



# Extreme value laws for dynamical systems under observational noise



Davide Faranda<sup>a,\*</sup>, Sandro Vaienti<sup>b,c</sup>

<sup>a</sup> Laboratoire SPHYNX, Service de Physique de l'Etat Condensé, DSM, CEA Saclay, CNRS URA 2464, 91191 Gif-sur-Yvette, France

<sup>b</sup> Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France

<sup>c</sup> Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France

## HIGHLIGHTS

- We provide a full extreme value theory for dynamical systems perturbed with instrument-like-error.
- Numerical experiments support the theoretical findings.
- Fractal dimensions can be recovered in perturbed systems.
- The theory allows for studying recurrences on finite time series.

## ARTICLE INFO

### Article history:

Received 27 August 2013

Received in revised form

29 April 2014

Accepted 30 April 2014

Available online 10 May 2014

Communicated by I. Melbourne

### Keywords:

Generalized extreme value distribution

Observational noise

Extreme value theory

Maps

Fractal dimension

## ABSTRACT

In this paper we prove the existence of extreme value laws for dynamical systems perturbed by the instrument-like-error, also called observational noise. An orbit perturbed with observational noise mimics the behavior of an instrumentally recorded time series. Instrument characteristics – defined as precision and accuracy – act both by truncating and randomly displacing the real value of a measured observable. Here we analyze both these effects from a theoretical and a numerical point of view. First we show that classical extreme value laws can be found for orbits of dynamical systems perturbed with observational noise. Then we present numerical experiments to support the theoretical findings and give an indication of the order of magnitude of the instrumental perturbations which cause relevant deviations from the extreme value laws observed in deterministic dynamical systems. Finally, we show that the observational noise preserves the structure of the deterministic attractor. This goes against the common assumption that random transformations cause the orbits asymptotically fill the ambient space with a loss of information about the fractal structure of the attractor.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

In two previous works [1,2], we investigated the persistence of Extreme Value Laws (EVLs) whenever a dynamical system is perturbed throughout random transformations. We considered an i.i.d. stochastic process  $(\omega_k)_{k \in \mathbb{N}}$  with values in the measurable space  $Q_\varepsilon$  and with probability distribution  $\theta_\varepsilon$ . After associating to each  $\omega \in Q_\varepsilon$  a map  $T_\omega$  acting on the measurable space  $\Omega$  into itself, we considered the random orbit starting from the point  $x$  and generated by the realization  $\underline{\omega}_n = (\omega_1, \omega_2, \dots, \omega_n)$ :

$$T_{\underline{\omega}_n} := T_{\omega_n} \circ \dots \circ T_{\omega_1}(x).$$

In this setting the transformations  $T_\omega$  are taken close to each other and the suitably rescaled scalar parameter  $\varepsilon$  is the strength of such a distance. We could therefore define a Markov process  $\mathcal{X}_\varepsilon$  on  $\Omega$  with transition function

$$P(x, A) = \int_{Q_\varepsilon} \mathbf{1}_A(T_\omega(x)) d\theta_\varepsilon(\omega), \quad (1.1)$$

where  $A \in \Omega$  is a measurable set,  $x \in \Omega$  and  $\mathbf{1}_A$  is the indicator function of a set  $A$ . We recall that a probability measure  $\mu_\varepsilon$  is called a *stationary measure* if for any measurable  $A$  we have:

$$\mu_\varepsilon(A) = \int_{\Omega} P_\varepsilon(x, A) d\mu_\varepsilon(x).$$

Moreover, we call it an absolutely continuous stationary measure (*acsm*), if it has a density with respect to the Lebesgue measure whenever  $\Omega$  is a metric space.

In this work we consider a different type of perturbation, the observational noise, which consists in replacing the orbit of the

\* Corresponding author.

E-mail addresses: [davide.faranda@cea.fr](mailto:davide.faranda@cea.fr) (D. Faranda), [vaienti@cpt.univ-mrs.fr](mailto:vaienti@cpt.univ-mrs.fr) (S. Vaienti).

point  $x \in \Omega$  at time  $i$ , namely  $T^i x$ , with  $T^i x + \omega_i$ . There are several physical motivations to investigate the behavior of this kind of perturbation. In fact, as Lalley and Noble wrote in [3]:

“...In this model our observations take the form  $y_i = T^i x + \omega_i$ , where  $\omega_i$  are independent, mean zero random vectors. In contrast with the dynamical noise model (e.g.; the random transformations), the noise does not interact with the dynamics: the deterministic character of the system, and its long range dependence, are preserved beneath the noise. Due in part to this dependence, estimation in the observational noise model has not been broadly addressed by statisticians, though the model captures important features of many experimental situations.”

Judd [4], quoted in [5], also pointed out that:

“...the reality is that many physical systems are indistinguishable from deterministic systems, there is no apparent small dynamic noise, and what is often attributed as such is in fact model error.”

Moreover, a system contaminated by the observational noise raises the natural and practical question whether it would be possible to recover the original signal, in our case the deterministic orbit  $\{T^i x\}_{i \geq 1}$ . In the last years a few techniques have been proposed for such a noise reduction [6]: we remind here the remarkable Schreiber–Lalley method [7–10], which provides a very consistent algorithm to perform the noise reduction when the underlying deterministic dynamical system has strong hyperbolic properties. Another interesting work shows that in the computation of some statistical quantities, the dynamical noise corresponding to random transformations could be considered as an observational noise with the Cauchy distribution [11]. Finally, the paper [12] proves concentration inequalities for systems perturbed by observational noise.

The present work tries to re-frame the previous findings in terms of extreme value theory (EVT) by adding a further motivation driven by the applicability of the whole EVT for dynamical systems to experimental data. It should be a general concern to check the role of instrument-like-perturbations before applying dynamical systems techniques to experimental datasets. In this sense, the dynamical systems considered in this paper share several properties with observed time series, as the observational noise acts exactly as a physical instrument. The goal is to exploit the recent advancements of the EVT for dynamical systems to define in a more rigorous way the extremes of time series. A successful application of the theory presented in this paper to experimental datasets is given in [13], where temperature data are analyzed with the algorithmic procedure presented in Section 4.2. More specifically, our interest is to understand which way the results obtained on deterministic dynamical systems are altered by the addition of observational noise and in which cases one can recover classical EVLs. We start the discussion by summarizing the main findings of the EVT for dynamical systems.

The first rigorous mathematical approach to EVT in dynamical systems goes back to the pioneer paper by Collet [14]. Important contributions have successively been given in [15–17] and in [18]. Here we briefly recall the main findings deferring to the previous papers for the full demonstrations.

Let us consider a dynamical system  $(\Omega, \mathcal{B}, \nu, T)$ , where  $\Omega$  is the invariant set in some manifold, usually  $\mathbb{R}^d$ ,  $\mathcal{B}$  is the Borel  $\sigma$ -algebra,  $T : \Omega \rightarrow \Omega$  is a measurable map and  $\nu$  a probability  $T$ -invariant Borel measure.

In order to adapt the EVT to dynamical systems, we follow [15]. We consider the stationary stochastic process  $X_0, X_1, \dots$  given by:

$$X_m(x) = w(\text{dist}(T^m x, z)) \quad \forall m \in \mathbb{N}, \tag{1.2}$$

where ‘dist’ is a distance on the ambient space  $\Omega$ ,  $z$  is a given point and  $w$  is a suitable function which will be specified later. This

particular functional form has been introduced first by Collet [14] and allows for a direct connection between recurrence properties around a point of the phase space  $z$  and the existence of EVLs. The object of interest is the distribution of  $\mathbb{P}(M_m \leq u_m)$ , where  $M_m := \max\{X_0, \dots, X_{m-1}\}$ ; we have an EVL for  $M_m$  if there is a non-degenerate distribution function  $H : \mathbb{R} \rightarrow [0, 1]$  with  $H(0) = 0$  and, for every  $\tau > 0$ , there exists a sequence of levels  $u_m = u_m(\tau)$ ,  $m = 1, 2, \dots$ , such that

$$m \mathbb{P}(X_0 > u_m) \rightarrow \tau, \quad \text{as } m \rightarrow \infty, \tag{1.3}$$

and for which the following limit holds:

$$\mathbb{P}(M_m \leq u_m) \rightarrow 1 - H(\tau), \quad \text{as } m \rightarrow \infty.$$

