Unveiling coherent structures through entropy-complexity analysis

Antoine Chrisment and Marie-Christine Firpo

Laboratoire de Physique des Plasmas, CNRS - Ecole Polytechnique, 91128 Palaiseau cedex, France
(E-mail: `marie-christine.firpo@lpp.polytechnique.fr`)

Abstract. The entropy-complexity plane has been shown in [O.A. Rosso *et al.*, Phys. Rev. Lett. **99**, 154102 (2007)] to be able to discriminate between low-dimensional deterministic chaos and noise, since chaotic maps and stochastic processes occupy different regions of the plane. For time continuous signals, a Poincaré section approach is proposed that provides a natural way to resolve the arbitrariness of the sampling time and return to the former discrete time case. This approach is illustrated on the repulsive Hamiltonian Mean Field (HMF) model that is a prototype of long-range systems. The permutation entropy and complexity are shown to capture the low energy out-of-equilibrium bicluster coherent structure.

Keywords: Deterministic chaos, stochasticity, long-range interacting systems, mean–field models, quasi-stationary states.

1 Introduction

In a recent work, Rosso *et al.* introduced a representation space, the entropy-complexity causality plane [1], as a novel tool to evaluate the chaotic and/or stochastic nature of dynamical systems. They considered some well-known chaotic maps, such as the Hénon or Schuster maps, and stochastic processes ranging from fractional Brownian motion with different Hurst parameters to white noise. The Bandt-Pompe permutation entropy [2] and Jensen-Shannon complexity [3] of their iterations were computed. The examination of the ensuing entropy-complexity plane, where the permutation entropy is normalized to its maximal value, revealed a quite appealing feature, namely a geography of chaos and stochasticity and a segregation between both regions. The accessible entropy-complexity region takes the form of a crescent joining the two limits of zero complexity associated to exactly constant or strictly monotonous time series (yielding a zero permutation entropy) in the left corner to the white noise limit maximizing the permutation entropy in the right corner. On this graph, the chaotic maps appear to have usually higher complexities than stochastic processes and to locate mostly in the left half-plane whereas stochastic processes occupy the right half-plane.

The application of this framework to time-continuous signals remains however problematic. Indeed, let us have in mind the time evolution of an observable of a deterministic system. If this time evolution is sampled with a too large sampling period, the resulting time series may eventually appear as a list of uncorrelated data, and therefore as noise. Conversely, a small sampling period will drive the location in the entropy-complexity plane to the left [4]. Therefore, one needs some procedure to overcome as much as possible the arbitrariness of the sampling period in time-continuous systems. Following our recent work [4], we propose to use a Poincaré approach by working on sub-series corresponding to the signal relative maxima (or minima).

So doing, some evidence will be given that the entropy-complexity frame can be used as a new complementary tool to analyze the dynamics and transport properties of N-body systems. We have in mind long-range systems, and in particular mean-field systems, that are well-known to exhibit both special equilibrium properties (e.g. with the possibility of ensemble inequivalence [5]) and relaxation properties such as long-lived out-of-equilibrium states supported by collective waves or other quasi-stationary states [6]. These are ergodicity-breaking features due to low collisionality impeding the relaxation towards Gibbs-Maxwell equilibrium and/or due to possibly insufficient intrinsic stochastic properties. Therefore these systems have a central role to play in the identification of the dynamical requirements for the validity and extensions of statistical mechanics [7].

Because much of the information on the dynamics of N-body globally-coupled systems is usually contained in the time evolution of some low-dimensional subset of collective macroscopic variables, it is meaningful to focus on the characterization of the chaotic or stochastic properties of these few relevant collective variables. A prototype of such systems is the well-known Hamiltonian Mean Field (HMF) model [8] that will be introduced in Section 2. If the all-to-all particle coupling is repulsive, the HMF system was shown to exhibit some puzzling out-of-equilibrium dynamics in the low energy regime with the emergence of long-lived bicluster patterns whereas the equilibrium statistical mechanics predicts an homogeneous phase for all energies [9]. This transition will serve here to probe the entropy-complexity analysis. This will be introduced and discussed in Section 3. In Section 4, a Poincaré section approach, by considering the time series of the relative maxima of the mean-field, is shown to enable to single-out in the entropy-complexity plane the low-energy regime where dynamical anomalies take place. A short evocation of the potential applications and perspectives of this work is given in Section 5.

