Extreme events in bimodal systems

S. C. Nicolis
Mathematics Department, Uppsala University, P.O. Box 480 SE-751 06 Uppsala, Sweden

C. Nicolis
Institut Royal Météorologique de Belgique, 3 Avenue Circulaire, 1180 Brussels, Belgium

(Received 30 July 2008; published 29 September 2008)

The extreme value statistics of systems possessing a two-hump probability density of the relevant variable, in which the left peak is more pronounced than the right one, is studied. It is shown that systems of this type display a nontrivial transient behavior in the form of anomalous fluctuations around the mean, for certain (finite) ranges of observational time windows. The results are illustrated on independent identically distributed random variables, systems possessing two locally stable states and subjected to additive white noise, and dynamical systems in the regime of deterministic chaos.

DOI: 10.1103/PhysRevE.78.036222 PACS number(s): 05.45.−a, 02.50.−r, 05.20.−y

I. INTRODUCTION

The study of extreme events is a highly developed branch of mathematics and is widely recognized to be of paramount importance in a variety of contexts, from hydrology to engineering to finance [1]. In its classical setting it is formulated entirely in terms of three universal types of probability distributions attracting different families of stochastic processes, depending on the way the cumulative distribution of the process $F(x)$ behaves near the upper boundary of the domain of variation of the relevant variable [2].

Central to the above remarkable universality is the passage to the asymptotic limit $n \to \infty$ of an infinite time observational window. Now the universality achieved thanks to this passage is at the expense of erasing information on the variability of extreme value related properties, as captured by their fluctuations around the asymptotic means. The situation is somewhat similar to early studies of phase transitions or chaos theory where by focusing offhand on the thermodynamic limit or on the invariant measure, key aspects related to the approach to criticality or to possible inhomogeneities in the structure of the underlying attractor and of the Lyapunov spectrum were overlooked.

In the present work we report results pertaining to the transient behavior of extremes when finite observational windows are considered, a case expected to be of high relevance in most of real world situations. We focus on the class of systems characterized by a bimodal probability distribution, which is known to encompass a wide spectrum of problems of interest in, among others, fluid mechanics, optics, electrical engineering, chemical kinetics, and atmospheric dynamics [3]. We first consider the case of independent identically distributed random variables and show that for such systems the variance of the relevant observable exhibits under well-defined conditions a maximum, thereby setting limits in the relevance of the information contained in the mean values afforded by classical theory. This property, which turns out to be generic, is illustrated in the subsequent sections on several types of more intricate dynamical systems, both stochastic and deterministic. The main conclusions are summarized in the last section.

II. INDEPENDENT IDENTICALLY DISTRIBUTED BIMODAL RANDOM VARIABLES

Consider an observable x $(a \leq x \leq b)$ whose probability density $P(x)$ exhibits two well-separated local maxima at $x = x_1$ and $x = x_2$ such that $x_1 < x_2$ and $P(x_1) > P(x_2)$. The associated cumulative distribution, $F(x) = \text{Prob}(x' \leq x)$ will start having small values for x below x_1, will possess a first inflexion point at x_1 after which its value will be subjected to a more or less abrupt increase, and will finally present a further inflexion point at x_2 before leveling off at unity for x values beyond x_2. We assume that the successive values $x^{(1)}, \ldots, x^{(N)}$ as x evolves in time, taken to be separated by a fixed observational window τ are independent identically distributed random variables. The (cumulative) probability $G_n(x)$ of the largest value x found in a subsequence $x^{(1)}, \ldots, x^{(n)}$ of the full time series is then $G_n(x) = F^n(x)$. Clearly, as n is increased, the values of $G_n(x)$ for x below x_2 will be gradually depressed and $G_n(x)$ will be increasingly displaced towards the upper boundary, the corresponding probability density $\rho_n(x)$ being increasingly closer to a δ peak concentrated on this boundary. As a corollary, the mean $\overline{x}_n$ of x will tend to the upper boundary and its variance $\overline{\delta x_n^2}$ around $\overline{x}_n$ will tend to zero. Now, if the initial $(n=1)$ peaks happen to be narrow around the values x_1 and x_2, the variance of the distribution $P(x)$ will also be small. Under these conditions then the variance starts $(n=1)$ and ends (as n becomes large) being small, and one may legitimately expect that there will exist some intermediate n for which it will go through a maximum. At this point the value of $\overline{x}_n$—one of the principal predictors in the theory of extremes—will be subjected to a maximum uncertainty, and predictions based on averages will have to be complemented by information pertaining to the fluctuations.