The motivation for using a normalizing sequence  $u_m$  satisfying (1.3) comes from the case when  $X_0, X_1, \dots$  are independent and identically distributed (i.i.d.). In this setting, it is clear that  $\mathbb{P}(M_m \leq u) = (F(u))^m$ , being  $F(u)$  the cumulative distribution function for the variable  $u$ . Hence, condition (1.3) implies that

$$\mathbb{P}(M_m \leq u_m) = (1 - \mathbb{P}(X_0 > u_m))^m \sim \left(1 - \frac{\tau}{m}\right)^m \rightarrow e^{-\tau},$$

as  $m \rightarrow \infty$ . Note that in this case  $H(\tau) = 1 - e^{-\tau}$  is the standard exponential distribution function. By choosing the sequence  $u_m = u_m(y)$  as one parameter families like  $u_m = y/a_m + b_m$ , where  $y \in \mathbb{R}$  and  $a_m > 0$ , for all  $m \in \mathbb{N}$  and  $w$  as above, we have  $\mathbb{P}(a_m(M_m - b_m) \leq y) \rightarrow G(y)$  whenever the variables  $X_i$  are i.i.d., if for some constants  $a_m > 0, b_m$ . When the convergence occurs at continuity points of  $G$  ( $G$  is non-degenerate) then  $G_m$  converges to one of the three EVLs rewritable in terms of the Generalized Extreme Value (GEV) distribution as:

$$G(y; \kappa) = \exp \{ [1 + \kappa y]^{-1/\kappa} \}. \tag{1.4}$$

Here  $\kappa \in \mathbb{R}$  is the shape parameter also called the tail index: when  $\kappa \rightarrow 0$ , the distribution corresponds to a Gumbel EVL; when the tail index is positive, it corresponds to a Fréchet EVL; when  $\kappa$  is negative, it corresponds to a Weibull EVL. The EVL obtained depends on the kind of observable chosen. In particular, in [14,15] the authors have shown that, once taken the observable:

$$w(y) = -\log(y), \tag{1.5}$$

one gets a Gumbel EVL, here  $y = \text{dist}(T^m x, z)$ . In the next section we prove the existence of Gumbel law for the maps perturbed with observational noise. It is in fact possible to introduce other observables than the one specified above in order to get convergence towards Fréchet and Weibull EVLs. However, for any choice different from  $w(y) = -\log(y)$ , the tail index can be written in terms of the local dimension (see Eqs. (4.2)–(4.4) in [19]). For a sequence  $(u_m)_{m \in \mathbb{N}}$  satisfying (1.3) we define:

$$U_m := \{X_0 > u_m\}. \tag{1.6}$$

When  $X_0, X_1, X_2, \dots$  are not independent, the standard exponential law still applies under some conditions on the dependence structure. These conditions are the following:

**Condition** ( $D_2(u_m)$ ). We say that  $D_2(u_m)$  holds for the sequence  $X_0, X_1, \dots$  if for all  $\ell, t$  and  $m$ ,

$$\begin{aligned} & |\mathbb{P}(X_0 > u_m \cap \max\{X_t, \dots, X_{t+\ell-1} \leq u_m\}) \\ & - \mathbb{P}(X_0 > u_m)\mathbb{P}(M_\ell \leq u_m)| \leq \gamma(m, t), \end{aligned} \tag{1.7}$$

where  $\gamma(m, t)$  is decreasing in  $t$  for each  $m$  and  $m\gamma(m, t_m) \rightarrow 0$  when  $m \rightarrow \infty$  for some sequence  $t_m = o(m)$ .

Now, let  $(k_m)_{m \in \mathbb{N}}$  be a sequence of integers such that

$$k_m \rightarrow \infty \quad \text{and} \quad k_m t_m = o(m). \tag{1.8}$$

**Condition** ( $D'(u_m)$ ). We say that  $D'(u_m)$  holds for the sequence  $X_0, X_1, X_2, \dots$  if there exists a sequence  $(k_m)_{m \in \mathbb{N}}$  satisfying (1.8) and such that

$$\lim_{m \rightarrow \infty} m \sum_{j=1}^{\lfloor m/k_m \rfloor} \mathbb{P}(X_0 > u_m, X_j > u_m) = 0. \quad (1.9)$$

By following Freitas and Freitas [20, Theorem 1], if conditions  $D_2(u_m)$  and  $D'(u_m)$  hold for  $X_0, X_1, X_2, \dots$ , then there exists an EVL for  $M_m$  and  $H(\tau) = 1 - e^{-\tau}$ .

In the paper [15], Freitas, Freitas and Todd made the interesting observation that the extreme value laws are intimately related to the concept of local recurrence, in particular to the first hitting time function in small sets. Their analysis has been brought to systems perturbed with random transformations in [1]. In Section 3 we will show that this kind of results holds also for systems perturbed with observational noise, first by adapting the definition of first hitting time and then by showing that it follows an exponential law tempered by the strength of the perturbation.

We remark that the analogy between extreme value laws and local recurrences is possible for particular observables of the type  $X_0(\cdot) = w(\text{dist}(\cdot, z))$ , where  $z$  is a given point. As explained in [20], the additional choice  $w(y) = -\log(y)$ , allows to get the Gumbel law (a direct proof of this fact is given after Proposition 2), and moreover it brings information on the local structure of the invariant measure, as we will explain in a moment. This fact represents the main motivation for using such observable although a more detailed discussion and other motivations can be found in [2,19].

We conclude the introduction by stressing what we believe is an interesting and very general result. The probability  $\mathbb{P}$ , used to rule out the distribution of the maxima, is the product between the invariant measure  $\nu$  and the measure of the noise  $\theta$ . Whenever one is able to prove the existence of an extreme value law for the process  $X_m$  with the observable (1.5), then Proposition 2 shows how the two sequences  $a_m$  and  $b_m$  appearing in the affine choice for  $u_m$ , are related to the local behavior of the measure  $\nu$  at a local scale given by the intensity of the noise. This local behavior is related to the fine structure of the measure  $\nu$ . We have therefore a useful tool to detect the fine geometric properties of the invariant measure by calibrating the normalizing sequence  $u_m$  until we get the Gumbel law at different scales for the noise.

## 2. Recurrences for time series: a theoretical approach

We now show how to adapt the EVT for orbits of dynamical systems perturbed by instrument-like-error.

Although the paper is dedicated to the observational noise, for completeness and in light of applications to time series, we want remark also some relevant properties of truncations. Each instrument introduces an effect related to the combined accuracy and precision of the measure by replacing the real (unknown) value with the biased indicated by the instrument itself. In general, if the real dynamics can be represented by the map  $T$ , what one observes is formally:

$$\varphi(i) = \text{trunc}(T^i x + \epsilon \xi_i, q),$$

where  $\text{trunc}(x, q) = \frac{\lfloor 10^q \cdot x \rfloor}{10^q}$  is the truncation introduced by the instrument precision and  $\epsilon \xi_i$  is a random displacement from the real value. This displacement is what we have defined as observational noise. Here it is rewritten as  $T^i x + \epsilon \xi_i$ , with the parameter  $\epsilon \in \mathbb{R}^+$ .

The role of the truncation, important also for numerical computations, has been discussed in the book by Knuth [21] and then analyzed, among others, by [22–24]. On the  $q$  digit, the truncation acts essentially as a random noise of variance  $\sigma = 10^{-q}$ . If  $q \gg 1$  the measure underlying  $\varphi(i)$  will match the one of the

original dynamics given by the map  $T$ , if not the support will appear as a collection of Dirac's deltas, precluding the convergence to the GEV distribution [25]. We defer to Section 4.1 for a numerical study on the truncation error showing for which values of  $q$  one should take truncations into account and whether  $q \simeq 7$  (the common truncation corresponding to a double precision representation) is a good choice for representing the properties of a deterministic dynamics.

We thus proceed to show that EVLs persist for chaotic dynamical systems, whenever they are perturbed with the observational noise [26,12]. The proof we give is reminiscent of that of Theorem D in [1]. We first point out that, in order to guarantee the stationarity of the random process involved in the distribution of the maxima, we need to evaluate the observable  $w$  at the point  $x + \xi$ , where  $x \in \Omega$  and  $\xi$  is a random vector. For this reason and in order to avoid ambiguities, we will choose  $\Omega$  as a torus. Another alternative would be to take  $\Omega$  which is strictly sent into itself by  $T$ ,  $T\Omega \subset \Omega$ , and with all the components of  $\xi$  small enough (see, for instance, Proposition 4.5 in [1]). This choice is often invoked for random transformation acting of bounded domains of  $\mathbb{R}^n$ . The paragraph below contains the assumptions on the systems which allow us to prove our main results.