2   Dynamics of the globally-coupled repulsive HMF model

A paradigmatic conservative globally-coupled system is the so-called Hamiltonian Mean Field (HMF) model where N particles are moving on a circle being

globally coupled by a cosine interaction with trajectories deriving from

$$H(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{N} \frac{p_i^2}{2} + \frac{c}{2N} \sum_{i,j=1}^{N} \left[1 - \cos(q_i - q_j) \right]. \tag{1}$$

The interacting potential is of the same form as in the non-conservative Kuramoto model and both systems bear similarities. For instance, a puzzling $N^{1.7}$ scaling of the lifetimes of homogeneous quasistationary states was reported for the attractive HMF model [10] strangely resonating with the $N^{-1.69}$ scaling of the diffusion coefficient reported in chimera states of the Kuramoto model [11]. An explanation for this strange scaling was proposed in Ref. [12] in the HMF frame on the basis of a stochastic, diffusive, approach.
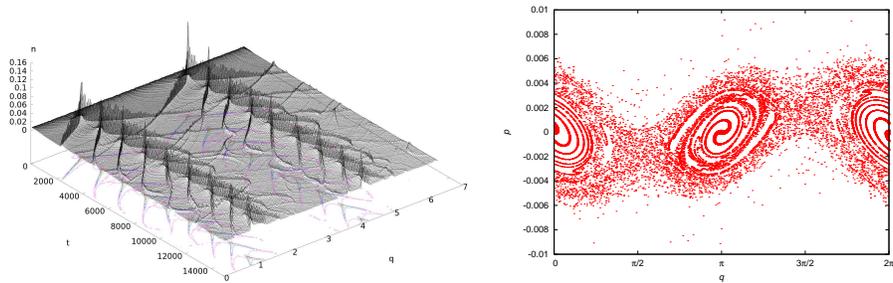
Fig. 1. (left) Early time evolution of the density of the N particles on the $\mathbb{R}/2\pi\mathbb{Z}$ circle and (right) snapshot of the one-particle (q, p) space at time $t = 30000$ for the energy density $\varepsilon = H/N = -0.4999$. In these simulations, $N = 20000$ particles were used.

Introducing the collective variable usually called the magnetization, using an analogy of the HMF potential with the X-Y spin model, defined by

$$\mathbf{M} = (M_x, M_y) = \left(\frac{1}{N} \sum_{i=1}^{N} \cos q_i, \frac{1}{N} \sum_{i=1}^{N} \sin q_i \right), \tag{2}$$

the equation of motion of any particle i may be written

$$\frac{d^2 q_i}{dt^2} = -cM_x(t) \sin q_i + cM_y(t) \cos q_i. \tag{3}$$

For a negative coupling constant c, the equilibrium statistical mechanics of the HMF model for this system predicts a vanishing canonical ensemble average of the modulus M of the mean-field. Moreover, the equivalence of the canonical and microcanonical ensembles has been proved for this system. Therefore, from an equilibrium statistical mechanics point of view, the repulsive HMF model appears as trivial, and consequently, uninteresting. However, the numerical symplectic computations of its dynamics revealed some puzzling out-of-equilibrium features. Putting the constant c equal to -1, the minimal accessible energy

density is $\varepsilon = H/N = -1/2$. For very low energy simulations, when ε approaches $-1/2$, a very robust biclustered state was observed to form on the $\mathbb{R}/2\pi\mathbb{Z}$ circle, the initiation of which being represented on Figure 1.

   This phenomenon was mostly elucidated as a nonlinear collective effect and an effective picture of the quasistationary biclustered states was proposed in Refs. [13,14]. In the present study, our objective is to use time series of the mean-field modulus $M(t) = (M_x^2 + M_y^2)^{1/2}$, a collective observable, in the quasisteady states at different energy densities, to probe the entropy-complexity analysis and extract information on the nature of the dynamics from it. Numerical results are obtained using a fourth-order symplectic integrator [15] using a time-step $\Delta\tau = 10^{-2}$ ensuring notably a very robust conservation of the total energy.

   Finally, let us make an important point. From Poincaré's recurrence theorem, it follows that, in the volume-preserving, and thus Hamiltonian, context, almost any point is recurrent. This a priori prevents the existence of genuine, possibly low-dimensional, chaotic attractors that contrarily may exist in dissipative systems. It is therefore not a priori possible to reconstruct the phase space of a $N \gg 1$-body Hamiltonian system with a low embedding dimension, d. However, the HMF model considered here is a globally-coupled, mean-field system. Discarding the self-consistency, needed to compute the time behavior of $\mathbf{M}(t)$ in Eq. (3), this N-body model would amount to just a one-and-a-half degrees of freedom Hamiltonian, or more exactly, a collection of N uncoupled one-and-a-half degrees of freedom Hamiltonian systems. This leads us to conjecture that it can display some low dimensional dynamic properties that may be captured within a complexity-entropy approach. This is introduced in the following Section.

