Robust extremes in chaotic deterministic systems

Renato Vitolo,[a] Mark P. Holland, and Christopher A. T. Ferro
School of Engineering, Computing and Mathematics, University of Exeter, Exeter, Devon EX4 4QF, United Kingdom

(Received 4 June 2009; accepted 10 November 2009; published online 9 December 2009)

This paper introduces the notion of robust extremes in deterministic chaotic systems, presents initial theoretical results, and outlines associated inferential techniques. A chaotic deterministic system is said to exhibit robust extremes under a given observable when the associated statistics of extreme values depend smoothly on the system's control parameters. Robust extremes are here illustrated numerically for the flow of the Lorenz model [E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963)]. Robustness of extremes is proved for one-dimensional Lorenz maps with two distinct types of observables for which conditions guaranteeing robust extremes are formulated explicitly. Two applications are shown: improving the precision of the statistical estimator for extreme value distributions and predicting future extremes in nonstationary systems. For the latter, extreme wind speeds are examined in a simple quasigeostrophic model with a robust chaotic attractor subject to nonstationary forcing. © *2009 American Institute of Physics*. [doi:10.1063/1.3270389]

The statistical theory of extreme events aims at the probabilistic prediction of events of unusual intensity (e.g., intense rainfall/wind, insurance or financial losses). A standard approach is to model the frequency distribution of such events through so-called extreme value probability distributions. Recent theoretical work has proved that extreme value laws also hold in certain chaotic deterministic dynamical systems. How do the extreme value laws depend on the control parameters of physical systems and on the used observables? A smooth dependence has been recently observed in a model of the atmospheric jet at midlatitudes. This paper aims to make first steps toward a theory of such "robustness of extremes" and to highlight its potential usefulness for statistical inference and prediction in nonlinear systems. Two applications are shown: (1) to reduce uncertainty in statistical estimation and extrapolation beyond experimentally used control parameter values and (2) to build predictive probabilistic models for extremes in nonstationary deterministic systems. We illustrate point (1) for the Lorenz 1963 model, also showing that robustness of extremes holds in a suitable parameter range. For point (2), we examine wind speeds in the atmospheric model discussed above, which is much more complex than the Lorenz model and whose ergodic properties are unknown. These ideas are potentially useful for the study of extremes and their trends in the climate system.

I. INTRODUCTION

The response of physical systems to variations in external control parameters is the subject of many fields of research, including Kolmogorov–Arnold–Moser (KAM) theory, bifurcation theory, and phase transitions; see, e.g., Refs. 1–3 and references therein. For nonlinear systems, bifurcations are associated with qualitative changes in dynamical behavior, which can be sudden and dramatic in the case of global bifurcations or "crises," see, e.g., Refs. 1, 4, and 5. However, chaotic deterministic systems also often display smooth responses to parameter variation. Notions of robust chaos and statistical stability have been developed to understand this latter phenomenon. Robust chaos is defined by the absence of periodic windows and coexisting attractors in some neighborhood of the parameter space for a given system,[6] which means that chaotic behavior is not destroyed by small parameter variations. This is a qualitative feature and is found in several systems from neural networks[7] to (piecewise) smooth maps,[8,9] see Elhadj and Sprott[10] and references therein. A related notion is that of a robust strange attractor (entailing C^1-openness of the defining conditions). Tucker[11] showed that the system of Lorenz[12] has a robust strange attractor for a suitable range of parameter values. Moreover, all robust strange attractors containing an equilibrium are singular hyperbolic (see Morales *et al.*[13] for a definition and proof), the attractor of the Lorenz[12] model being a main example. Statistical stability is a more quantitative notion entailing continuity of the Sinai–Ruelle–Bowen (SRB) measure (and associated metric entropies) with respect to the parameters. Statistical stability has been proved for the logistic and Hénon families when parameters are restricted to a nowhere dense subset and for other classes of one-dimensional (1D) maps; see Refs. 14 and 15 and references therein.

In this paper we investigate how the extremal properties of chaotic deterministic systems respond to parameter variation. We consider the statistical properties of extreme values that arise when the system enters asymptotically small regions of phase space. Such extreme values can be highly significant when modeling physical systems such as the climate system. The statistical theory of extremes was originally developed for stochastic processes[16–19] but substantial progress has been made recently in transferring this theory to chaotic deterministic systems.[20–23] We make first steps to-

[a]Electronic mail: r.vitolo@exeter.ac.uk.

ward a theory of robust extremes. In particular, we formulate the notion of robust extremes, prove robustness of extremes for a simple class of systems, and demonstrate how knowledge of robust extremes can be used to improve inferences about the extreme values of complex systems. Note that the robustness of extremes depends on both short-range and long-range recurrence time statistics and therefore is not equivalent to statistical stability or robust chaos, as discussed in Sec. VI.

The classical statistical theory of extremes is briefly reviewed in Sec. II and the notion of robust extremes is formulated in Sec. III with a formal definition given in Sec. IV. Section IV discusses how to derive the dependence of a system's extremal properties on its control parameters, thereby determining whether or not the system exhibits robust extremes, and demonstrates the procedure for 1D Lorenz maps. Section V discusses the implication of these results for the three-dimensional (3D) flow of the Lorenz63 model, also illustrating lack of robustness and the link with lack of hyperbolicity. Extensions of this analysis to more complex systems are discussed in Sec. VI. Two applications demonstrating the benefits accruing from robust extremes are presented in Sec. VII. The first reduces uncertainty in estimates of the statistical properties of extreme values; the second predicts extremal properties in nonstationary systems. The applications are illustrated with extreme wind speeds from a quasi-geostrophic model.[24]