A simple illustration capturing the essence of the above ideas is provided by a distribution in the form of two delta peaks at x_1 and x_2 of weights a and $1-a$, respectively, with $1/2 < a < 1$

$$P(x) = a\delta(x - x_1) + (1 - a)\delta(x - x_2) \qquad (1)$$

the corresponding cumulative probability distribution being a step function with two discontinuous jumps at x_1 and x_2:

$$F(x)=0, \quad 0<x<x_1,$$
$$=a, \quad x_1<x<x_2,$$
$$=1, \quad x_2<x<b. \tag{2}$$

In this setting $G_n(x)=F^n(x)$ can be determined straightforwardly. By construction, it will keep its step like form the difference with Eq. (2) being that the intermediate level value a will now be a^n. Differentiating with respect to x one obtains the associated probability density

$$\rho(n)=\frac{dG_n(x)}{dx}=a^n\delta(x-x_1)+(1-a^n)\delta(x-x_2). \tag{3}$$

The mean and the variance of x can now be evaluated straightforwardly. One finds

$$\overline{x_n}=a^n x_1+(1-a^n)x_2, \tag{4a}$$

$$\overline{\delta x_n^2}=(x_2-x_1)^2(a^n-a^{2n}) \tag{4b}$$

from which the limiting behavior $\overline{x_n}\to x_2$ and $\overline{\delta x_n^2}\to 0$ as $n\to\infty$ conjectured above follows. Equations (4) allows us, however, to go one step further and investigate the behavior as a function of a and n. Specifically, (i) $\overline{x_n}$ is a monotonically increasing function of n for given a, and monotonically decreasing function of a for fixed n, (ii) $\overline{\delta x_n^2}$ exhibits a maximum with respect to n for given $a(a>1/2)$ as well as with respect to a for given n for $a^n=1/2$ or

$$n_{\max}=\ln 2/\ln\left(\frac{1}{a}\right) \tag{5}$$

independent of x_1 and x_2. As a gets closer to unity $n_{\max}$ increases, the value of $\overline{\delta x_n^2}$ itself at maximum being an increasing function of the distance separating the peaks of $P(x)$,

$$\overline{\delta x^2}_{n,\max}=\frac{(x_2-x_1)^2}{4}. \tag{6}$$

Notice that at $n_{\max}$ the two peaks of $\rho_n(x)$ [Eq. (3)] around x_1 and x_2 have equal weights.

A better representation of a generic distribution consisting of two well-defined peaks separated by a deep minimum is provided by two square pulses extending over an interval ϵ on either side of the points x_1 and x_2:

$$P(x)=a\frac{1}{2\epsilon}\theta[x-(x_1-\epsilon)]\theta(x_1+\epsilon-x)$$
$$+(1-a)\frac{1}{2\epsilon}\theta[x-(x_2-\epsilon)]\theta(x_2+\epsilon-x)$$
$$\left(\epsilon<\frac{x_2-x_1}{2}\right). \tag{7}$$

Following the same procedure as in Eq. (3) one obtains straightforwardly a cumulative distribution $G_n(x)=F^n(x)$ in the form of a piecewise differentiable function

$$G_n(x)=0, \quad x<x_1-\epsilon,$$
$$=\left(\frac{a(x-x_1+\epsilon)}{2\epsilon}\right)^n, \quad x_1-\epsilon<x<x_1+\epsilon,$$
$$=a^n, \quad x_1+\epsilon<x<x_2-\epsilon,$$
$$=\left(a+\frac{(x-x_2+\epsilon)(1-a)}{2\epsilon}\right)^n, \quad x_2-\epsilon<x<x_2+\epsilon,$$
$$=1, \quad x>x_2+\epsilon. \tag{8}$$