3   Permutation entropy and construction of complexity

In order to incorporate the intuitive notion of complexity of a physical system, in relation with the amount of order/disorder present in its states, and the hierarchy between them, a statistical measure of *complexity* was proposed in terms of access probabilities [16]. Its construction started from the assessment that two opposite fundamental N-body systems could be viewed as being simple, and be defined as two limits of zero-complexity, namely: i) the perfect crystal and ii) the isolated ideal gas. The perfect crystal is completely ordered and therefore a state is privileged : the hierarchy between accessible states is strong, such that, for this system, the distance to the equiprobable distribution is maximal. Conversely, the isolated ideal gas is completely disordered and all its microstates are equiprobable.

   These limit-zero complexity states led to the introduction of the notion of *disequilibrium*: Whereas entropy measures the degree of disorder, disequilibrium is defined to be the distance between the equiprobable distribution and the distribution of microstates of the studied system, such that complexity is the product of those two measures

$$[\text{Complexity}] = [\text{Disequilibrium}] \times [\text{Entropy}]. \tag{4}$$

A measure of complexity [3] is also expected to satisfy the following requirements: it should be intensive, be able to distinguish among different degrees of periodicity and it should give the most exactly possible an indication of the nature (chaotic or integrable) of the dynamics. From those constraints, a relevant choice made by many authors [1,17] has been to choose in the complexity definition's (4) the Jensen-Shannon disequilibrium and the Bandt-Pompe's permutation entropy, the definition of those will be recalled in the following.

Let us start to introduce the Bandt-Pompe's permutation entropy. Our working objects are time series from some physical observable. Given a time series, one examines a sequence, chosen randomly, of d successive points.

Here d denotes the so-called embedding dimension, chosen to be rather small [1]. In particular, Bandt and Pompe suggest for practical purposes working with $3 \leq d \leq 7$ [2], and in the present work we shall use $d = 5$ (while e.g. Rosso *et al.* use $d = 6$ in [1]).

Then one asks: What is the permutation of those d points which sorts them by increasing order? The occurrence probability μ_i ($i \in \{1, ..., d!\}$) of a permutation π_i is defined as being the probability that the answer to the later question is π_i. Then the permutation distribution of the time series is defined through the vector of $[0; 1]^{d!}$

$$\mathbf{P} = (\mu_1, ..., \mu_{d!}) \text{ with } \sum_{i=1}^{d!} \mu_i = 1. \tag{5}$$

Therefore the uniform permutation probability distribution $\mathbf{P}_e$ corresponds to $\mu_i = \frac{1}{d!}$, $i \in \{1, ..., d!\}$. Let Ω_d be the set of permutation probability distributions of size d.

Definition 1 (Bandt-Pompe's permutation entropy). The Bandt-Pompe's permutation entropy, S_{BP}, [2] of order d is the Shannon entropy, S_S, of a d-order permutation probability distribution

$$\forall \mathbf{P} \in \Omega_d, \quad S_{BP}(\mathbf{P}) = S_S(\mathbf{P}) = -\sum_{i=1}^{d!} \mu_i \ \log \ \mu_i \tag{6}$$

and therefore the Bandt-Pompe's (permutation) entropy is the restriction of the Shannon's one on Ω_d.

Consequently, the permutation probability distribution which maximizes the Bandt-Pompe entropy is the uniform one, $\mathbf{P}_e$, corresponding to a time series of the highest degree of randomness

$$\max_{\mathbf{P} \in \Omega_d} S_{BP}(\mathbf{P}) = S_{BP}(\mathbf{P}_e) = \log(d!). \tag{7}$$

The permutations distribution which minimize the Bandt-Pompe entropy are those for which $(d! - 1)$ occurrence probabilities are equal to 0 (so one of them is equal to 1), corresponding to monotonic time series, so that

$$\min_{\mathbf{P} \in \Omega_d} S_{BP}(\mathbf{P}) = 0 - 1 \ \log \ 1 = 0. \tag{8}$$