**Assumption M.** We consider maps  $T$  defined on the torus  $\Omega = \mathbb{T}^d$  with norm  $\|\cdot\|$  and satisfying:

- There exists a finite partition (mod-0) of  $\Omega$  into open sets  $Y_j, j = 1, \dots, p$ , namely  $\Omega = \bigcup_{j=1}^p \bar{Y}_j$ , such that  $T$  has a Lipschitz extension on the closure of each  $Y_j$  with a uniform and strictly larger than 1 Lipschitz constant  $\eta$ ,  $\|T(x) - T(y)\| \leq \eta \|x - y\|, \forall x, y \in \bar{Y}_j, j = 1, \dots, p$ .
- $T$  preserves a Borel probability measure  $\nu$  which is also mixing with decay of correlations given by

$$\left| \int f \circ T^m h d\nu - \int f d\nu \int h d\nu \right| \leq C \|h\|_{\mathcal{B}} \|f\|_1 m^{-2} \quad (2.1)$$

where the constant  $C$  depends only on the map  $T$ ,  $\|\cdot\|_1$  denotes the  $L^1_\nu$  norm with respect to  $\nu$  and finally  $\mathcal{B}$  is a Banach space included in  $L^\infty_{\mathcal{L}}$ , where  $\mathcal{L}$  denotes the Lebesgue (Haar) measure on  $X$ : the corresponding norm will be denoted with  $\|\cdot\|_\infty$ . We will also need  $\nu$  to be equivalent to  $\mathcal{L}$  with density in  $L^\infty_{\mathcal{L}}$ .

**Assumption N.** We consider a sequence  $\xi_i$  of i.i.d. vector-valued random variables which take values in the hypersphere  $S := S^d \subset \mathbb{R}^d$  centered at 0 and of radius 1,  $S := \{u \in \mathbb{R}^d; \|u\| \leq 1\}$ , and with common distribution  $\theta$ , which we choose absolutely continuous with density  $\rho \in L^\infty_{\mathcal{L}}$ , namely  $d\theta(\xi) = \rho(\xi)d\mathcal{L}(\xi)$ , with  $\int_S \rho(\xi)d\mathcal{L}(\xi) = 1$ .<sup>1</sup>

**Remark 1.** The paper [1] contains examples of endomorphisms verifying the Assumption M, in particular one-dimensional Rychlik maps endowed with bounded variation functions and multidimensional piecewise uniformly expanding maps endowed with quasi-Hölder observables. In order to get the decay of correlations (2.1), one needs more regularity for  $T$ , usually  $C^{1+\alpha}$ . Our next proofs crucially depend on the decay against  $L^1_\nu$  functions<sup>2</sup>; we believe that one could weaken such assumption and extend the theory to invertible maps, but that would need a different approach: in order to support this claim, Section 4 will contain numerical computations on examples which are not covered by our analytical results, but which show similar behaviors. We will comment further on these issues in Section 3—Consequence 3.

<sup>1</sup> Each  $\xi$  is a vector with  $d$  components; all these components are independent and distributed with common density  $\rho$ ; the product of such marginals  $\rho$ 's gives  $\rho$ .

<sup>2</sup> Actually, we will stress in Section 3 that what is really needed is the  $L^1_\nu$  property for characteristic functions which, for some systems, is easier to show.

The random orbits  $T^i x + \epsilon \xi_i$  generates a new random process when an observable is computed along them. Suppose that  $w$  is a measurable real function defined on  $\Omega$ ; we take  $w(x) = -\log(\|x - z\|)$ , where  $z$  is a given point in  $\Omega$ . The process:

$$X_0 = w(x + \epsilon \xi_0), X_1 = w(Tx + \epsilon \xi_1), \dots, X_m = w(T^m x + \epsilon \xi_m)$$

is endowed with the probability  $\mathbb{P} = \nu \times \theta^{\mathbb{N}}$  defined on the product space  $\Omega \times S^{\mathbb{N}}$  with the product  $\sigma$ -algebra; a point in this space is the couple  $(x, \bar{\xi} := \{\xi_0, \xi_1, \dots, \xi_m, \dots\}) \in X \times S^{\mathbb{N}}$ <sup>3</sup>.

**Remark 2.** Before checking conditions  $D_2(u_m)$  and  $D'(u_m)$ , we notice that, contrarily to the random setting studied in [1], the random variable  $X_0$  depends now not only on the initial condition  $x \in \Omega$ , but also on the random variable  $\xi$  and this makes  $\mathbb{P}$  stationary.

With the given choice of the observable  $w$ , the set  $U_m$  is explicitly given by

$$U_m = \{(x, \xi); \|x + \epsilon \xi - z\| \leq e^{-u_m}\}.$$

For convenience, we will also set  $V_m := B(z, e^{-u_m})$ , the ball of center  $z$  and radius  $e^{-u_m}$ .

**Proposition 1.** *Let us suppose that our dynamical systems verifies the Assumption M and it is perturbed with observational noise satisfying the Assumption N. Then conditions  $D_2(u_m)$  and  $D'(u_m)$  hold for the observable  $w$ .*

**Proof.** We will give the proof when  $T$  is continuous on the torus. The extension to the piecewise Lipschitz case is straightforward and it could be done as explained in the analogous extension proofs of Propositions 4.2 and 4.5 in [1] to which we defer for further details. We begin to check condition  $D_2(u_m)$  by estimating the contribution given by the first term on the l.h.s. of Eq. (1.7):

$$\begin{aligned} &\mathbb{P}(X_0 > u_m \cap \max\{X_t, \dots, X_{t+\ell-1} \leq u_m\}) \\ &= \iint d\nu d\theta^{\mathbb{N}} \mathbf{1}_{\{g(x+\epsilon\xi_0) > u_m\}} \mathbf{1}_{\{g(T^t x + \epsilon\xi_t) \leq u_m\}} \dots \\ &\quad \mathbf{1}_{\{g(T^{t+\ell-1} x + \epsilon\xi_{t+\ell-1}) \leq u_m\}}, \end{aligned} \tag{2.2}$$

the set of integration variables here being  $(x, \bar{\xi}^{(t+l)})$  with  $\bar{\xi}^{(t+l)} = (\xi_0, \xi_1, \dots, \xi_{t+l-1})$ . We apply Fubini's theorem and factorize the integrals by exploiting the independence of the variables  $\xi_i$ , so that the previous expression becomes:

$$\int d\nu \int \mathbf{1}_{\{g(x+\epsilon\xi_0) > u_m\}} d\theta(\xi_0) \prod_{i=1}^{l-1} \int \mathbf{1}_{\{g(T^{t+i} x + \epsilon\xi_i) \leq u_m\}} d\theta(\xi_i).$$

Let us introduce the measurable functions:

$$H_m(l, x) = \prod_{i=1}^{l-1} \int \mathbf{1}_{\{g(T^{t+i} x + \epsilon\xi_i) \leq u_m\}} d\theta(\xi_i);$$

$$G_m(x) = \int \mathbf{1}_{\{g(x+\epsilon\xi_0) > u_m\}} d\theta(\xi_0)$$

with  $G_m(x) \in L^1_\nu$  and  $H_m(l, x) \in \mathcal{B}$ . Then Eq. (2.2) can be rewritten as  $\int d\nu G_m(x) H_m(l, T^t x)$ . By the decay of correlations assumption, we get

$$\begin{aligned} &\left| \int d\nu G_m(x) H_m(T^t(x)) - \int d\nu G_m(x) \int d\nu H_m(l, x) \right| \\ &\leq C \|G_m\|_1 \|H_m\|_\infty t^{-2} \leq C t^{-2}, \end{aligned}$$

where  $C$  is a constant depending only on  $T$ . In the previous equation, the second term on the l.h.s. corresponds to  $\mathbb{P}(X_0 > u_m) \mathbb{P}(M_\ell \leq u_m)$  which is the second term on the l.h.s. of the condition  $D_2(u_m)$ . Let us note, and this will be useful later, that  $D_2(u_m)$  holds with  $\gamma(m, t) = \gamma(t) = C^* t^{-2}$  for some  $C^* > 0$  and  $t_m = m^{-\beta}$ , with  $1/2 < \beta < 1$ .