II. THE GENERALIZED EXTREME VALUE DISTRIBUTION

Extreme value theory traditionally studies the tail behavior of stochastic processes with extensive applications in hydrology, climatology, finance, and insurance.[16–18] The principal statistical model in this theory is the generalized extreme value (GEV) family of distribution functions,

$$H(x;\xi)=\begin{cases}\exp[-(1+\xi x)_+^{-1/\xi}], & \xi\neq 0\\ \exp[-\exp(-x)], & \xi=0,\end{cases} \tag{1}$$

where $w_+=\max\{w,0\}$. The GEV is justified theoretically by classical results of Fisher/Tippett and Gnedenko (see the above references). Let $\{X_j: j\geq 1\}$ be a sequence of independent, identically distributed random variables and let $M_N=\max(X_1,\ldots,X_N)$. If there exist normalizing constants $a_N>0$ and b_N such that

$$P[(M_N-b_N)/a_N\leq x]\overset{d}{\to}H(x) \quad \text{as} \quad N\to\infty \tag{2}$$

for a nondegenerate distribution function H (where $\overset{d}{\to}$ denotes convergence in distribution), then H belongs to the GEV family (1). The parameter ξ is called the *tail index* and determines the behavior of the upper tail of the GEV distribution, which is unbounded if $\xi\geq 0$ and has a finite upper end point x_+ if $\xi<0$. The tail index can be obtained analytically given the distribution function F of the X_j: indeed condition (2) holds if and only if

$$g(ut)/g(u)\to t^{-1/\xi}, \tag{3}$$

where $g(x)=1-F(x)$ if $\xi>0$ and $g(x)=1-F(x_+-1/x)$ if $\xi<0$. Condition (3) is called *regular variation*.[25] The status of the GEV distribution as the only nondegenerate limiting distribution for linearly normalized maxima of sequences of independent and identically distributed random variables extends to maxima of stationary stochastic processes under two additional conditions,[26] known as the $D(u_n)$ and $D'(u_n)$ conditions in the statistical literature. Loosely speaking, the $D(u_n)$ is a condition that requires weak mixing of the stochastic process and $D'(u_n)$ controls the tendency of large values to form clusters in time, see Ref. 26 for details.

This theory provides a statistical framework with which to make inferences about values of an observed process which fall beyond the range of the available observations. The frequency distribution of observed maxima in blocks of large but finite length N is modeled with the GEV distribution function

$$G(x;\mu,\sigma,\xi)=\exp\left[-\left(1+\xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right] \tag{4}$$

when $\xi\neq 0$, where the normalizing constants b_N and a_N are subsumed into the definition of the GEV distribution to yield the location parameter μ and scale parameter $\sigma>0$. The observed block maxima are then considered as a sample drawn from Eq. (4) and the parameters (μ,σ,ξ) are estimated by, e.g., maximum likelihood. In meteo-climatic applications, where the block length is often 1 yr, this is called the annual maximum method. A quantity of interest for planning purposes is the value x_p that has a probability p to be exceeded by the annual maximum: this is called the *return level* associated with the *return period* of $1/p$ yr. The return level is simply the quantile of Eq. (4),

$$x_p=\mu-\frac{\sigma}{\xi}\{1-[-\log(1-p)]^{-\xi}\}, \tag{5}$$

when $\xi\neq 0$. Uncertainty in the estimation of the tail index ξ can have a large impact on predicting return levels beyond the observed range or with return periods that are larger than the record length.[19]

III. ROBUST EXTREMES

To motivate the notion of robust extremes consider the model (referred to as Lorenz63 hereafter) of Lorenz,[12]

$$\dot{x}=(y-x)P, \quad \dot{y}=x(\rho-z)-y, \quad \dot{z}=xy-\beta z, \tag{6}$$

derived from the Rayleigh equations for convection in a fluid layer between two plates. Here P is the Prandtl and ρ is the Rayleigh number; see Sparrow[27] for a dynamical study. For the "classical" values $P=10$, $\beta=8/3$, and $\rho=28$, Eq. (6) has a strange attractor[11] that is robust in the sense of Morales *et al.*[13]

We study the extremes of the variable x proportional to the intensity of the convective motion.[12] Orbits of length 10^n time units are generated from Eq. (6) with both $n=5$ and $n=8$. Values of x are sampled every 0.5 time units along an orbit.[28] Maxima over blocks of 1000 time units are extracted from each series and their frequency distributions are modeled with the GEV distribution function (4), which is fitted by maximum likelihood.[18,19] The procedure is repeated for

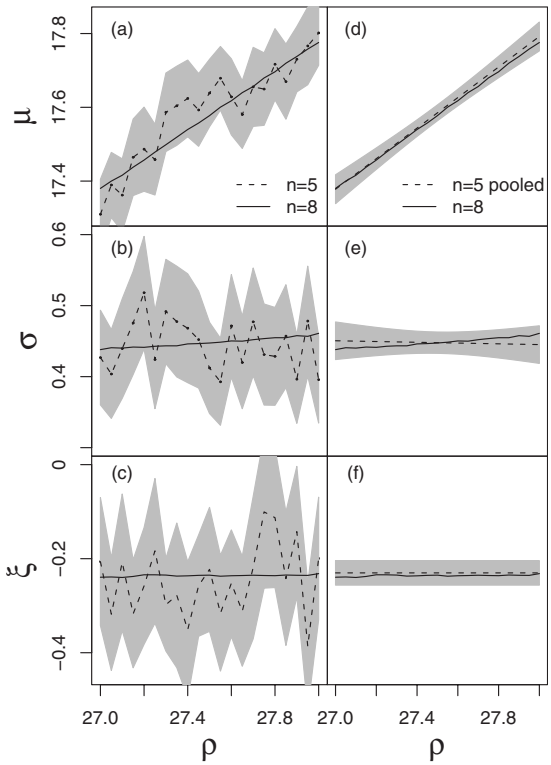

FIG. 1. [(a)–(c)] Estimates of the GEV parameters μ,σ,ξ (respectively) vs the value ρ_j used in Eq. (6) for time series length 10^n with $n=5$ (dashed) and $n=8$ solid. Approximate 95% confidence intervals are added for $n=5$ (gray shading). [(d)–(f)] Pooled estimates (dashed) and confidence intervals (gray shading) obtained with Eq. (12) are plotted together with the true values obtained for $n=8$.

each of 21 values of ρ ($\rho_j=27+0.05j$ for $j=0,1,\ldots,20$) and the estimates of the GEV parameters are plotted against ρ in Figs. 1(a)–1(c).

For the short time series (with $n=5$ corresponding to a sample of 100 block maxima) the estimates oscillate wildly around the smoothly varying "true" values obtained with $n=8$ (corresponding to 10^5 maxima). This smooth variation in the statistics of extreme values is what we term *robust extremes* in a broad, operational sense; a rigorous definition is used in Sec. IV requiring continuity of the tail index ξ with respect to the control parameters of the system.