The mean $\overline{x_n}$ and variance $\overline{\delta x_n^2}$ of the associated probability density can be evaluated explicitly. The expressions, which are rather cumbersome, have the general structure

$$\overline{x_n}=a^n x_1+(1-a^n)x_2+\epsilon C_1(a,n), \tag{9a}$$

$$\overline{\delta x_n^2}=(a^n-a^{2n})(x_2-x_1)^2+\epsilon D_1(a,n)+\epsilon^2 D_2(a,n). \tag{9b}$$

As an example,

$$C_1(a,2)=\frac{1}{3}(1-2a+2a^2),$$

$$D_1(a,2)=\frac{4}{3}(x_2-x_1)(a^3+a^4),$$

$$D_2(a,2)=\frac{2}{9}(1+2a-4a^2+4a^3-2a^4). \tag{10}$$

We recognize in the ϵ-independent part of Eqs. (9) the expressions of $\overline{x_n}$ and $\overline{\delta x_n^2}$ for the two delta peak case, Eqs. (4). The presence of correction terms in ϵ and ϵ^2 entails that, contrary to Eqs. (4), the state of equipartition (which still corresponds to $a^n=1/2$) does not coincide here with the (a,n) values yielding the extremum of $\overline{\delta x_n^2}$. For instance, in the $n=2$ case Eqs. (10) yield for $x_1=-5$, $x_2=5$, and $\epsilon=2$ an extremum of $\overline{\delta x_n^2}$ for $a\approx 0.699$, which is slightly less than the value $a^2=1/2$ or $a=\sqrt{2}/2$ and corresponds to a total weight of the left pulse equal to 0.489 rather than 1/2. Notice that $a_{\max}$ increases as $|x_2-x_1|$ increases.

Coming back to expressions (9), numerical evaluation of $\overline{\delta x_n^2}$ as a function of n for various a's complemented by direct simulation of the process, viewed as a superposition of two uniform noises, confirms this view. Figure 1(a) depicts the main result. The dependence of $n_{\max}$ on the width ϵ of the original pulses for a given a value, shown in Fig. 1(b), displays a thresholdlike behavior such that the deviation from $a^n=1/2$ relation begins to show up beyond a (rather substantial) value of ϵ. The variance $\overline{\delta x^2}_{n,\max}$ itself is a decreasing function of ϵ (not shown). Notice that under the same conditions $\overline{x_n}$ increases monotonically with n, as expected.

The above conclusions extend to the more generic case where $P(x)$ is the superposition of two narrow Gaussians centered on x_1 and x_2,