In order to deal with Condition  $D'(u_m)$ , we follow the same strategy as in [1]. We begin to define the *approximated first return time of the point  $x$  in  $V_m$*  in the following way: we fix the couple  $(\xi, \xi') \in S^2$  and we set

$$r_{V_m, \xi, \xi'}(x) := \min\{j \geq 1, T^j x + \epsilon \xi' \in V_m; x + \epsilon \xi \in V_m\}.$$

Notice that we keep fixed the variables  $\xi, \xi'$  while iterating the point  $x$ . Moreover, instead of  $x$ , we require the initial condition  $x + \epsilon \xi$  to be in  $V_m$ . Then we define the *approximated first return time of the set  $V_m$  into itself* as:

$$R_{V_m, \xi, \xi'} := \min_{\{x, x+\epsilon\xi \in V_m\}} \{r_{V_m, \xi, \xi'}(x)\}.$$

We observe that  $\|T^j(V_m - \epsilon\xi) + \epsilon\xi'\| \leq \eta^j \|V_m - \epsilon\xi\|$ , where  $\eta > 1$  is the highest rate of separations for the points. We use the symbol  $\|\cdot\|$  indifferently to denote *distance* and *diameter*. The notation  $z + B$ , where  $z \in M$  and  $B$  is a subset of  $M$ , stands for the set  $(\cup_{w \in B} \{z+w\})$ . We now fix some sequences  $(\alpha_m)_{m \in \mathbb{N}}$  going to infinity and such that  $\alpha_m = o(\log k_m)$ , where  $k_m$  is the sequence defined in (1.6). Therefore, whenever

$$\|T^j(z - \epsilon\xi) + \epsilon\xi' - (z - \epsilon\xi)\| > 2\eta^j \|V_m - \epsilon\xi\|,$$

$\forall j = 1, \dots, \alpha_m$ , then

$$T^j(V_m - \epsilon\xi) + \epsilon\xi' \cap (V_m - \epsilon\xi) = \emptyset,$$

which in turn implies that  $R_{V_m, \xi, \xi'} > \alpha_m$ . Since

$$\begin{aligned} &\{(\xi, \xi') \in S^2; R_{V_m, \xi, \xi'} \leq \alpha_m\} \\ &\subset \cup_{j=1}^{\alpha_m} \{(\xi, \xi' \in S^2); \|T^j(z - \epsilon\xi) + \epsilon\xi' - (z - \epsilon\xi)\| \\ &\leq 2\eta^j \|V_m - \epsilon\xi\|\}, \end{aligned}$$

we have

$$\begin{aligned} &\theta^2 \{(\xi, \xi') \in S^2; R_{V_m}(\xi, \xi') \leq \alpha_m\} \\ &\leq \sum_{j=1}^{\alpha_m} \int d\xi \rho(\xi) \int d\xi' \rho(\xi') \mathbf{1}_{\{\xi', \|\xi' - \frac{(z-\epsilon\xi - T^j(z+\epsilon\xi))}{\epsilon}\| \leq \frac{2\eta^j}{\epsilon} \|V_m - \epsilon\xi\|\}} \\ &\leq \mathcal{O} \left( \|\rho\|_\infty \sum_{j=1}^{\alpha_m} 2^d \eta^{jd} \|V_m\|^d \epsilon^{-d} \right) \\ &\leq \mathcal{O} (\|g\|_\infty \|V_m\|^d \epsilon^{-d} \eta^{d\alpha_m}). \end{aligned}$$

We now have:

$$\begin{aligned} &m \sum_{j=1}^{\lfloor \frac{m}{k_m} \rfloor} (\nu \times \theta^{\mathbb{N}}) \{(x, \bar{\xi}); X_0 > u_m; X_j > u_m\} \\ &= m \sum_{j=\alpha_m}^{\lfloor \frac{m}{k_m} \rfloor} \int d\nu \int d\theta(\xi) \mathbf{1}_{\{g(x+\epsilon\xi) > u_m\}} \end{aligned}$$

<sup>3</sup> With this final notation the random process is better defined as  $T^i x + \epsilon \Pi_i(\bar{\xi})$ , where  $\Pi_i(\bar{\xi})$  projects onto the  $i$ th component  $\xi_i$ .

$$\begin{aligned} & \times \int d\theta(\xi') \mathbf{1}_{\{g(T^j x + \varepsilon \xi') > u_m\}} \} \\ & + m \sum_{j=1}^{\alpha_m} \iint \mathbf{1}_{\{(\xi, \xi') : R_{V_m, \xi, \xi'} \leq \alpha_m\}} d\theta(\xi) d\theta(\xi') \cdot A(j, \xi, \xi') \\ & = I + II \end{aligned}$$

where  $A(j, \xi, \xi') = \left\{ \int d\nu \mathbf{1}_{\{x + \varepsilon \xi \in V_m\}} \mathbf{1}_{\{T^j x + \varepsilon \xi' \in V_m\}} \right\}$ .

Again, the first term (I) on the l.h.s. can be estimated by using decay of correlations applied to the (same) observable  $\tilde{H}(x) = \int d\theta(\xi) \mathbf{1}_{\{g(x + \varepsilon \xi) > u_m\}}$ ; we easily get:

$$\begin{aligned} I & \leq m \sum_{j=\alpha_m}^{\lfloor \frac{m}{k_m} \rfloor} \{ \mathbb{P}(U_m)^2 + C \mathbb{P}(U_m) j^{-2} \} \\ & \leq \frac{(n \mathbb{P}(U_m))^2}{k_m} + m C \mathbb{P}(U_m) \sum_{j=\alpha_m}^{\lfloor \frac{m}{k_m} \rfloor} j^{-2} \\ & = \mathcal{O} \left( \frac{\tau^2}{k_m} + \tau \sum_{j=\alpha_m}^{\lfloor \frac{m}{k_m} \rfloor} j^{-2} \right) \xrightarrow{m \rightarrow +\infty} 0 \end{aligned}$$

since  $m \mathbb{P}(U_m) \rightarrow \tau$ .

For the second term (II) we use Hölder's inequality and the fact that  $\nu$  is equivalent to  $\mathcal{L}^d$  with essentially bounded density and  $\mathcal{L}$  is translationally invariant:

$$\begin{aligned} II & \leq \mathcal{O} \left( m \sum_{j=1}^{\alpha_m} \mathcal{L}(V_m) \theta^2 \{ (\xi, \xi') \in S^2; R_{V_m, \xi, \xi'} \leq \alpha_m \} \right) \\ & \leq \mathcal{O} \left( m \sum_{j=1}^{\alpha_m} \mathcal{L}(V_m) \|\rho\|_\infty \|V_m\|^d \varepsilon^{-d} \eta^{d\alpha_m} \right). \end{aligned}$$

Since  $\mathbb{P}(U_m) \approx \mathcal{L}(V_m)$  and  $\|V_m\| \approx \mathcal{L}(V_m)^{1/d}$  we finally have

$$II = \mathcal{O} \left( \frac{m^2}{k_m} \mathbb{P}(U_m)^2 \eta^{d\alpha_m} \right) = \mathcal{O} \left( \tau^2 \frac{\eta^{d\alpha_m}}{k_m} \right),$$

which goes to zero with the prescribed assumptions for  $\alpha_m$  and  $k_m$ .  $\square$

### 3. Generalizations and consequences

In this section we would like to point out a few interesting properties of the observational noise. We start by an explicit calculation of the quantity  $\tau$  defined in (1.3) for the observable (1.5). The computation is done in  $d = 1$ , but the generalization to higher dimensions is trivial. As we have anticipated in Section 1, the following proposition requires only the existence of the Gumbel law for the process under investigation, which is of course true for systems verifying Proposition 1.

**Proposition 2.** *Let us suppose that the one-dimensional dynamical systems  $(\Omega, \mathcal{B}, \nu, T)$  verifies a Gumbel law for the process  $X_m(x, \xi) := -\log(|T^m x + \Pi_m(\xi) - z|)$  endowed with the probability  $\mathbb{P} = \nu \times \theta^{\mathbb{N}}$ , and also that  $\theta$  is the Lebesgue measure on  $S$ . Then the linear sequence  $u_m := u/a_m + b_m$ , defined by (1.3), verifies:*

$$a_m = 1; \quad b_m = \log \left( \frac{m\nu(B(z, \varepsilon))}{\varepsilon} \right).$$

**Proof.** We begin to observe that

$$\begin{aligned} m \mathbb{P}(X_0 > u_m) & = m \int d\nu(x) \left( \int d\theta(\xi) \mathbf{1}_{\{\xi; \|\xi - \frac{(z-x)}{\varepsilon}\| < \frac{e^{-u_m}}{\varepsilon}\}} \right) \\ & = m \int d\nu(x) \theta \left( B \left( \frac{z-x}{\varepsilon}, \frac{e^{-u_m}}{\varepsilon} \right) \cap B(0, 1) \right) \end{aligned}$$

since the variable  $\xi$  must stay in the ball of center 0 and radius 1.