Felici *et al.*[29] found similarly smooth behavior in a model for the atmospheric circulation at midlatitudes in the Northern Hemisphere. This is a 192-dimensional system of ordinary differential equations providing a simplified representation of relevant atmospheric physical processes (baroclinic conversion, barotropic stabilization, thermal diffusion, and viscouslike dissipation). The chaotic attractor of the model is numerically shown to be robust with respect to a parameter T_E representing baroclinic forcing.[30] The extremes of the system's total energy are found to have a smooth scaling law with respect to T_E.[29] That study was in turn motivated by a question in climate science: Are meteo-climatic extremes robust with respect to variations in atmospheric CO_2, or do they exhibit abrupt changes or "tipping points"?[31] This issue is of great relevance in climate science for prediction and adaptation purposes, as well as for the rigorous quantification of change and trends in extremes. We conjecture that the phenomenon of robust extremes may occur in many classes of systems (e.g., lasers, networks) beyond the two examples of geophysical flows considered here.

IV. ROBUST EXTREMES IN 1D LORENZ MAPS

This section contains the main theoretical results of our paper and the exposition is necessarily more technical. Suppose that the evolution of a model of interest is described by a parametrized family $f_\alpha^t:\mathbb{R}^n\to\mathbb{R}^n$ of continuous-time dynamical systems, smoothly depending on the parameter $\alpha\in\mathbb{R}$. For example, let the model be defined by the solution of a system of ordinary differential equations in the phase space $\mathbb{R}^n$. Let $\phi:\mathbb{R}^n\to\mathbb{R}$ be an observable (e.g., ϕ is the projection on x for Lorenz63 in Fig. 1 or the total energy of the system in Felici *et al.*[29]). Suppose that for fixed α the system has an attractor Λ_α: this is defined as a compact invariant subset of $\mathbb{R}^n$ which is transitive (there exists $x\in\Lambda_\alpha$ with a dense orbit) and has an open trapping basin [there exists a neighborhood U of Λ_α such that $f_\alpha^t(U)\subset U$ for all $t>0$, see Morales *et al.*[13]].

To define the statistical properties of the attractor in a physically relevant way, one must assume the existence of a SRB measure μ_α: indeed, SRB measures have basins of attraction with a positive volume, meaning that they are observable in concrete physical experiments.[32] The existence of a unique SRB measure μ_α allows a unique stationary stochastic process to be associated with the observed evolution of a deterministic system: indeed, by the invariance of μ_α the random variables $X_{j,\alpha}=\phi\circ f_\alpha^{(j-1)t_0}$, with $t_0>0$ and $j\in\mathbb{N}$, form a stationary stochastic process on the probability space $(\Lambda_\alpha,\mu_\alpha)$. As explained in Sec. II, if the normalized maxima of this stochastic process converge in distribution to a nondegenerate limit H_α and Leadbetter's conditions hold,[26] then H_α is a GEV distribution function whose tail index ξ_α is determined by regular variation in the distribution function F [condition (3)]. In this case, we say that the pair (f_α^t,ϕ) has robust extremes if ξ_α depends continuously on α.

We now show that 1D Lorenz maps have robust extremes. A C^1 map $T_\alpha:[-1,1]\backslash\{0\}\to[-1,1]$ is a Lorenz map if it satisfies (i) $|(T_\alpha^n)'(x)|\geq c\lambda^n$ for some $\lambda>1$ (provided x is not a preimage of 0); (ii) T_α is continuous except at 0, where $T_\alpha(0^-)=1$ and $T_\alpha(0^+)=1$; (iii) $T_\alpha'(x)=l(x)|x|^{\alpha-1}$ near 0, for some $\alpha\in(0,1)$ and some function $l(x)$, which is slowly varying at 0 and Hölder continuous; and (iv) T_α admits an absolutely continuous invariant measure with density $\theta_\alpha(x)=d\mu_\alpha/dx$ of type *bounded variation*. In particular the density has at most a countable number of discontinuities and its support has upper end point at $x^+=1$, where $\theta_\alpha(x)$ vanishes. For T_α, the attractor Λ_α is the whole interval $[-1,1]$.

Our proof of robustness of extremes for parametric families of 1D Lorenz maps is based on regular variation (3), which yields the actual value of the tail index ξ_α as a function of the parameter α. The two additional conditions $D(u_n)$ and $D'(u_n)$ must be verified (see Sec. II), since the stochastic process arising from observing a dynamical system is not formed by independent random variables. To this end, we use a recent result by Gupta *et al.*:[33] the conditions $D_2(u_n)$ and $D'(u_n)$ hold for 1D Lorenz maps and for a large class of

observables, where $D_2(u_n)$ is a weaker form of Leadbetter's $D(u_n)$ condition. Freitas and Freitas[34] showed that $D(u_n)$ can be replaced by $D_2(u_n)$, which is more suitable for stochastic processes arising from chaotic deterministic systems since it follows from sufficiently fast (e.g., exponential) decay of correlations. We also show that ξ_α may depend on the parameter α in two different ways according to where the observable ϕ is maximized:

(I) ϕ is maximized at the upper end point $\phi_\alpha^+ = \phi(x^+)$ of the distribution function F of $X_1 = \phi$.
(II) ϕ has a maximum at $x_0 \in (-1,1)\backslash\{0\}$ with $\theta_\alpha(x_0) \neq 0$.

In loose terms, the maximum is "on the peel" of the attractor for (I) and is "within the attractor" for (II). For simplicity, specific examples of observables are used for each of the above two cases. Our arguments easily generalize to other functional forms.

To examine case (I) we work with the observable $\phi(x) = x$. Using invariance of μ_α and the Lebesgue differentiation theorem, one can show that there exists a function h, with $h(0)$ strictly bounded away from zero and from infinity (that is, $h(0) \in [a,b]$ for some $0 < a < b < +\infty$) and $h(1/u)$ slowly varying as $u \to \infty$, such that for $t, u > 0$,

$$\begin{aligned}\mu_\alpha[\phi(x) > \phi_\alpha^+ - 1/ut] &= \mu_\alpha[x > 1 - 1/ut] \\ &= \mu_\alpha[T_\alpha^{-1}[x^+ - 1/tu, x^+]] \\ &= \theta_\alpha(0) h(1/tu)(tu)^{-1/\alpha} \\ &\quad + o(h(1/tu)(tu)^{-1/\alpha}) \end{aligned} \tag{7}$$

as $u \to \infty$. Defining $g(u) = 1 - F(\phi_\alpha^+ - 1/u)$, we have

$$\frac{g(ut)}{g(u)} = \frac{\mu_\alpha[\phi(x) > \phi_\alpha^+ - 1/ut]}{\mu_\alpha[\phi(x) > \phi_\alpha^+ - 1/u]} = \frac{\theta_\alpha(0) + o(1)}{\theta_\alpha(0) + o(1)} t^{-1/\alpha} \tag{8}$$

as $u \to \infty$, provided $\theta_\alpha(0) \neq 0$ (which numerical experiments indicate it does indeed hold). This shows that F is regularly varying with index $-1/\alpha$. Combining this with Ref. 33, Theorem 3.2, we find that F is in the domain of attraction of a GEV with tail index $\xi = -\alpha$.[25]