$$P_1(x)=a\phi_1(x)+(1-a)\phi_2(x), \quad -\infty<x<\infty \tag{11a}$$

with

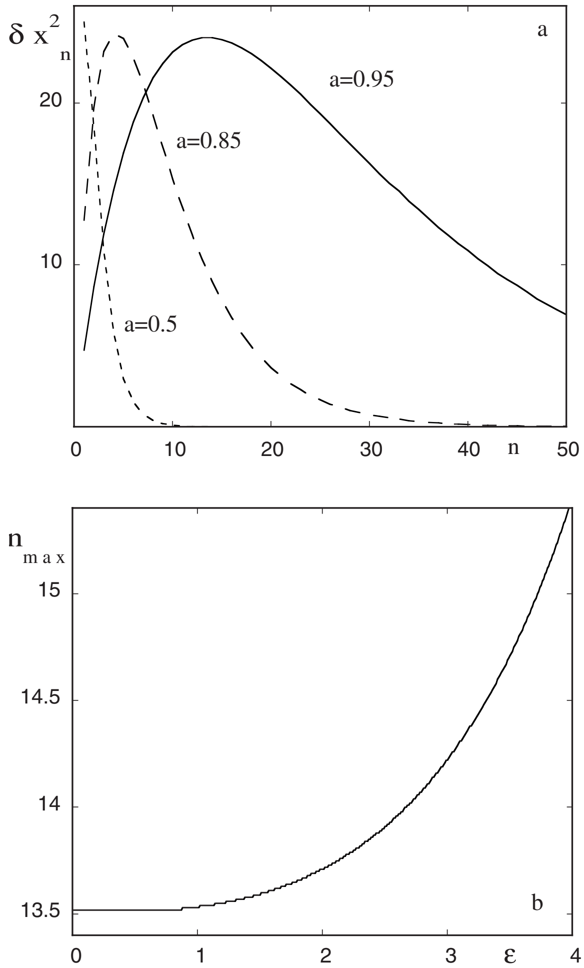

FIG. 1. (a) Variance of extremes versus the time window n in the case of an independent bimodal random variable distributed according to two square pulses centered on $x_1=-5$ and $x_2=5$ extending over an interval $\epsilon=0.25$ on either side of x_1 and x_2 as obtained by direct simulation of the process for three different weights of the leftmost pulse. Number of realizations for the statistical averaging is 10^6. (b) Dependence of the time window n for which the variance displays a maximum on the width of the pulse ϵ as obtained from the analytic expression (9b). Parameter values as in (a) with $a=0.95$.

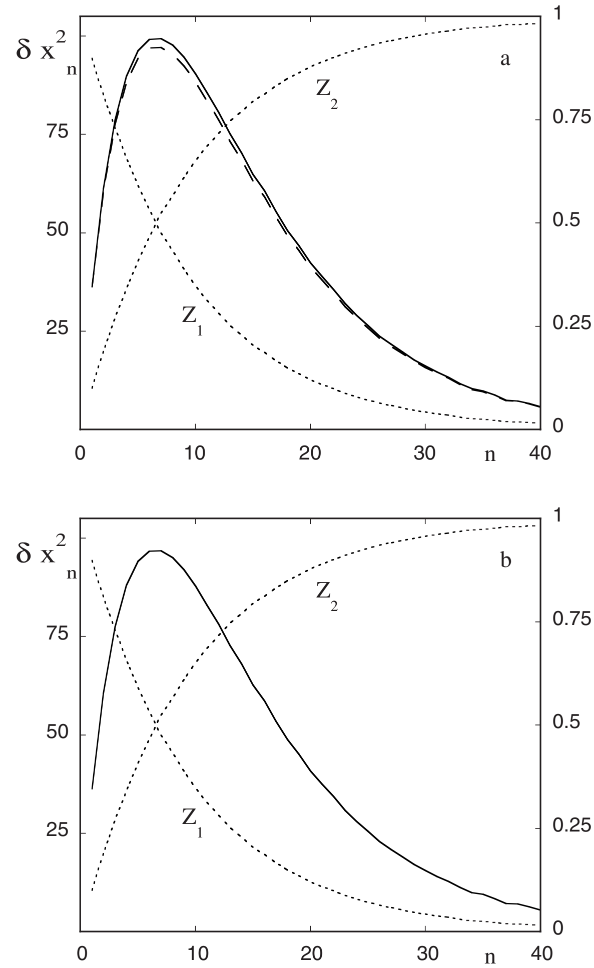

FIG. 2. (a) As in Fig. 1(a) but in the case of two narrow Gaussians of equal variances, $\sigma^2=10^{-3}$ (full line) and $\sigma^2=5\times10^{-2}$ (dashed line), centered on $x_1=-10$ and $x_2=10$ with $a=0.9$ [Eq. (11a)]. Dotted lines stand for the dependence of the probability masses Z_i around x_1 and x_2 on the time window n. Number of realizations is 10^5. (b) As in (a) but variances of the Gaussians are unequal $\sigma_1^2=5\times10^{-2}$ and $\sigma_2^2=10^{-3}$.

$$\phi_i = \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{(x-x_i)^2}{2\sigma_i^2}\right)\ (i=1,2). \tag{11b}$$