Let us now introduce

$$u_m := -\log \left( \frac{\varepsilon \tau}{m\nu(B(z, \varepsilon))} \right) = u/a_m + b_m,$$

with  $u := -\log \tau$ ;  $a_m := 1$ ;  $b_m := \log \left( \frac{m\nu(B(z, \varepsilon))}{\varepsilon} \right)$ ; and observe that

$$m \theta \left( B \left( \frac{z-x}{\varepsilon}, \frac{e^{-u_m}}{\varepsilon} \right) \cap B(0, 1) \right) \leq m \frac{e^{-u_m}}{\varepsilon} \leq \frac{m}{\varepsilon} \frac{\varepsilon \tau}{m\nu(B(z, \varepsilon))}.$$

This bound is independent from  $m$  and integrable.<sup>5</sup>

We can apply the theorem of dominated convergence since

$$\lim_{m \rightarrow \infty} m \theta \left( B \left( \frac{z-x}{\varepsilon}, \frac{e^{-u_m}}{\varepsilon} \right) \cap B(0, 1) \right) = \mathbf{1}_{B(z, \varepsilon)}(x) \frac{\tau}{\nu(B(z, \varepsilon))}.$$

Having passed the limit inside, the integral finally gives  $\tau$ :

$$m \mathbb{P}(X_0 > u + b_m) \rightarrow \tau$$

and therefore

$$\mathbb{P}(M_m \leq u + b_m) \rightarrow \exp(-e^{-u})$$

which is exactly the Gumbel law.

Whenever  $d > 1$  a similar computation immediately gives that the linear sequence  $u_m := u/a_m + b_m$  verifies:

$$a_m = d; \quad b_m = \frac{1}{d} \log \left( \frac{m\nu(B(z, \varepsilon))}{\varepsilon^d} \right). \quad \square$$

The following useful consequences will be exploited in the next section:

- **Consequence 1**

The scaling parameter  $b_m$  depends on the target point  $z$  via the local density of the invariant measure in a ball of radius given by the  $\varepsilon$ . Let us start by considering the case of absolutely continuous invariant measures. If the point  $z$  is visited with less frequency, the local density will be of lower order in  $\varepsilon$ , which means that one should go to higher values of  $m$  in order to have a reliable statistics. This is the case, for instance, for the points  $\pm 1$  for the map introduced by Hemmer [27]:

$$T(x) = 1 - 2\sqrt{|x|} \tag{3.1}$$

and defined on the interval  $[-1, 1]$ . The invariant density  $\rho$  can be computed directly by inspection and reads:  $\rho(x) = \frac{1}{2}(1-x)$ . Therefore  $\nu(B(-1, \varepsilon)) \approx \varepsilon^2$  and  $b_m \approx \log(m\varepsilon)$ .

A complementary issue will appear whenever the map exhibits a laminar behavior in some regions of the phase space and therefore it will spend there a lot of time. This happens, for instance, for the well-known map of Pomeau–Manneville [28], which could be written as:

$$\begin{cases} T_1(x) = x + 2^\alpha x^{1+\alpha}, & 0 \leq x \leq 1/2 \\ T_2(x) = T(x) = 2x - 1, & 1/2 \leq x \leq 1. \end{cases} \tag{3.2}$$

<sup>4</sup> We will use the symbol “ $\approx$ ” to signify that equivalence, namely there exists a positive constant  $\iota$  such that for any measurable set  $A$  we have that  $\iota^{-1} \mathcal{L}(A) \leq \nu(A) \leq \iota \mathcal{L}(A)$ .

<sup>5</sup> Here we use crucially the fact that  $\theta$  is exactly Lebesgue, since its translational invariance property allows us to get rid of the variable  $x$  in the center of the ball.

The origin 0 is a neutral fixed point and, for  $0 < \alpha < 1$ , the density of the absolutely continuous invariant measure  $\nu$  behaves like  $x^{-\alpha}$  in the neighborhood of 0 [29]. Therefore,  $\nu(B(0, \varepsilon)) \approx \varepsilon^{1-\alpha}$  and  $b_m \approx \log(m\varepsilon^{-\alpha})$ . We will analyze in the next section which finite size effects arise for these two examples.

*Warning:* strictly speaking the previous two maps do not fit with the assumptions of Proposition 1. In fact, in both cases it is not possible to prove a polynomial decay of correlations against all  $L^1_\nu$  functions as it was shown in [30]. On the other hand, an inspection of the proof shows that we only require the  $L^1_\nu$  property for characteristic functions, so that, in principle, such a decay could be obtained. This was achieved, for instance, in the case of rotations by using a Fourier series technique [19]. An additional problem concerns the rate of decay. For the Pomeau–Manneville map it is of order  $n^{-\frac{1}{\alpha}+1}$  [29], so that it fits with our assumption whenever  $\alpha < 1/3$ ; for the Hemmer’s map the situation is worse since the correlations decay as  $\frac{1}{n}$  [31]. Nevertheless, we conjecture that Proposition 1 could be applied to the latter case as well, as we will argue in Consequence 3, and as we will show numerically in the next section.

• **Consequence 2**

Let us observe that, if we define the linear scaling factor as above, we have convergence towards the Gumbel law  $e^{-e^{-u}}$  for any  $z$ . This shows that the extremal index defined and studied in [17,1] is 1 everywhere implying that there are no points which behave like unstable fixed points of deterministic dynamical systems.

• **Consequence 3**

Although we were able to prove Proposition 1 whenever the invariant measure is equivalent to Lebesgue, we pointed out that Proposition 2 is true for any invariant measure, providing that one can show the existence of an EVL. Let us therefore suppose that a given dynamical system with invariant measure  $\mu$ , not necessarily equivalent to Lebesgue, admits an extreme value law for the observable  $w$  under the observational noise. Then we can apply Proposition 2, which only requires that  $\theta$  is Lebesgue. We remark that in the expression for  $b_m$  we are considering a fixed positive size for the noise  $\varepsilon$ . We remind that the scaling parameters  $b_m$  are linked to the location parameter of the EVLs. Therefore, the dependence of  $b_m$  on the noise structure gives hint on how the extreme value distribution drifts by changing the noise intensity. Let us suppose that at this scale we have  $\nu(B(z, \varepsilon)) \approx \varepsilon^D$ , where  $D$  is usually an estimation of the geometric and fractal properties of  $\nu$  at the point  $z$ .<sup>6</sup> In this case, if the ambient space has dimension  $d$ , the linear scaling parameter  $b_m$  has the form:

$$b_m \sim \frac{1}{d} \log(m\varepsilon^{D-d}). \tag{3.3}$$

Therefore, we have an useful technique to detect the local dimensions of the measure, with a *finite* resolution given by the strength of the noise. This will also allow us to compute directly

the distribution of the maxima with the linearization, given the explicit expression of the  $u_m$ . This would be particularly useful whenever the invariant measure is singular and therefore the GEV distribution does not admit a probability density function: see Section 3.1 in [2] for a detailed discussion on this point.

• **Consequence 4**

As we said in the Introduction, another interesting property of the observational noise is its direct relationship with the statistics of first hitting times in small sets. We first define the *first hitting time of the set  $V_m$*  in a slightly different manner which respect to the quantity  $r_{V_m}$  introduced above. Given the triple  $(V_m, x \in M, \bar{\xi} \in S^{\mathbb{N}})$ , we set:

$$\mathcal{R}_{V_m}(x, \bar{\xi}) := \min\{j \geq 1, T^j x + \varepsilon \xi_j \in V_m\}.$$

Let us notice that, contrarily to  $r_{V_m}$ , the error increment changes at each step since we are now dealing with a *true* random orbit; this easily implies that

$$\mathbb{P}(M_m \leq u_m) = \mathbb{P}(\mathcal{R}_{V_m} > m).$$

By using the expression of  $u_m$  found above and setting consequently  $V_m = B(z; \frac{\varepsilon t}{m\nu(B(z, \varepsilon))})$ , in dimension 1, we could rewrite the previous formula as, for  $t \in \mathbb{R}$ :

$$\mathbb{P}\left(\frac{t}{m} \mathcal{R}_{B(z; \frac{\varepsilon t}{m})} > t\right) \rightarrow e^{-\frac{t\nu(B(z, \varepsilon))}{\varepsilon}} \tag{3.4}$$

which shows that the first hitting time follows an exponential law tempered by the strength noise  $\varepsilon$ .