To examine case (II) we consider the example $\phi(x) = C - D|x - x_0|^\delta$ with $D, \delta > 0$. Using an argument as above we obtain

$$\frac{g(ut)}{g(u)} = \frac{\theta_\alpha(x_0) + o(1)}{\theta_\alpha(x_0) + o(1)} t^{-1/\delta} \tag{9}$$

as $u \to \infty$. The tail index $\xi = -\delta$ is thus constant in α: it only depends on the "curvature" of the observable ϕ near the extremal point x_0. Both Eqs. (8) and (9) can be generalized to observables with a strict absolute maximum at x_0 [where $x_0 = x^+$ for Eq. (7)] such that $\phi(x) = C - D|x - x_0|^\delta + o(|x - x_0|^\delta)$ as $x \to x_0$ with $D, \delta > 0$.

In summary, if the maxima in the 1D Lorenz maps are GEV distributed, then the tail index varies smoothly with the system parameter α for large classes of observables. The regularity of the density $\theta_\alpha(x)$ plays a significant role in case (I): here $\theta_\alpha(0) = 0$, and the value of $\xi(\alpha)$ continuously depends on α through the density's Hölder constant at $x = 0$. For case (II), we have shown that the tail index depends on the form of the observable (and not on α) for values of x_0 with $\theta_\alpha(x_0) \neq 0$. For Lorenz maps, the density $\theta_\alpha(x)$ is nonzero for μ-almost all values of x_0. Hence for generic maxima within the attractor the value of ξ would appear constant as α is varied.

This analysis clarifies that robust extremes are not equivalent to or implied by any of the concepts of robust chaos mentioned in Sec. I. Indeed, neither the "geometric" definition of robust strange attractors by Ref. 13 nor the "empirical" definition by Banerjee *et al.*[6] entails continuous dependence on the parameter for the SRB measure at the upper end point [see condition (7)]. A particular form of statistical stability holds for the geometrical models of the Lorenz attractor: the continuity with respect to the parameter of the SRB measure in the weak-* topology.[35] However, not even this property is sufficient for robustness of extremes. The sensitivity of the GEV parameters is dependent on the sensitivity of the SRB measure density in the vicinity of the set where the observable is maximized: this is a local property of the system [compare again with Eq. (7)]. Moreover, for weakly dependent stochastic processes (such as those associated with chaotic deterministic dynamical systems) Leadbetter's conditions $D'(u_n)$ and $D(u_n)$ need to be verified and in general this is not straightforward. We return to this point at the end of Sec. VI.

V. THE LORENZ FLOW, REVISITED

What do the results of Sec. IV imply for the flow of the Lorenz63 model (6)? The Lorenz63 flow has an attractor $\Lambda_\rho \subset \mathbb{R}^3$ which supports a unique SRB measure μ_ρ for $\rho \approx 28$, see Ref. 36. A standard approach for the study of the Lorenz63 flow is to consider 1D Lorenz maps, which arise as simplified geometric models of the full 3D flow.[1,27,37] Specifically, a 1D Lorenz map T_α is obtained for the flow of Eq. (6) via an identification map S in a Poincaré section. The map S is formed by identifying points within each stable manifold of the strong stable foliation, see Ref. 11. It can be shown that the parameter α is given by $|\lambda_s|/\lambda_u$, where λ_u is the unstable and λ_s the weakly stable eigenvalue of the Lorenz63 flow at the origin.[27,38] Therefore, the results of Sec. IV can be applied to the flow of Eq. (6). Again, we distin-

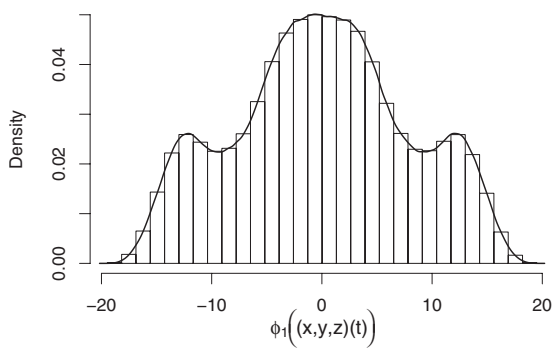

FIG. 2. Histogram and smoothed density from an orbit of Lorenz63 observed through $\phi_1(x,y,z) = x$ for $\rho = 28$, and computed as in Sec. III with $n = 6$, starting from point $(x,y,z) = (0.1, 0.2, 0.3)$ with a transient of 10 000 time units.

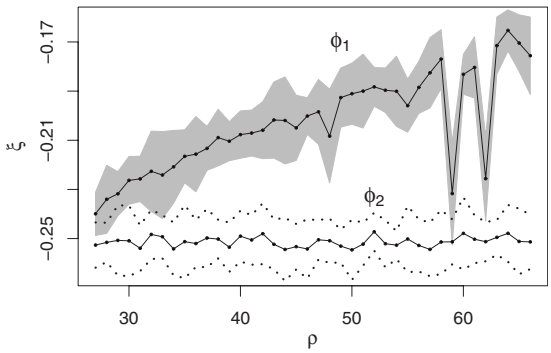

FIG. 3. Maximum likelihood estimates of the tail index ξ for two observables for a time series of length 10^n with $n=8$ (corresponding to 10^5 block maxima, see Sec. III for details). For each $\rho_j \in$ (Ref. 25), the uncertainty bands are computed as twice the standard deviation of the ten distinct estimates obtained with ten "chunks" of 10^4 maxima each.