The disequilibrium, noted Q, is proportional to the distance between the probability distribution $\mathbf{P}$ and the uniform one : $Q(\mathbf{P}) = \eta D(\mathbf{P}, \mathbf{P}_e)$, where $\eta \in \mathbb{R}_+^*$ is such that $0 \leq Q \leq 1$. The choice of D is crucial and in this study, according to constraints evoked before, the measure of the disequilibrium has consisted of using a Jensen-Shannon divergence defined by [3]

$$\forall \mathbf{P}_1, \mathbf{P}_2 \in \Omega_d ~,~ J(\mathbf{P}_1, \mathbf{P}_2) = S_S\left(\frac{1}{2}\mathbf{P}_1 + \frac{1}{2}\mathbf{P}_2\right) - \frac{1}{2}S_S(\mathbf{P}_1) - \frac{1}{2}S_S(\mathbf{P}_2) \quad (9)$$

such that the distance is $D(\mathbf{P}, \mathbf{P}_e) = J(\mathbf{P}, \mathbf{P}_e)$. It remains then to estimate η. As Q measures a deviation from the uniform distribution, Q is maximal for distributions for which $(d! - 1)$ occurrence probabilities are equal to 0, so that one of them is equal to 1. Designing by $\mathbf{P}_m = (0, ...0, 1, 0, ..., 0)$ one of such a distribution (the 1 is at the mth position), from

$$\frac{1}{2}\mathbf{P}_m + \frac{1}{2}\mathbf{P}_e = \left(\frac{1}{2d!}, ...\frac{1}{2d!}, \frac{1 + \frac{1}{d!}}{2}, \frac{1}{2d!}, ..., \frac{1}{2d!}\right), \quad (10)$$

one deduces $S_S\left(\frac{1}{2}\mathbf{P} + \frac{1}{2}\mathbf{P}_e\right)$, and, because $S_S(\mathbf{P}_m) = 0$ and $S_S(\mathbf{P}_e) = \log(d!)$, one obtains

$$1 = \max_{\mathbf{P} \in \Omega_d} Q(\mathbf{P}) = Q(\mathbf{P}_m)$$

$$= -\frac{\eta}{2}\left(\frac{d! + 1}{d!}~\log(d! + 1) - 2\log(2d!) + \log(d!)\right). \quad (11)$$

Consequently, the Jensen-Shannon's statistical disequilibrium of any probability distribution $\mathbf{P} \in \Omega_d$ reads

$$Q_{JS}(\mathbf{P}) = \frac{S_S\left(\frac{1}{2}\mathbf{P} + \frac{1}{2}\mathbf{P}_e\right) - \frac{1}{2}S_S(\mathbf{P}) - \frac{1}{2}S_S(\mathbf{P}_e)}{-\frac{1}{2}\left(\frac{d!+1}{d!}~\log(d! + 1) - 2\log(2d!) + \log(d!)\right)}, \quad (12)$$

where the JS subscript has been added. From this follows, the definition of the complexity used in our study.

Definition 2 (Jensen-Shannon's statistical complexity). The Jensen-Shannon's statistical complexity is defined by

$$C_{JS} = Q_{JS} \times s_{BP} \quad \text{where} \quad s_{BP} = \frac{S_{BP}}{\log(d!)}, \quad (13)$$

where Q_{JS} is given in Eq. (12).

4   Permutation entropy and complexity as dynamical markers of coherent structures

The impact on the entropy-complexity location of the time resolution in the time series of the observable was shown and discussed in our recent work [4]. This dependency on the time resolution makes for instance quite problematic

the comparison in the entropy-complexity frame of time series coming from different systems. It is therefore desirable to search for routines diminishing as much as possible this dependency. A possibility is to extract the maxima (or minima) of the time series and to proceed to the entropy-complexity analysis on these sub-time series. This amounts to work on a zero-time derivative Poincaré section of the time signal. This approach is shown below to enable to single out the energy domain where dynamic anomalies take place.

We performed numerical simulations of the repulsive HMF model using waterbag conditions in p and a uniform space distribution so that the initial distribution functions are of the type $f_0(p,q) = (2\Delta p)^{-1}1_{[-\Delta p;\Delta p]}(p)$.