**4. Discussion and numerical results**

In this section we discuss some important implications connected to the introduction of observational noise in finite time series. In particular, through numerical experiments devised on low dimensional maps, we show that the influence of truncations and observational noise is related to the intensity of the perturbation applied. Some of the theoretical findings presented in the previous sections present a practical interest in a wide range of applications namely the analysis of the role of truncation errors for instrument with low accuracy, the statistics of points visited sporadically in the analysis of recurrence of time series and the possibility of computing attractor dimension by using Eq. (3.3) as an alternative way to other techniques. For each maps, we present and comment numerical experiments and outline possible further applications.

Before introducing such examples, let us make a useful comment. A close inspection to the proof of Proposition 1, shows that the parameter  $\varepsilon$  appears in the denominator of one factor in the r.h.s. of the term (II) at the end of the proof. This means that the convergence gets better when  $\varepsilon$  is large, which is not surprising since a large value of the perturbation implies a more stochastic independence of the process. Moreover, if we want to use the form of the linear scaling parameter  $b_m$  to catch the local properties of the invariant measure  $\nu$ , we also need large values of  $m$ . A judicious balance of the value of  $\varepsilon$  between these two regimes is therefore necessary when we pursue such numerical analysis.

**4.1. Truncation**

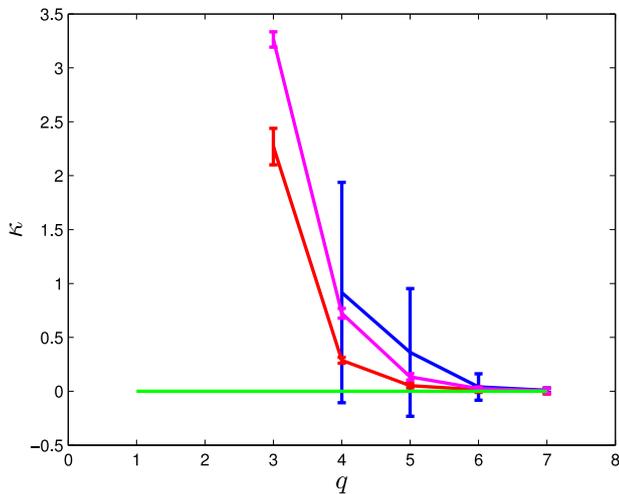
We perturb a ternary shift map with truncation error on the different digits  $q$ :

$$\varphi(i) = \text{trunc}(3x_i \text{ mod } 1, q).$$

The experiment consists in producing 30 orbits starting from different initial conditions taken on the support of the truncated measure. The length of the orbit is fixed according to the results presented in [25] to be such that  $n = 1000$ . In order to analyze

<sup>6</sup> We give an example: a measure  $\nu$  on  $\mathbb{R}^+$  is called *Ahlfors upper semi-regular* if there is a constant  $C > 0$  and a real number  $\alpha > 0$  such that for all non-empty open intervals  $I \subset \mathbb{R}^+$   $\frac{\nu(I)}{(\text{diam}(I))^\alpha} < C$ .

Ahlfors upper semi-regular measures include fractal measures like the measures of maximal dimension of dynamically defined Cantor sets, i.e. Cantor sets that arise from smooth expanding repellers. More generally, given an invariant measure  $\mu$  and whenever the limit  $\lim_{r \rightarrow 0^+} \frac{\log \mu(B(x, r))}{\log r}$  exists  $x - \mu$ -almost everywhere, then this limit equals the *Hausdorff dimension of the measure  $\mu$* ,  $HD(\mu) := \inf\{\text{Hausdorff dimension of } Y, \mu(Y) = 1\}$  [32], also called *information dimension*. In some cases  $HD(\mu)$  is given by suitable relations between Lyapunov exponents and entropies, formulae better known as *Kaplan–Yorke* and *Ledrappier–Young*: see [33] for a detailed exposition of these issues.

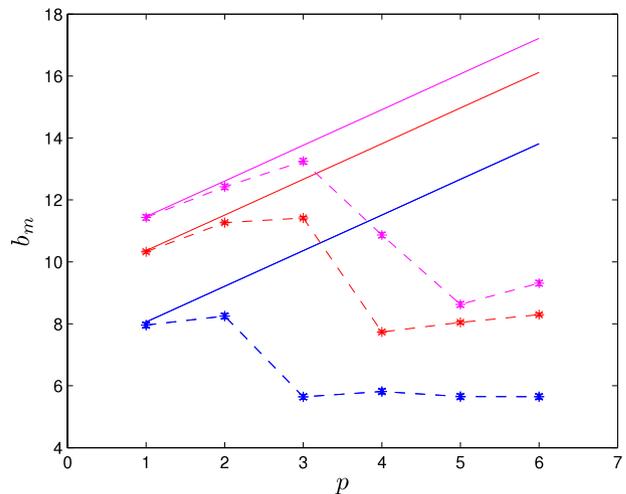


**Fig. 1.** Shape parameter  $\kappa$  vs.  $q$ , the digit where the truncation has been applied. Error-bars display the average of  $\kappa$  over 30 realizations and the standard deviation of the samples for  $m = 300$  (blue),  $m = 1000$  (magenta) and  $m = 3000$  (red). The green line corresponds to the Gumbel law ( $\kappa = 0$ ).  $n = 1000$  for all the cases considered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the effect of varying the bin length combined to the order of the truncation, we consider three different values of  $m$ . Each series of  $w$  is therefore divided in  $n$  bins and in each of them the maxima  $M_j$  are extracted and then fitted to the GEV model via the L-moments procedure described in [19]. Note also that the points  $z$  are chosen after applying the truncation. The results are shown in Fig. 1 where the behavior of the shape parameter  $\kappa$  is compared to the asymptotic Gumbel law  $\kappa = 0$ . In the deterministic limit  $q \gg 1$  the usual Gumbel EVL is recovered, whereas for  $q \rightarrow 1$  the discretization of the invariant measure becomes relevant and the asymptotic EVL appears as a collection of Dirac deltas thus producing a divergence of the shape parameter from 0. This is exactly what is visible in Fig. 1: the convergence gets worse when  $q < 6$  and at  $q = 3$  we are already unable to fit the GEV distribution for all the  $z$  points considered, although by increasing the bin length  $m$  the convergence improves as one would expect. These results, as we tested, are reproducible in other maps and gives a more general indication that computing asymptotic properties on truncated series leads to estimation errors and divergence even at high order of truncation.

#### 4.2. Highly recurrent and sporadic points

With the introduction of the observational noise, the scaling parameter  $b_m$  depends on the target point  $z$  via the local density of the invariant measure in a ball whose radius is given by the error  $\varepsilon$ . In *Consequence 1* we said that if the point  $z$  is visited with less frequency, the local density is of lower order with respect to  $\varepsilon$ , which means that one should go to higher values of  $m$  in order to have a reliable statistics. Here we want to test that the order of  $m$  needed to get convergence to the asymptotic  $b_m$  is lower for a highly recurrent point than for a sporadic one. As highly recurrent point we choose  $z = 0$  for the Pomeau–Manneville map in Eq. (3.2) and as sporadic one the point 1 of the map introduced by Hemmer and reported in Eq. (3.1). The experiment consists in computing 30 realizations of the maps perturbed with observational noise. Again, we fit the maxima of the observable  $w$  to the GEV distribution by using the L-moments procedure and compare the values of  $b_m$  obtained experimentally to the theoretical ones stated in *Proposition 2*. We report here the results for three different bin lengths  $m = 1000, 10\,000, 30\,000$  in Fig. 2 for the Pomeau–Manneville map and in Fig. 3 for Hemmer map. The figures show how  $b_m$  varies as a function of the noise  $\varepsilon = 10^{-p}$ , in terms of  $p$ . In both cases we