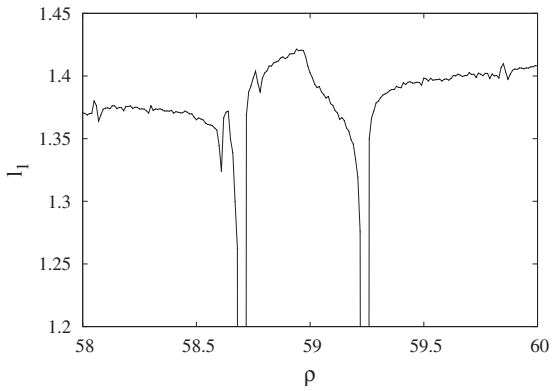

FIG. 5. Maximal Lyapunov exponent ℓ_1 of the attractor of Eq. (6) as a function of ρ.

guish two cases according to whether the observable is maximized on the peel of or "within" the attractor Λ_ρ. We consider two examples:

$$\phi_1(x,y,z) = x \quad \text{and} \quad \phi_2(x,y,z) = 1 - (x-5)^{1/4}. \tag{10}$$

Figure 2 shows a histogram corresponding to the projected measure $\mu_\rho \circ \phi_1^{-1}$. The tail index depends smoothly on ρ through the ratio $|\lambda_s|/\lambda_u$ [obtained in Eq. (8) for T_α with $\phi(x)=x$] and the identification map S (see above). The robustness of the tail index is preserved, since S is smooth in ρ. This is confirmed by the numerical results in Sec. III. Note that the tail index is not constant for the observable ϕ_1: this is visualized in Fig. 3 by scanning a wider ρ-interval than in Fig. 1. Also, the tail index for ϕ_1 is *nonrobust* for large ρ. The sharp drop near $\rho=59$ in Fig. 3 corresponds to a sudden change in the empirical distribution function of the block maxima: the normalized quantiles of the block maxima for $\rho=58$ look much more similar to those for $\rho=60$ than for $\rho=59$ (Fig. 4). The explanation is that folds appear in the Poincaré map as ρ increases beyond 32. The folds correspond to critical points in the 1D Lorenz map T_α,[37] whereby hyperbolicity is destroyed and Eq. (8) no longer holds in this parameter regime. Appearance of *windows of periodicity*[6] and sudden changes in the attractor are typical symptoms of hyperbolicity loss. Two windows of periodicity are identified in Fig. 5 by a vanishing maximal Lyapunov exponent ℓ_1.[39] Attracting periodic orbits with different topological structures occurs within each of these parameter intervals, see Fig. 6. The value $\rho=59$ falls in the transition region between the two windows of periodicity, where ℓ_1 has a peak which is larger than for $\rho=58,60$. Closeness to the periodic windows induces intermittency in the strange attractor for $\rho=59$. This is visualized in Fig. 7 through a smoothed projection of the invariant measure of Λ_ρ on the (x,z)-plane. Intermittency induces a rather different spatial distribution of the "orbit visits" for $\rho=59$ than for $\rho=58,60$; the former is characterized by a steeper, stepwise decrease in the invariant measure near the "edge" of the attractor (compare with the left panel of Fig. 4).

The observable ϕ_2 in Eq. (10) corresponds to the case within the attractor of Sec. IV: indeed, ϕ_2 is maximized on the plane $\{x=5\} \subset \mathbb{R}^3$, which intersects Λ_ρ for all $\rho \in [27,66]$. Constancy of the tail index ξ is readily verified for this observable. An average estimate of $\xi=-0.251$ with standard deviation of 0.002 is obtained over $\rho \in [27,66]$ (Fig. 3), in good agreement with Eq. (9). Note that constancy of ξ holds for the observable ϕ_2 also in the parameter region where the tail index for the observable ϕ_1 is nonrobust due to lack of hyperbolicity.

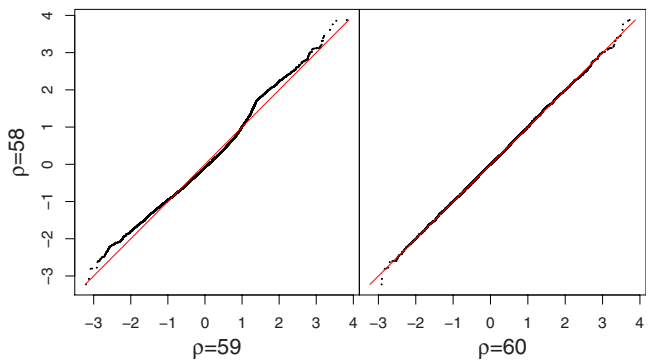

FIG. 4. (Color online) Quantile-quantile plots of 10 000 block maxima of Lorenz63 for $\rho=58$ vs $\rho=59$ (left) and $\rho=60$ (right). The block maxima have been standardized by subtracting the mean and dividing by standard deviation.

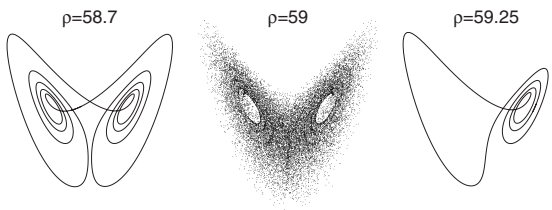

FIG. 6. Projection on (x,z) of orbits of Eq. (6) in the window $(x,z) \in [-30,30] \times [10,100]$ for $\rho=58.7,59,59.25$.

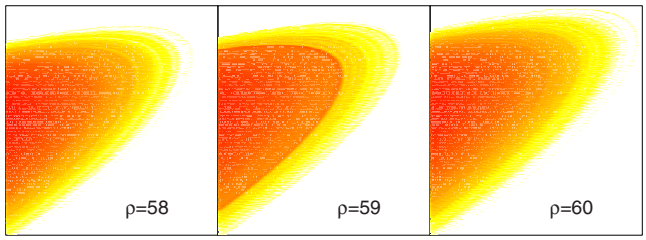

FIG. 7. (Color online) Projection on (x,z) of the invariant density of Eq. (6) obtained by kernel-smoothing 10^6 points along an orbit within the window $(x,z)\in[26,32]\times[60,105]$ for $\rho=58,59,60$ (same scale is used for the densities).

VI. HOW GENERAL ARE ROBUST EXTREMES?

Do the results of Sec. IV hold for other observables and other systems? As shown in Secs. IV and V, the nature of the observable ϕ plays a crucial role in robustness of extremes. If ϕ is maximized within the attractor, the argument leading to Eq. (9) seems to be valid under weak conditions and is insensitive to loss of hyperbolicity: the resulting constancy of ξ is a rather general property. For constancy to hold it is sufficient that the unprojected invariant measure μ_ρ provides good sampling near the set where ϕ is maximized. Caution is required if this set is a point x_0: in general μ_ρ is supported on a fractal set of zero Lebesgue measure in phase space. Even if x_0 belongs to the attractor for some parameter value, typically parameter variations will bring x_0 "outside of the attractor." This problem is avoided for the observable ϕ_2 in Fig. 3 because the attractor Λ_ρ always intersects the extremal set.