Let us first assume that the time resolution used in the time series of the observable under consideration is sufficiently fine, so that one remains far away from the maximal entropy case mentioned above where the too large sampling period make successive sampled values appear unrelated. Then, instead of using the full time series in the entropy complexity analysis, one proceed to a Poincaré section of the signal by restricting to the series of its relative maxima (or minima). This amounts to doing a Poincaré section of the signal $M(t)$ on the zero time derivative section $dM/dt = 0$ [18]. In the present model, times series of the mean-fields were recorded every time step $\Delta\tau = 10^{-2}$. We extracted from these time series the relative maxima of the mean-fields. Figure 2 represent the corresponding time series of the relative maxima of the modulus of the mean-field, M, for three different energy densities close to the fundamental one.
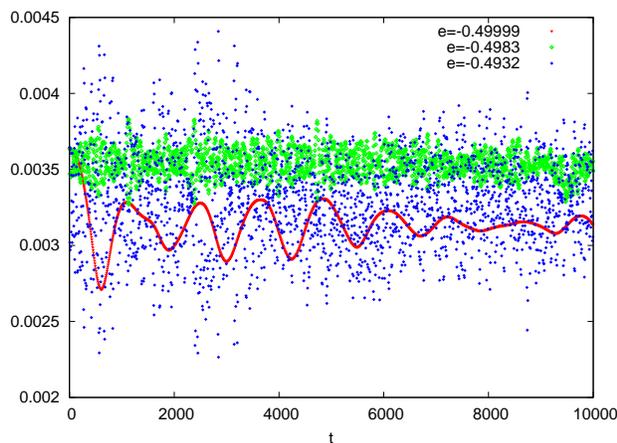

Fig. 2. Each point represents the value of a relative maximum of M (y-axis) at its occurrence time (x-axis). Three different simulations with different energy densities are represented. Waterbag initial conditions were used.

At low energy, the biclustered phase illustrated on Figure 1 may be quantified by the mean-field

$$\mathbf{M_2} = (M_{2x}, M_{2y}) = \left(\frac{1}{N} \sum_{i=1}^{N} \cos 2q_i, \frac{1}{N} \sum_{i=1}^{N} \sin 2q_i \right), \tag{14}$$

the two components of which are the $n = 2$ Fourier coefficients of the spatial distribution, corresponding to the spatial scale π. It is instructive to compute the time averages of the mean-fields M and M_2 as a function of the energy density. This has been done using the initial waterbag distributions and results are shown on Figure 3. Equilibrium statistical mechanics would predict
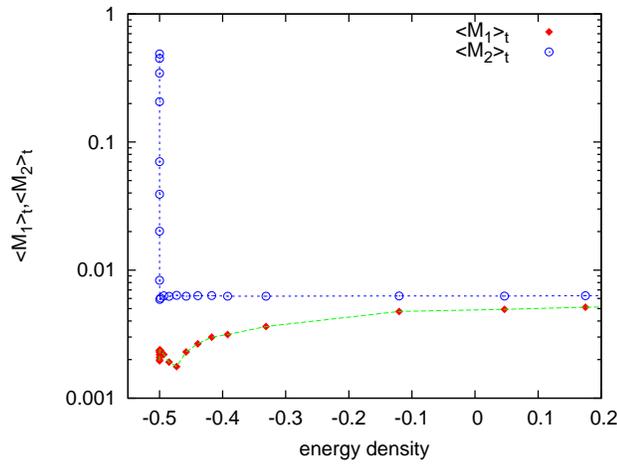
Fig. 3. Time averages as a function of the energy densities of the moduli of the mean-fields M (noted here M_1) and M_2 obtained from waterbag initial conditions. The plot is in lin-log scale.

vanishing mean-fields in the large N limit which is at odds with the low energy values of the time averages of M_2. One may remark that the repulsive HMF model is presumably, and not surprisingly, non ergodic at low energies since changing the initial conditions to the previous sinusoidal conditions in the impulsion, that served to push the collective mode, could broaden the energy domain associated to the biclustered phase. Here the biclustered phase appears as a zero-temperature out-of-equilibrium effect that may have some applications to cold atoms [19]. From the observation of Figure 3, one can also remark that, whereas the time average of the mean-field M_2 rapidly changes from a macroscopic value signaling the biclustered phase about the fundamental state to attain its $O(N^{-1/2})$ large energy value, the time average of the mean-field M presents some anomalous behaviour in a somehow wider energy domain above the fundamental state.