Figures 2(a) and 2(b) depict the dependence of $\overline{\delta x_n^2}$ on n for the two cases of equal and unequal variances σ_1^2 and σ_2^2, respectively. In both cases a clearcut maximum, which in the case of Fig. 2(a) tends to increase with decreasing variance values, is observed. The dotted lines in the same figures represent the n dependence of the probability masses Z_1 and Z_2 in the intervals $[-\infty,0]$ and $[0,\infty]$, respectively. As can be seen the value $n_{\max}$ of $\overline{\delta x_n^2}$ is very close to the case of equipartition, i.e., the role of the variance (as long as it remains weak) is here less pronounced than in the previous case of two pulses. We argue that this may be due to the fact that in the case of Fig. 2 one deals with a distribution defined on an infinite support, whereas in the case of Fig. 1 the support is finite. Now, a probability density defined on an infinite support and possessing two maxima is asymptotically equivalent (as far as the computation of its moments is concerned), in the limit of small local widths around each of the maxima, to the exponential of a quartic function of x times a normalization factor. This reduces, in turn, by a steepest descent type of argument to two Gaussian peaks similar to Eq. (11a) but valid now for higher windows n as well, for which the maximum variance is indeed attained in the state of equipartition. Notice that the exponential of a quartic function is also the state that maximizes information entropy—essentially the delocalization in state space—at given second and fourth moment values.

III. STOCHASTICALLY FORCED DYNAMICAL SYSTEMS

We now place ourselves in a dynamical perspective, in which bimodality and extreme value properties are generated by an underlying evolution law. Specifically, we consider an overdamped one-variable system driven by a potential $V(x)$ and subjected also to an additive Gaussian white noise. Assuming that $V(x)$ possesses two minima at x_1 and x_2 separated by a maximum located (without loss of generality) at $x=0$, the evolution equation takes the form [4]

$$\frac{dx}{dt} = -\frac{\partial V}{\partial x} + F(t). \tag{12}$$

Here

$$V = \lambda\frac{x^2}{2} - \mu\frac{x^3}{3} + \frac{x^4}{4} \tag{13a}$$

with

$$\lambda = x_1x_2, \quad \mu = x_1 + x_2, \tag{13b}$$

and

$$\langle F(t)\rangle = 0, \quad \langle F(t)F(t')\rangle = q^2\delta(t-t'), \tag{13c}$$

where the brackets denote average over the different realizations of the noise.

To secure the bistable character of the potential we take $x_1<0$ and $x_2>0$. The relative stability of x_1 and x_2, also reflected by the relative magnitudes of the peaks of the invariant probability density

$$P(x) = Z^{-1}\exp\left[-\frac{V(x)}{q^2/2}\right] \tag{14}$$

is determined by the distances of x_1 and x_2 from $x=0$. We here choose $|x_1|>x_2$, which guarantees that state x_1 is more probable than x_2.

From the standpoint of dynamics one is in the presence of two types of processes, characterized by widely separated time scales: a local, small scale diffusion around x_1 and x_2 whose characteristic time is

$$\tau_i^{-1} = \left(\frac{d^2V}{dx^2}\right)_i = 3x_i^2 - 2\mu x_i + \lambda \tag{15a}$$

and a sequence of transitions between x_1 and x_2. The mean sojourn time around each of the states x_i prior to a transition is given by the Kramers formula

$$\tau_K = \frac{\pi}{\sqrt{-V''(0)V''(x_i)}}\exp\frac{[V(0)-V(x_i)]}{q^2/2}, \tag{15b}$$

where $V(0)-V(x_i)$ is the potential barrier separating each x_i from the other locally stable state and $V''(0)$, $V''(x_i)$ are the second derivatives of the potential V evaluated at $x=0$ and $x=x_i$, respectively.