To have robustness of extremes in the case on the peel (i.e., when ϕ is not maximized within Λ_ρ), the projected measure must be robust at its upper end point x^+. This condition is specified in Eq. (7) for the 1D Lorenz map. Typically, physical observables such as energy, vorticity, or wind speeds are unbounded in phase space (see Sec. VII and also Ref. 24); therefore, a condition like Eq. (7) is required. However, robustness at the upper end point cannot be expected in general, as shown in Fig. 3. Windows of periodicity are to be expected in the parameter sets of Hénon-like and nonuniformly hyperbolic systems.[1] The limit distribution H (see Sec. IV) is singular for periodic attractors; therefore extreme statistics vary discontinuously with respect to the control parameter.

It is reasonable to expect robust extremes in systems with an axiom-A attractor: indeed, such systems possess a SRB measure which is differentiable in the parameter.[40,41] Unfortunately, for many physical systems encountered in applications it is exceedingly difficult to prove the existence of a SRB measure, let alone its differentiability. One can adopt a pragmatic point of view and simply assume that the system under consideration has a SRB measure which depends smoothly on control parameters, provided that numerical explorations suggest that this is the case: this will be our approach in Sec. VII B. This assumption is compatible with the *chaotic hypothesis*,[42,43] stating that a many particle system in a stationary state can be regarded, for the purpose of computing macroscopic properties, as a smooth dynamical system with a transitive axiom-A global attractor (a version exists for fluid dynamical systems[44]). In line with this, we expect that various types of system may display robust extremes at the experimental, *observable* level: robust chaotic systems,[6,10] systems with a robust attractor,[13] many-particle and fluid dynamical systems,[42,44] high-dimensional systems,[45] and geophysical flows.[24] This may also explain the robustness which is observed in Fig. 3 for ρ between 32 and 58, where the attractor of the Lorenz63 system is not robust (see Sec. V).

From this discussion it seems evident that the tail index depends on the precise form of the SRB measure and local behavior of the observable in the vicinity of the maximum. Although this can be conjectured to be true in many applications the rigorous verification depends on an analysis of the recurrence statistics of typical orbits. A systematic approach is not yet available but recent studies have involved checking Leadbetter's $D(u_n)$ conditions[26] (or weaker versions) directly. Leadbetter's conditions generally fail if the maximum is on a strongly recurrent orbit (e.g., a periodic orbit). In these situations informed conjectures on the tail index cannot be based solely on the local form of the observable and the SRB measure near the maximum.

The verification of condition $D'(u_n)$ depends on both the short and long time range recurrence statistics. In general $D'(u_n)$ is harder to check in specific applications. For logistic maps, $D'(u_n)$ has been shown to hold for observables with a unique maximum along the critical orbit,[34] while Refs. 20 and 21 show that $D'(u_n)$ holds in a broad class of nonuniformly expanding dynamical systems for observables taking maxima at μ-typical values. These results provide a means to investigate the robustness of extremes.

VII. APPLICATIONS

A. Enhancing statistical estimation

We now show how robustness of extremes can be used to improve the precision of estimates of GEV parameters. Consider again the setup of Sec. III. Robust extremes justify the adoption of functional forms (at least locally) to approximate the smooth dependence on ρ of the GEV parameters. Consider, for example,

$$\mu(\rho)=\mu_0+\mu_1\rho,\quad \sigma(\rho)=\sigma_0+\sigma_1\rho,\quad \xi(\rho)=\xi_0. \tag{11}$$

The sequences of maxima $\{z_t^j, t=1,\ldots,m\}$ obtained for $\rho=\rho_j$ $(j=0,1,\ldots,20)$ can then be pooled to estimate these functions. In particular, the parameters $(\mu_0,\mu_1,\sigma_0,\sigma_1,\xi_0)$ can be estimated by maximizing the likelihood

$$\prod_{j,t}\frac{\partial G}{\partial x}(z_t^j;\mu(\rho_j),\sigma(\rho_j),\xi(\rho_j)). \tag{12}$$

"Pooled" estimates obtained from the "short" sequences of $m=100$ maxima [dashed lines in Figs. 1(d)–1(f)] are much closer to the "true values" (solid lines) than the fits carried out for each value of ρ independently. The global structure expressed by Eq. (11) complements the local information contained in the individual sequences of maxima, leading to reduced uncertainty in the parameter estimates: compare the width of the gray bands in the left and right columns of Fig. 1.

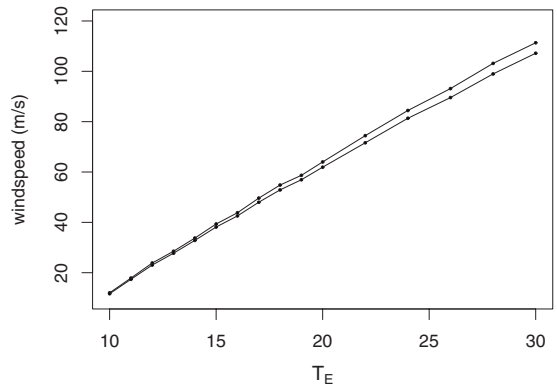

FIG. 8. Confidence intervals at the 95% level for the 100-yr return levels of the wind speed measured at the center of the domain in the lower layer of the quasigeostrophic model. The different values of the control parameter T_E used to generate the time series are given on the horizontal axis.

The above approach can be applied to probabilistic predictions of climate change based on sets of ensembles of simulations with climate models where the individual runs are obtained by perturbing control parameters (the so-called *perturbed physics* approach[46]). Pooling extremes from the different runs as above can reduce uncertainty in the estimated return periods of extreme hazardous events. Of course, robustness of extremes must be assumed or assessed: current research is aiming to develop suitable statistical tests and procedures to be used in such applications.