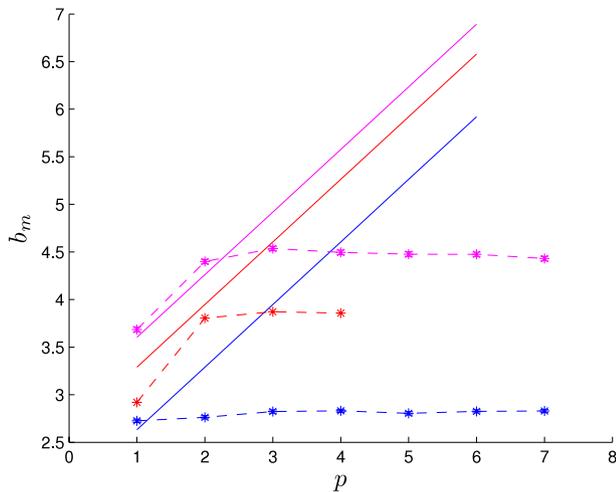


**Fig. 2.** Normalizing sequence  $b_m$  vs. intensity of the noise in terms of  $p$  (we recall that  $\varepsilon = 10^{-p}$ ) for the Pomeau–Manneville map (Eq. (3.2)). We recall that dashed error-bars display the average of  $b_m$  over 30 realizations and the standard deviation of the sample. Solid lines indicate the theoretical values. The blue, red and magenta curves respectively refers to  $m = 1000, 10\,000, 30\,000, z = 0, n = 1000$  for all the cases considered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

observe convergence towards the theoretical values (solid lines) for high values of  $\varepsilon$  (low orders in  $p$ ) whereas in the limit of weak noise one must increase the bin lengths to get convergence. The main result to be highlighted here is the better convergence of highly recurrent points with respect to the ones visited sporadically. This important property can be used to study time series recurrences and identify extremes as the points visited rarely for which the convergence towards the asymptotic parameters is bad. The main advantage of studying recurrence properties in this way over applying other techniques is due to the built-in test of convergence of this method: even for a point rarely recurrent there will be a time scale  $m$  such that the fit converges. For smaller  $m$ , we can therefore consider such a  $z$  as a *sporadically recurrent* point of the orbit as explained in [13]. There we show how to use this property to define rigorous recurrences in long temperature records collected at several weather stations. The convergence or divergence of the fit allows us for discriminating between temperatures belonging to the normal variability associated to the time scales defined by the bin length (e.g. the seasonal cycle) or as extremes temperature if there are no or few recurrences in  $m$ . In another study [34], we suggested a quantitative way to discriminate between highly recurrent points and sporadic points of the dynamics in a rather algorithmic way. Basically, one can assess the minimum bin length  $m$  such that the fit to one of the EVL converges, that is the value of  $m$  such that a sporadic point becomes a normally recurring one. For highly recurring points this typical value of  $m$  is order  $10^3$ , whereas for quasiperiodic dynamics can be larger than  $m = 10^9$ .

#### 4.3. Attractor dimensions

As we have already said in *Consequence 3*, if the invariant measure is not absolutely continuous, one could still perform the previous analysis, but both the ambient space dimension  $d$  and the local dimension  $D$  will enter in the computation of  $b_m$  via Eq. (3.3). This formula can be used in principle to test whether a map has a fractal support by comparing the local dimensions with the ambient space dimension: it is enough to check how the obtained  $b_m$  depend on the intensity of the noise  $\varepsilon = 10^{-p}$ . We should point out that, as we said in the footnote 1, we are targeting the Hausdorff dimension of the measure. We test this idea on the classical



**Fig. 3.** Normalizing sequence  $b_m$  vs. intensity of the noise in terms of  $p$  (we recall that  $\epsilon = 10^{-p}$ ) for the Hemmer map (Eq. (3.1)). Dashed error-bars display the average of  $b_m$  over 30 realizations and the standard deviation of the sample. Solid lines indicate the theoretical values. The blue, red and magenta curves respectively refers to  $m = 1000, 10\,000, 30\,000, z = 1, n = 1000$  for all the cases considered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

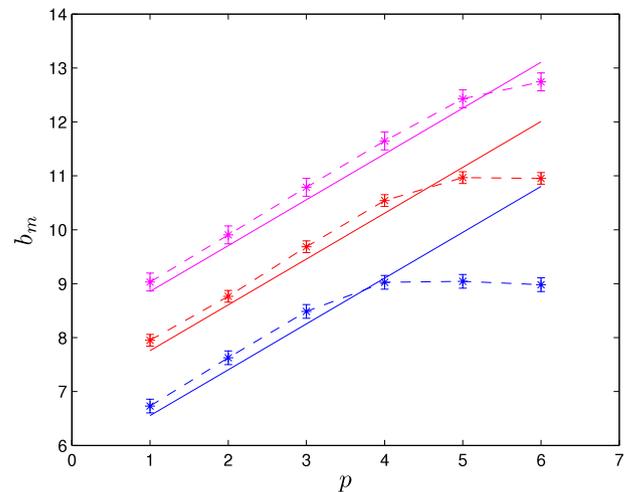
Iterated Function System  $\{T_1, T_2\}$  used to produce a Cantor set:

$$\begin{cases} T_1(x) = x/3 \text{ with weight } q_1 \\ T_2(x) = (x+2)/3 \text{ with weight } q_2 \end{cases} \quad (4.1)$$

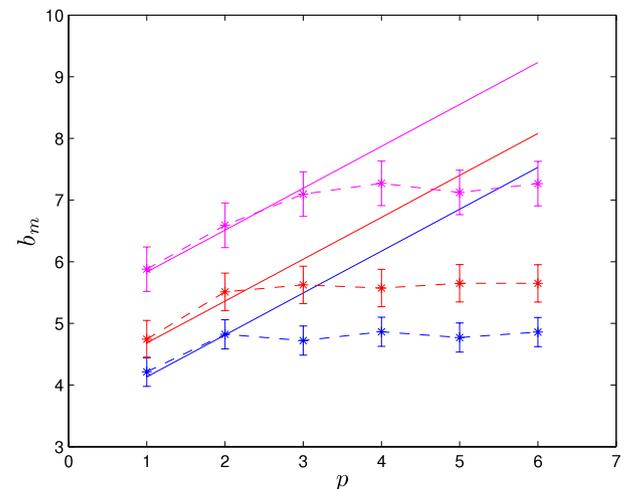
where  $x \in [0, 1]$ , and we set  $q_1 = q_2 = 1/2$ . Therefore, at each time step, we have the same probability to iterate  $T_1(x)$  or  $T_2(x)$ ; the *balanced* invariant measure associated to the map and the way to construct the orbit to detect the maxima are described in [2]. The results for different  $m$  are reported in Fig. 4 for an average among 30 different realizations and three different bin lengths  $n = 1000, m = 1000, 10\,000, 30\,000$ . Experimental data follow the prediction of Eq. (3.3) with  $D = \log(2)/\log(3)$  as Hausdorff dimension and  $d = 1$  as ambient space dimension but only up to a certain noise intensity  $p$  beyond which the values of  $b_m$  reach a *plateau*. A justification for this behavior is that when the noise intensity is very small, the system needs longer trajectories – higher  $m$  – to explore the ball of radius  $\epsilon$ . This gives an implicit criterion for the selection of the bin length  $m$  needed to observe reliable results and it tells that one should be careful in applications where the intensity of the observational noise is small compared to the scale of the dynamics. This analysis is confirmed by the results obtained for the Lozi map:

$$\begin{cases} x_{t+1}^{(1)} = x_t^{(2)} + 1 - a|x_t^{(1)}| \\ x_{t+1}^{(2)} = bx_t^{(1)} \end{cases} \quad (4.2)$$

for which we consider the classical set of parameter  $a = 1.7$  and  $b = 0.5$ . Young [35] proved the existence of the SRB measure for the Lozi map and found the value  $D = 1.40419$  for the Hausdorff dimension of the measure by computing the Lyapunov exponents and using a Kaplan–Yorke like formula. The experiments are exactly the same described for the Cantor set and the results, obtained by using the value of  $D$  found by York, are presented in Fig. 5 for the values of  $m = 1000, 10\,000, 30\,000$ . Agreement with the theoretical behavior of  $b_m$ , as expected from Eq. (3.3) represented in Fig. 5 by the solid straight lines, is found for small values of  $p$  which means for large  $\epsilon$ . For the Lozi map the convergence is worse than for the Cantor case. This phenomenon has been already observed in [2] and, up to now, there is not a clear explanation for it. However, by scanning numerically the  $(m, p)$  space, one can infer the intervals of such parameters such that the  $b_m$  converge towards



**Fig. 4.** Normalizing sequence  $b_m$  vs. intensity of the noise in terms of  $p$  (we recall that  $\epsilon = 10^{-p}$ ) for the Cantor IFS (Eq. (4.1)). Dashed error-bars display the average of  $b_m$  over 30 realizations and the standard deviation of the sample. Solid lines indicate the theoretical values. The blue, red and magenta curves respectively refers to  $m = 1000, 10\,000, 30\,000$ , the points  $z$  randomly chosen on the attractor.  $n = 1000$  for all the cases considered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Normalizing sequence  $b_m$  vs. intensity of the noise in terms of  $p$  (we recall that  $\epsilon = 10^{-p}$ ) for the Lozi map (Eq. (4.2)). Dashed error-bars display the average of  $b_m$  over 30 realizations and the standard deviation of the sample. Solid lines indicate the theoretical values. The blue, red and magenta curves respectively refers to  $m = 1000, 10\,000, 30\,000$ . The points  $z$  are randomly chosen on the attractor.  $n = 1000$  for all the cases considered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the asymptotic results: for example, when  $m = 30\,000$ , an order of noise intensity  $p \leq 3$  is needed to get convergence to the predicted theoretical values. Since there is only a limited range of  $p$  such that the  $b_m$  convergence to the prediction of Eq. (3.3), one has to take extreme care on using the results obtained with this method to estimate fractal dimensions. A good strategy to overcome this problem is to discard the value of  $b_m$  which show no dependence on  $p$  and check that the remaining points are sufficient to perform a linear fit of  $b_m$  vs.  $p$ .

