *Statistics & Probability Letters*, 78(9):1088–1093, 2008.

[42] N. Chernov. Decay of correlations. *Scholarpedia*, pages 3(4):4862, Created: 26 August 2007, reviewed: 22 January 2008, accepted: 10 April 2008, 2008.

[43] P. Marie, G. Turchetti, S. Vaienti, and F. Zanlungo. Error distribution in randomly perturbed orbits. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19:043118, 2009.

[44] V. Baladi. Decay of correlations. In *Smooth Ergodic Theory and Its Applications: Proceedings of the AMS Summer Research Institute on Smooth Ergodic Theory and Its Applications, July 26-August 13, 1999, University of Washington, Seattle*, volume 69, page 297. Amer Mathematical Society, 1999.

[45] D. Ruelle. Resonances of chaotic dynamical systems. *Physical review letters*, 56(5):405–407, 1986.

[46] M. Pollicott. Meromorphic extensions of generalised zeta functions. *Inventiones Mathematicae*, 85(1):147–164, 1986.

[47] L.S. Young. Recurrence times and rates of mixing. *Israel Journal of Mathematics*, 110(1):153–188, 1999.

[48] B. Hasselblatt and AB Katok. *A first course in dynamics: with a panorama of recent developments*. Cambridge Univ Pr, 2003.

[49] N. Chernov and HK Zhang. Billiards with polynomial mixing rates. *Nonlinearity*, 18:1527, 2005.

[50] A.C.M. Freitas and J.M. Freitas. On the link between dependence and independence in Extreme Value Theory for Dynamical Systems. *Statistics and Probability Letters*, 78(9):1088–1093, 2008.

[51] W.L. Martinez and A.R. Martinez. *Computational statistics handbook with MATLAB*. CRC Press, 2002.

[52] P. Manneville. Understanding the sub-critical transition to turbulence in wall flows. *Pramana*, 70(6):1009–1021, 2008.