We come now to extreme values and their probabilistic properties. Solving the Langevin Eq. (12) numerically we generate a time series of the variable x and monitor its successive values at times $t=\tau, 2\tau, \ldots, N\tau$. For each given (long) such series we identify the largest value found in successively larger windows n along the series and deduce its probability density and its first few moments. Figure 3 summarizes the main result of this evaluation for parameter values $x_1=-1$, $x_2=1/2$, $q^2=0.08$, $\tau=1$ time unit. As can be seen the variance displays a maximum at $n_{\max}\tau\approx 226$ time units,

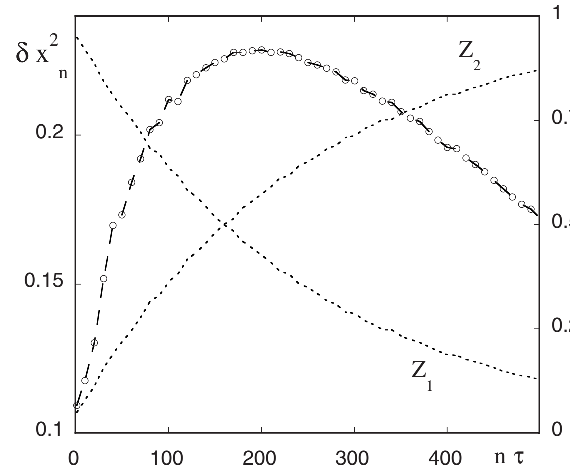

FIG. 3. As in Fig. 2(a) but for a stochastically forced dynamical system, Eq. (12). Parameter values are $x_1=-1$, $x_2=1/2$, $q^2=0.08$, and $\tau=1$ time unit yielding a probability mass around state x_1 for $n=1$ $Z_1\approx 0.96$. The number of realizations is 2×10^4.

whereas the equality of the probability masses in the intervals $(-\infty,0)$ and $(0,\infty)$ is achieved for $n_{\max}\tau\approx 160$ time units. Under the same conditions the correlation and Kramers times [Eqs. (15a) and (15b)] are $\tau_1^{-1}=3/2$, $\tau_2^{-1}=3/4$, $\tau_K=234$ time units. The similarity between the values of $n_{\max}\tau$ and τ_K can be understood qualitatively by noticing that the crossing of the barrier is a necessary condition for the transfer of probability mass towards higher values of x as the window n is increased.

IV. DETERMINISTIC DYNAMICAL SYSTEMS

Fundamentally, the laws governing the evolution of natural systems are deterministic. In the present section we investigate the transient behavior of extremes for deterministic dynamical systems generating nonlinear behavior responsible for a bimodal structure of the invariant probability distribution of a relevant variable. Previous work by the present authors and co-workers has shown that the structure of the n-time probability density $\rho_n(x)$ of deterministic systems presents some fundamental differences from those featured in classical theory, in the form of distinct plateaus formed at discrete (generally n-dependent) sets of values [5,6]. Here we focus on the specific role of bimodality of the probability density $P(x)$ of the process in the behavior of extremes.

The systems to be considered are chosen among the class of one-dimensional chaotic maps subjected (in order to achieve bimodality) to a multiplicative periodic forcing. They are constructed around a skeleton consisting of an antisymmetric discontinuous tent map in the interval $-1\leqslant x\leqslant 1$, whose parameters are subsequently modified in order to control the relative values of the probability masses in the left and right subintervals separated by $x=0$. Specifically, setting

$$f_+(x_k) = [a_1 + d_1\sin(2\pi\omega k)](0.5 - x_k),$$
$$f_-(x_k) = [a_2 - d_2\sin(2\pi\omega k)](0.5 + x_k) \tag{16}$$

the map is defined by [7]