B. Predicting extremes in nonstationary systems

We show here how robustness can be used in interpreting and predicting nonstationary extremes. We consider the simplified two-layer quasigeostrophic model of Felici *et al.*[29] First we examine the stationary setting by generating several simulations of length 1000 model yr, each with a different value for the baroclinic forcing parameter T_E (note that these simulations do not attempt to reproduce the Earth climate in a specific period of time). Time series of wind speeds are computed for a point at the center of symmetry of the model domain in the lower layer. The GEV distribution (4) is fitted independently for each value of T_E through the block maximum method with annual blocks. Wind speed return levels are obtained as quantiles of the fitted GEV distribution[29] independently for each T_E. Figure 8 displays 100-yr return levels as functions of the corresponding value of T_E. Note that we have chosen a return period which is shorter than the length of the time series in order to limit the sampling variability. Indeed, convergence to the "truth" may require prohibitively long data sets, such as the 10^5 maxima for Fig. 1. The 100-yr return levels have a rather smooth dependence on the control parameter T_E. Figure 8 suggests that a property like Eq. (7) holds for the quasigeostrophic model, where the observable ϕ is the wind speed on the selected point of the domain (same for the total energy[24]). This property is now exploited to improve prediction in the nonstationary case.

Following Felici *et al.*,[24] the following linear time trend is imposed on T_E with speed $\Delta T_E = 2/100$ yr:

TABLE I. Empirical quantiles for the fitted GAMLSS model (see text).

Quantiles (%)	Training	Predicted
0.4	0.0	0.0
2	2.5	3.0
10	8.5	8.9
25	27.0	25.7
50	48.5	50.5
75	75.0	75.2
90	89.0	83.2
98	97.5	95.0
99.6	100.0	98.0

$$T_E(t) = (T_E^0 - 1) + t\Delta T_E, \quad t \in [0, t_0] \tag{13}$$

with $T_E^0 = 10$ and $t_0 = 2100$ yr. Felici *et al.*[24] show that the extremes of the total energy vary smoothly in time for this nonstationary setup. In loose words, nonstationary extremes remain close (locally in time) to those of the stationary system obtained by "freezing" $T_E(t)$ to a constant value, thereby robustness of extremes in T_E translates to smooth variation in extremes in time. An *adiabatic ansatz* is required for this: the trend speed ΔT_E must be sufficiently small with respect to the time which is necessary to sample the upper tail of the energy distribution, given by a projected measure as in Eq. (7).

Given robustness as in Fig. 8 and assuming validity of the adiabatic ansatz, nonstationary statistical models can be used to predict the future variation in extremes. A time-varying GEV distribution (4) is fitted using the generalized additive models for location, scale, and shape (GAMLSS) framework:[47,48] the parameter ξ is kept constant, whereas μ and σ are functions of time fitted by cubic splines with identity link functions.

A single simulation is generated with the quasigeostrophic model, choosing T_E^0 and t_0 in such a way that the range swept by $T_E(t)$ covers the interval[13,19] on the horizontal axis of Fig. 8. A time series of wind speeds is then computed for a point at the center of symmetry of the domain in the lower layer. Yearly maxima are extracted from this nonstationary time series. The sequence of yearly maxima is split into a training set (200–400 years, the initial 66% of the time series) and a prediction-test set (401–500 years, the final 33%). A nonstationary GEV-GAMLSS is fitted to the training set: this yields a time-dependent GEV distribution which is used to compute the quantiles in the left column of Table I. Prediction is performed by a generalization of the method described by Hastie.[49] These time-dependent quantiles are plotted as curves in Fig. 9: the fraction of the training and test data points lying below these curves yields empirical quantiles (center and right column in Table I), whose values compare well with the theoretical quantiles (left column). In summary, Table I shows both goodness of fit (center column) and predictive power (right column) of the GEV-GAMLSS. The fraction of predicted points between the 5% and 95% time-dependent quantiles is 0.85: closeness to the theoretical value of 0.9 gives a further measure of predictive power.

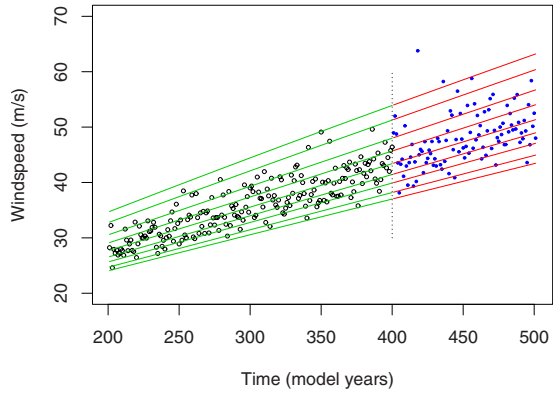

FIG. 9. (Color online) Wind speed maxima over blocks of 1 yr length for the nonstationary simulation of the quasigeostrophic model. The solid lines are the time-dependent quantiles in the left column of Table I estimated for a GEV-GAMLSS model. The vertical dashed line separates the training and prediction-test data sets (see text).

Remarks:

(1) The time-dependent quantiles do not define a full predictive distribution, since uncertainty in the estimation is not taken into account when GAMLSS are used with additive smoothers such as the cubic splines (see the GAMLSS manual, Sec. 3.4).
(2) Construction of a predictive distribution which incorporates uncertainty is easier if parametric models (say, through polynomials) are used for modeling μ and σ as functions of time. However, this yields functional shapes which are in principle more rigid and may not fully capture the nonlinear dependence of μ and σ on ρ (compare with Fig. 8), see Sec. VI. We believe that this is the reason for which cubic smoothing splines are found to work better in our example.
(3) In GEV modeling, a log link is often used for σ to preserve its positivity Coles.[18] However this works less well for extrapolation because the predicted σ values tend to grow excessively.

VIII. CONCLUSIONS

This paper introduces the notion of robust extremes in deterministic chaotic systems, presents initial theoretical results, and outlines associated inferential techniques. There is a considerable scope to develop the mathematical theory and thereby to obtain results for higher-dimensional systems such as uniformly or partially hyperbolic systems, see, e.g., Ref. 50. There is also a scope to develop the inferential techniques, for example, to develop statistical tests for robustness of extremes and methods for predicting extremes in both stationary and nonstationary systems.

ACKNOWLEDGMENTS

R.V. acknowledges the kind support by the Willis Research Network (www.willisresearchnetwork.com).