We extracted the time series given by the relative maxima of both the magnetization, M, and of the M_2 mean-fields. The result of the entropy-complexity analysis of these sub-time series is represented on Figure 4.
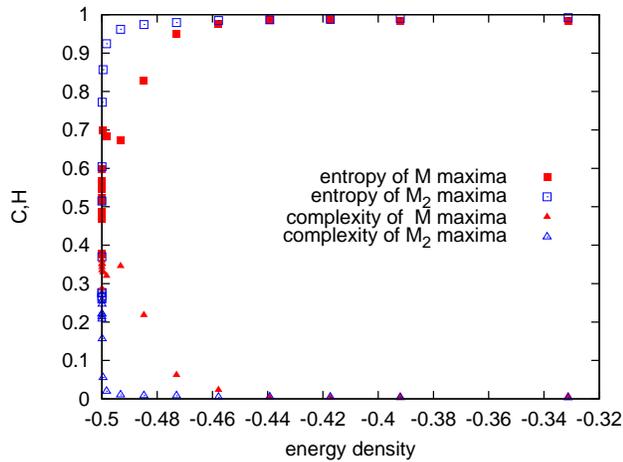
Fig. 4. Normalized Bandt-Pompe entropy and Jensen-Shannon complexity associated to the time series of the relative maxima of the magnetization $M(t)$ and of the relative maxima of $M_2(t)$ as a function of the energy density.

The comparison between the Figures 3 and 4 is quite explicit. The entropy-complexity analysis in the Poincaré section approach shows that the mean-fields, M and M_2, behave as white-noise on the timescale separating two successive maxima except in the energy domain where they display out-of-equilibrium features. This is very clear for the M_2 mean-field for which the time-average drops rapidly to its thermal value when the energy per body, ε, is just above $-1/2$. Correspondingly, there is a sharp variation of the entropy and of the complexity that becomes very low and quickly vanishes. The analysis of the magnetization M shows the same decay of complexity towards zero and growth of the normalized entropy towards one, yet taking place on a wider domain of energy that roughly corresponds to a non-thermal behavior of M presumably due to the presence of some collective mode. These results bring then some evidence that an entropy-complexity analysis is able to unveil out-of-equilibrium dynamical features. Further study will be conducted to assess these results.

5   Conclusion

A Poincaré section approach through the use of the sub-time series corresponding to the relative maxima (or minima) of the mean-field time series provides a frame in which the sampling period is, so to speak, automatically chosen by the system. Some numerical evidence has been given that the entropy-complexity results using this approach are able to single-out the low-energy anomalous dynamics. The entropy-complexity analysis is thus able to separate the low energy domain where the emergence of the collective mode induces some coherence effects from the upper-energy domain where it only capture finite-N noise.

This result offers interesting perspectives for this approach, e.g. as a novel - easy to handle - dynamical indicator to estimate the respective weights of collective modes versus turbulence in domains such as fluid dynamics or plasma physics.

References

1. O.A. Rosso *et al.*, Phys. Rev. Lett. **99**, 154102 (2007).
2. C. Bandt and B. Pompe, Phys. Rev. Lett. **88**, 174102 (2002).
3. P.W. Lamberti, M.T. Martin, A. Plastino, O.A. Rosso, Physica A **334**, 119131 (2004).
4. A. Chrisment and M.-C. Firpo, Physica A **460**, 162-173 (2016).
5. F. Leyvraz and S. Ruffo, J. Phys. A **35** (2002).
6. W. Ettoumi and M.-C. Firpo, Phys. Rev. E **84**, 030103(R) (2011).
7. C. Tsallis, Braz. J. Phys. **29**, 1 35 (1999).
8. M. Antoni and S. Ruffo, Phys. Rev. E **52**, 2361 (1995).
9. T. Dauxois, P. Holdsworth, and S. Ruffo, Eur. Phys. J. B **16**, 659667 (2000).
10. , Y.Y. Yamaguchi *et al.*, Physica A **337**, 36-66 (2004).
11. O.E. Omel'chenko, M. Wolfrum, Y.L. Maistrenko, Phys. Rev. E **81**, 065201 (2010).
12. W. Ettoumi and M.-C. Firpo, Phys. Rev. E **87**, 030102(R) (2013).
13. J. Barré, F. Bouchet, T. Dauxois, and S. Ruffo, Eur. Phys. J. B **29**, 577591 (2002).
14. F. Leyvraz, M.-C. Firpo, S. Ruffo, J. Phys. A **35**, 4413 (2002).
15. H. Yoshida, Phys. Lett. A **150**, 262 268 (1990).
16. R. López-Ruiz, H. Mancini, X. Calbet, Phys. Lett. A **209**, 321 326 (1995).
17. F. Olivares, A. Plastino, O.A. Rosso, Phys. Lett. A **376**, 15771583 (2012).
18. H. Kantz, T. Schreiber, Nonlinear time series analysis, Vol. 7 of Cambridge Nonlinear Science Series, Cambridge University Press, Cambridge (1997).
19. J. Maciel, M.-C. Firpo, M. Amato, Physica A **424**, 34 43 (2015).