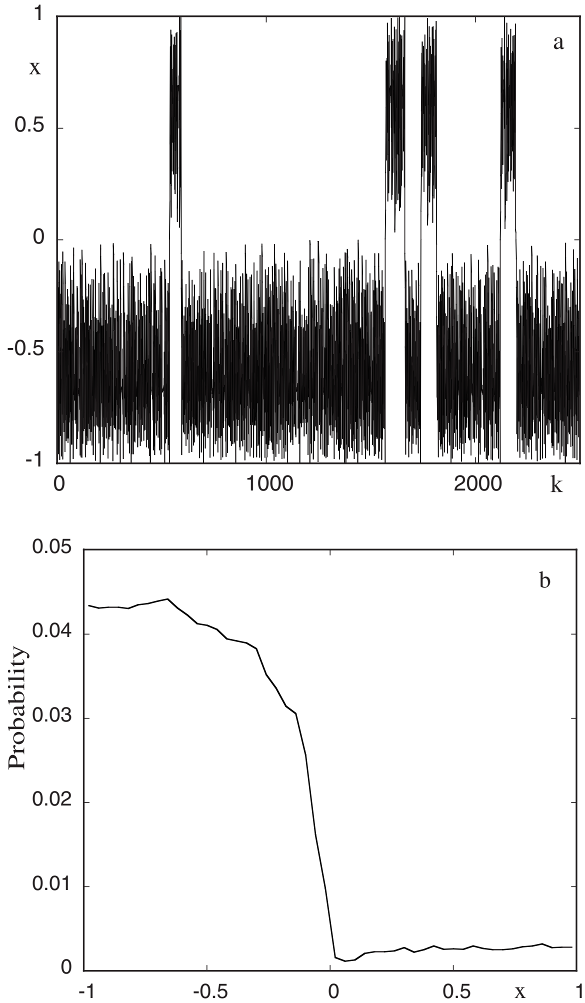

FIG. 4. Time evolution (a), and probability distribution of variable x (b) of the map of Eqs. (16) and (17) with parameter values $a_1=a_2=1.9$, $d_1=0.25$, $d_2=0.12$, and $\omega=0.1$.

$$x_{n+1} = -1 - f_-(x_k), \quad -1 \leqslant x < -0.5$$

$$= -1 + f_-(x_k), \quad -0.5 < x \leqslant 0$$

$$= 1 - f_+(x_k), \quad 0 < x \leqslant 0.5$$

$$= 1 + f_+(x_k), \quad 0.5 < x \leqslant 1. \tag{17}$$

Taking $a_1=a_2$ and d_1 sufficiently larger than d_2 leads, for a wide range of values of the frequency ω, to a clearcut bimodality in which the left part of the interval dominates as seen in Figs. 4(a) and 4(b), the mean sojourn time in this region being about 1050 time units. Coming next to the behavior of extremes, Fig. 5 depicts the cumulative probability $F_n(x)$ for increasing windows n. We recognize the piecewise differentiable structure advanced in the beginning of this section as the principal signature of the deterministic character of the dynamics. As one might expect the discontinuities are smeared out at the level of statistical averages. In particular, as seen in Fig. 6 the variance $\overline{\delta x_n^2}$ displays a smooth dependence on the window. One again sees in this Figure the existence of a maximum at some window value n_{max} of about 650 time units, close to the n value for which the equipartition of the probability masses Z_1 and Z_2 is achieved. We notice that $n_{max}\tau$ and the mean sojourn time are here significantly different, although they still are of the same order of magnitude. This difference with what was found in Sec. III is due to the fact that in a deterministic system there is no clearcut distinction between "systematic" and "random" behavior, owing to the presence of persistent correlations. Similar conclusions as above are reached for the much simpler class of deterministic systems showing periodic behavior, for which explicit analytic solutions can also be constructed. It suffices for this to tune, through appropriate parameter values, the parts of the overall periodicity that the system spends in different selected ranges of values of the variable x.

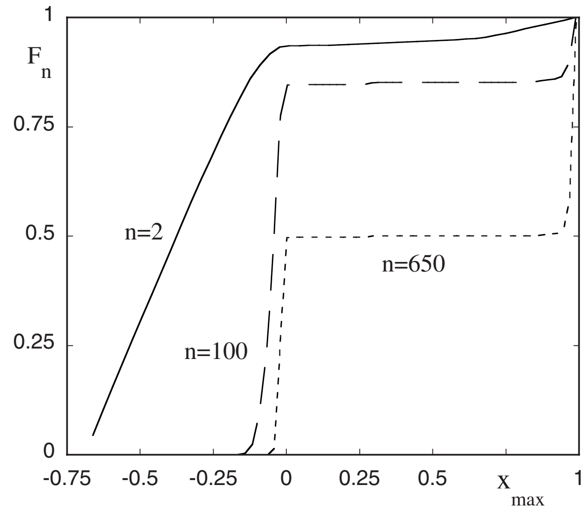

FIG. 5. Cumulative probability distribution for the extreme values of system (16) and (17) for increasing time windows n deduced numerically using 10^6 realizations. Parameter values as in Fig. 4.