[1]J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer-Verlag, Berlin, 1983).
[2]H. W. Broer, Bull., New Ser., Am. Math. Soc. **41**, 507 (2004).
[3]L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics. Vol. 5: Statistical Physics*, translated from the Russian by J. B. Sykes and M. J. Kearsley, second revised and enlarged edition (Pergamon, New York, 1968).
[4]C. Grebogi, E. Ott, and J. A. Yorke, Phys. Rev. Lett. **48**, 1507 (1982).
[5]J. Palis and F. Takens, *Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations* (Cambridge University Press, Cambridge, 1993).
[6]S. Banerjee, J. A. Yorke, and C. Grebogi, Phys. Rev. Lett. **80**, 3049 (1998).
[7]A. Potapov and M. K. Ali, Phys. Lett. A **277**, 310 (2000).
[8]M. Andrecut and M. K. Ali, Phys. Rev. E **64**, 025203(R) (2001).
[9]P. Kowalczyk, Nonlinearity **18**, 485 (2005).
[10]Z. Elhadj and J. Sprott, Fron. Phys. China **3**, 195 (2008).
[11]W. Tucker, Acad. Sci., Paris, C. R. Sér. I Math. **328**, 1197 (1999).
[12]E. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[13]C. A. Morales, M. J. Pacifico, and E. R. Pujals, Proc. Am. Math. Soc. **127**, 3393 (1999).
[14]J. F. Alves, M. Carvalho, and J. M. Freitas, "Statistical stability for Hénon maps of the Benedicks-Carleson type," Ann. Inst. Henri Poincare, Anal. Non Lineaire (to be published), e-print arXiv:math/0610602.
[15]J. M. Freitas and M. Todd, Nonlinearity **22**, 259 (2009).
[16]E. Castillo, *Extreme Value Theory in Engineering*, Statistical Modeling and Decision Science (Academic, New York, 1988).
[17]P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Applications of Mathematics (New York), Vol. 33 (Springer-Verlag, Berlin, 1997).
[18]S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics (Springer, New York, 2001).
[19]J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, *Statistics of Extremes: Theory and Applications* (Wiley, New York, 2004).
[20]P. Collet, Ergod. Theory Dyn. Syst. **21**, 401 (2001).
[21]M. Holland, M. Nicol, and A. Török, "Extreme value distributions for non-uniformly hyperbolic dynamical systems" (unpublished).
[22]A. C. M. Freitas and J. M. Freitas, Ergod. Theory Dyn. Syst. **28**, 1117 (2008).
[23]A. C. M. Freitas, J. M. Freitas, and M. Todd, "Hitting time statistics and extreme value theory," Probab. Theory Relat. Fields (to be published), e-print arXiv:0804.2887v2.
[24]M. Felici, V. Lucarini, A. Speranza, and R. Vitolo, J. Atmos. Sci. **64**, 2159 (2007).
[25]N. H. Bingham, J. Comput. Appl. Math. **200**, 357 (2007).
[26]M. R. Leadbetter, Z. Wahrscheinlichkeitstheor. Verwandte Geb. **65**, 291 (1983).
[27]C. Sparrow, *The Lorenz equations: bifurcations, chaos, and strange attractors*, Applied Mathematical Sciences, Vol. 41 (Springer-Verlag, New York, 1982).
[28]The variable-order adaptive-stepsize Taylor method of Jorba and Zou (Ref. 51) is used for integration of Eq. (6) with local absolute and relative truncation errors fixed at 10^{-8}.
[29]M. Felici, V. Lucarini, A. Speranza, and R. Vitolo, J. Atmos. Sci. **64**, 2137 (2007).
[30]V. Lucarini, A. Speranza, and R. Vitolo, Physica D **234**, 105 (2007).
[31]E. Kriegler, J. W. Hall, H. Held, R. Dawson, and H. J. Schellnhuber, Proc. Natl. Acad. Sci. U.S.A. **106**, 5041 (2009).
[32]L.-S. Young, J. Stat. Phys. **108**, 733 (2002).
[33]C. Gupta, M. Holland, and M. Nicol, "Extreme value theory for Lorenz maps and flows, Henon-like diffeomorphisms and a class of hyperbolic systems with singularities" (unpublished).
[34]A. C. M. Freitas and J. M. Freitas, Stat. Probab. Lett. **78**, 1088 (2008).
[35]L. Mendoza, *Mathematical Notes, No. 100 (Spanish)* (Univ. de Los Andes, Mérida, 1989), pp. 91–101.
[36]V. Araujo, M. J. Pacifico, E. R. Pujals, and M. Viana, Trans. Am. Math. Soc. **361**, 2431 (2009).
[37]S. Luzzatto and W. Tucker, Publ. Math., Inst. Hautes Etud. Sci. **89**, 179 (1999).
[38]S. Luzzatto, I. Melbourne, and F. Paccaut, Commun. Math. Phys. **260**, 393 (2005).
[39]The algorithm described in Broer *et al.* (Ref. 52) has been used.
[40]D. Ruelle, Commun. Math. Phys. **187**, 227 (1997).
[41]D. Ruelle, Ergod. Theory Dyn. Syst. **28**, 613 (2008).
[42]G. Gallavotti and E. G. D. Cohen, J. Stat. Phys. **80**, 931 (1995).
[43]F. Bonetto, G. Gallavotti, A. Giuliani, and F. Zamponi, J. Stat. Phys. **123**, 39 (2006).
[44]G. Gallavotti, *Foundations of Fluid Dynamics* (Springer-Verlag, Berlin, 2002).

[45]D. J. Albers, J. C. Sprott, and J. P. Crutchfield, Phys. Rev. E **74**, 057201 (2006).
[46]J. M. Murphy, B. B. B. Booth, M. Collins, G. R. Harris, D. M. H. Sexton, and M. J. Webb, Philos. Trans. R. Soc. London, Ser. A **365**, 1993 (2007).
[47]R. A. Rigby and D. M. Stasinopoulos, J. R. Stat. Soc., Ser. C, Appl. Stat. **54**, 507 (2005).
[48]D. Stasinopoulos and R. A. Rigby, J. Stat. Software **23**, 1 (2007).
[49]T. J. Hastie, in *Statistical Models in S*, edited by J. M. Chambers and T. J. Hastie (Brooks-Cole, Belmont, 1991), pp. 309–376.
[50]J. F. Alves, C. Bonatti, and M. Viana, Invent. Math. **140**, 351 (2000).
[51]À. Jorba and M. Zou, Exp. Math. **14**, 99 (2005).
[52]H. Broer, C. Simó, and R. Vitolo, Nonlinearity **15**, 1205 (2002).