The possibility of computing fractal dimensions by using a random perturbation of a dynamical system is an interesting fact: when random perturbations are applied, one usually expects the orbit to explore asymptotically the ambient space and this normally hide the fractal property of the measure. Instead, as we said at the beginning of this section, for reasonable choices of  $m$  and  $p$ ,

one gets accurate information on the fractal dimension by fitting the  $D$  parameter of Eq. (3.3). This finding opens new questions, for instance if it is possible to get analogous formulas in the case of random transformations. An application can be to adapt this theory for multi-scale systems whose description is usually made through an approximation of a dynamic consisting of a deterministic and a stochastic components. For example, by tuning the strength of the stochastic components, one can study how the noise affects the structure of the deterministic dynamics; these aspects will be explored in forthcoming works.

### Acknowledgments

SV was supported by the ANR-Project *Perturbations*, by the PICS (Projet International de Coopération Scientifique), Propriétés statistiques des systèmes dynamiques déterministes et aléatoires, with the University of Houston, no. PICS05968 and by the projet MODE TER COM supported by *Région PACA, France*. SV thanks support from FCT (Portugal) project PTDC/MAT/120346/2010. The authors acknowledge Jean-René Chazottes, Jorge Milhazes Freitas and Paul Manneville for useful discussions. DF and SV acknowledge the Newton Institute in Cambridge where this work was completed during the program *Mathematics for the Fluid Earth*.

### References

- [1] H. Aytaç, J.-M. Freitas, S. Vaienti, Laws of rare events for deterministic and random dynamical systems, *Trans. Amer. Math. Soc.* (2012) in press.
- [2] V. Lucarini, D. Faranda, G. Turchetti, S. Vaienti, Extreme value distributions for singular measures, *Chaos* 22 (2012) 023135.
- [3] S.P. Lalley, A.B. Nobel, Denoising deterministic time series, *Dyn. Partial Differ. Equ.* 3 (4) (2006) 259–279.
- [4] K. Judd, Chaotic-time-series reconstruction by the Bayesian paradigm: right results by wrong methods, *Phys. Rev. E* 67 (2003) 026212.
- [5] K. McGoff, S. Mukherjee, N.S. Pillai, Statistical inference for dynamical systems: a review, 2012. arXiv:1204.6265v1 [math.ST].
- [6] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, second ed., Cambridge University Press, 2004.
- [7] S.P. Lalley, Removing the noise from chaos plus noise, in: *Nonlinear Dynamics and Statistics* (Cambridge, 1998), Birkhäuser, Boston, MA, 2001, pp. 233–244.
- [8] S.P. Lalley, Beneath the noise, *chaos, Ann. Statist.* 27 (2) (1999) 461–479.
- [9] T. Schreiber, Extremely simple nonlinear noise-reduction method, *Phys. Rev. E* 47 (1993) 2401–2404.
- [10] T. Schreiber, P. Grassberger, A simple noise-reduction method for real data, *Phys. Lett. A* (1991) 411–418.
- [11] M. Strumik, W.M. Macek, S. Redaelli, Discriminating additive from dynamical noise for chaotic time series, *Phys. Rev. E* 72 (2005) 036219.
- [12] C. Maldonado, Fluctuation bounds for chaos plus noise in dynamical systems, *J. Stat. Phys.* 148 (2012) 548–564.
- [13] D. Faranda, S. Vaienti, A recurrence based technique for detecting genuine extremes in instrumental temperature records, *Geophys. Res. Lett.* 40 (21) (2013) 5782–5786.
- [14] P. Collet, Statistics of closest return for some non-uniformly hyperbolic systems, *Ergodic Theory Dynam. Systems* 21 (2) (2001) 401–420.
- [15] A.-C.-M. Freitas, J.-M. Freitas, M. Todd, Hitting time statistics and extreme value theory, *Probab. Theory Related Fields* 147 (3–4) (2010) 675–710.
- [16] A.-C.-M. Freitas, J.-M. Freitas, M. Todd, Extreme value laws in dynamical systems for non-smooth observations, *J. Stat. Phys.* 142 (1) (2011) 108–126.
- [17] A.-C.-M. Freitas, J.-M. Freitas, M. Todd, The extremal index, hitting time statistics and periodicity, *Adv. Math.* 231 (5) (2012) 2626–2665.
- [18] C. Gupta, Extreme-value distributions for some classes of non-uniformly partially hyperbolic dynamical systems, *Ergodic Theory Dynam. Systems* 30 (3) (2010) 757–771.
- [19] D. Faranda, J.M. Freitas, V. Lucarini, G. Turchetti, S. Vaienti, Extreme value statistics for dynamical systems with noise, *Nonlinearity* 26 (2012) 2597.
- [20] A.-C.-M. Freitas, J.-M. Freitas, On the link between dependence and independence in extreme value theory for dynamical systems, *Statist. Probab. Lett.* 78 (9) (2008) 1088–1093.
- [21] D.-E. Knuth, *The Art of Computer Programming. Vol. 2*, Addison-Wesley, 2006.
- [22] M. Thiel, M.-C. Romano, J. Kurths, R. Meucci, E. Allaria, F.-T. Arecchi, Influence of observational noise on the recurrence quantification analysis, *Physica D* 171 (3) (2002) 138–152.
- [23] T. Miyano, Time series analysis of complex dynamical behavior contaminated with observational noise, *Internat. J. Bifur. Chaos* 6 (11) (1996) 2031–2045.
- [24] D. Faranda, M.-F. Mestre, G. Turchetti, Analysis of round off errors with reversibility test as a dynamical indicator, *Internat. J. Bifur. Chaos* 22 (09) (2012).
- [25] D. Faranda, V. Lucarini, G. Turchetti, S. Vaienti, Numerical convergence of the block-maxima approach to the Generalized Extreme Value distribution, *J. Stat. Phys.* 145 (5) (2011) 1156–1180.
- [26] S. Lalley, A.-B. Nobel, Denoising deterministic time series, *Dyn. Partial Differ. Equ.* 3 (4) (2006) 259–279.
- [27] P.C. Hemmer, The exact invariant density for a cusp-shaped return map, *J. Phys. A* 17 (5) (1984) L247.
- [28] Y. Pomeau, P. Manneville, Intermittent transition to turbulence in dissipative dynamical systems, *Comm. Math. Phys.* 74 (2) (1980) 189–197.
- [29] C. Liverani, B. Saussol, S. Vaienti, A probabilistic approach to intermittency, *Ergodic Theory Dynam. Systems* 19 (1999) 671–685.
- [30] J. Alves, J. Freitas, S. Luzzatto, S. Vaienti, From rates of mixing to recurrence times via large deviations, *Adv. Math.* 228 (2011) 1203–123.
- [31] G. Cristadoro, N. Haydn, Ph. Marie, S. Vaienti, Statistical properties of intermittent maps with unbounded derivative, *Nonlinearity* 23 (2010) 1071–1096.
- [32] L.-S. Young, Dimension, entropy and Lyapunov exponents, *Ergodic Theory Dynam. Systems* 2 (1982) 109–124.
- [33] J.-P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Modern Phys.* 57 (1985) 617–656.
- [34] D. Faranda, X. Leoncini, S. Vaienti, Mixing properties in the advection of passive tracers via recurrences and extreme value theory, *Phys. Rev. E* 89 (2014) 052901-1–052901-7.
- [35] L.S. Young, Bowen–Ruelle measures for certain piecewise hyperbolic maps, *Trans. Amer. Math. Soc.* (1985) 41–48.