V. CONCLUSIONS

In this work we analyzed a class of systems showing nontrivial transient behavior in their extreme value properties, in

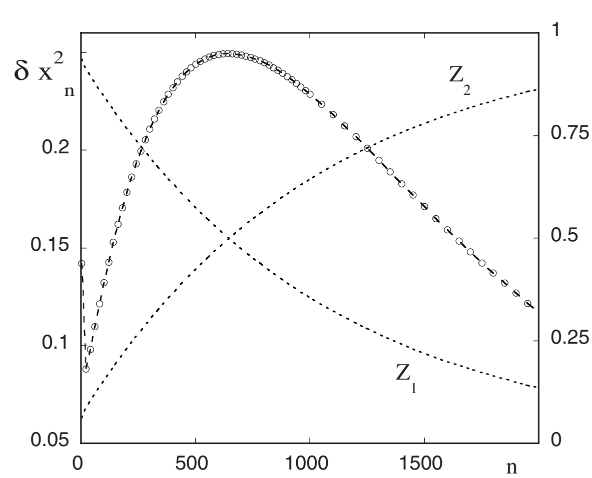

FIG. 6. As in Fig. 2(a) but for system (16) and (17). Parameter values as in Fig. 4.

the form of anomalous fluctuations for certain ranges of observational windows. These systems share the common feature of having a bimodal probability density of the relevant variable x, the leftmost peak being much more pronounced compared to the rightmost one. There are different dynamical scenarios for realizing such a distribution: the system possesses two locally stable states separated by an intermediate unstable one and is subjected to an additive white noise; or it operates in the regime of nonlinear oscillations or deterministic chaos, in which the attractor is highly nonuniform in the form of two "hot spots" monopolizing much of the probability mass.

As is well known, as the observational window is increased the probability mass of the n-fold density $\rho_n(x)$ tends to be displaced towards the upper boundary of the variable x. The principal role of bimodality is to postpone this process, by inducing intermediate regimes in which the probability mass of $\rho_n(x)$ in the range of moderate values of x remains substantial for observational windows that may be large and physically relevant. For such regimes mean value-related predictions need to be complemented with information pertaining to fluctuations.

A straightforward extension of this work would be to consider n-modal $(n>2)$ systems, where a further postponement for reaching the asymptotic regime can be expected. The case of multivariate systems would also be worth considering, since there may now be alternative (and competing) pathways for transitions between states.

ACKNOWLEDGMENTS

This work is supported, in part, by the Human Frontiers Science Programme and the European Commission under Contract No. 12975 (NEST).

[1] E. Montroll and W. Badger, *Introduction to Quantitative Aspects of Social Phenomena* (Gordon and Breach, New York, 1974).

[2] P. Embrechts, P. Küppelberg, and T. Mikosch, *Modelling Extremal Events* (Springer, Berlin, 1999).

[3] G. Nicolis, *Introduction to Nonlinear Science* (Cambridge University Press, Cambridge, 1995).

[4] N. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1981).

[5] V. Balakrishnan, C. Nicolis, and G. Nicolis, J. Stat. Phys. **80**, 307 (1995); T. Goychuk, J. Casado-Pascual, M. Morillo, J. Lehmann, and P. Hanggi, Phys. Rev. Lett. **97**, 210601 (2006).

[6] C. Nicolis and S. C. Nicolis, Europhys. Lett. **80**, 40003 (2007).

[7] S. Sinha and B. K. Chakrabarti, Phys. Rev. E **58**, 8009 (1998).