Generalized extensive entropies for studying dynamical systems in highly anisotropic phase spaces

Giorgio Sonnino[1,2,*] and György Steinbrecher[3,†]

[1]Université Libre de Bruxelles (ULB), Department of Physics, Bvd du Triomphe, Campus de la Plaine CP 231, 1050 Brussels, Belgium

[2]Royal Military School (RMS), Av. de la Renaissance 30, 1000 Brussels, Belgium

[3]Association EURATOM-MEdC, University of Craiova, A. I. Cuza 13, 200585 Craiova, Romania

(Received 19 November 2013; published 4 June 2014)

Starting from the geometrical interpretation of the Rényi entropy, we introduce further extensive generalizations and study their properties. In particular, we found the probability distribution function obtained by the MaxEnt principle with generalized entropies. We prove that for a large class of dynamical systems subject to random perturbations, including particle transport in random media, these entropies play the role of Liapunov functionals. Some physical examples, which can be treated by the generalized Rényi entropies, are also illustrated.

DOI: 10.1103/PhysRevE.89.062106 PACS number(s): 05.70.Ln, 52.25.Dg, 05.20.Dd, 02.30.Cj

I. INTRODUCTION

Generally, to characterize some unknown events with a statistical model, we choose the one that has maximum entropy (MaxEnt principle), i.e., the one that has maximum uncertainty. The corresponding probability distribution functions (PDFs) are obtained by maximizing entropy, under a set of constraints. In other terms, entropy is regarded as a measure of information, and it is a quantity able to quantify the uncertainty, or the randomness, of a system. Information theory was initially developed by Shannon to quantify the expected value of the information contained in a message, usually in units such as bits [1]. However, the Shannon entropy, used in standard top-down decision trees, does not guarantee the best generalization (see, for example, Ref. [2]). The Shannon entropy has then been successively generalized. Rényi introduced the most general definition of information measures that preserve the additivity for independent events and are compatible with the axioms of probability [3]. Tsallis, on the on the other hand, introduced a nonadditive entropy, such as nonextensive statistical mechanics, generalizing the Boltzmann-Gibbs theory [4]. Rényi's and Tsallis's entropies are algebraically related, and both definitions include the Shannon entropy as a limit case.

The structure of the work is the following. In Sec. II the main necessity to extend the definition of the Rényi entropy to the study of distributions in anisotropic phase spaces is discussed. In Sec. III the Rényi entropy is expressed by a distance in a metric function spaces. We give examples when the Shannon-Boltzmann entropy is infinite while the Rényi entropy is finite, and we give a two-variable PDF with all of the Rényi entropy infinite (Remark 2).

In Sec. IV the natural generalization of the Rényi entropy (GRE) is defined in terms of distance in more general functional spaces, suitable to treat the anisotropy of the phase space. In particular the problem encountered with PDFs with infinite Rényi entropy is solved: the counterexample given in (Remark III) has a sufficient finite GRE. In Sec. V the limit cases of the GRE and aspects of the MaxEnt principle with general linear constraints are discussed. In Sec. VI an explicit solution of a maximal GRE problem with scale-invariant linear constraints is given, as a candidate for representation of PDF. For a large class of Fokker-Planck equations, driven by a random external noise, a class of GRE is defined that obeys the generalized H theorem.

II. MOTIVATION

The independent discovery of functionally related entropies, from a pure mathematical axiomatic approach [3], respectively by the study of a heavy tail distribution in physics, motivates the importance of the Rényi entropy, its geometrical reformulation, and generalization via geometrical reformulation.

A wide spectrum of natural, artificial, and social complex systems are now analyzed by means of these two entropies. At present, information theory finds applications in broad areas of science such as in neurobiology [5], the evolution of molecular codes [6], model selection in ecology [7], thermal physics [8], plagiarism detection [9], or quantum information [10].

The purpose of this work is to introduce and study a class of *anisotropic* generalizations of Rényi's entropy (RE). An obvious reason is that, in some problems of complex systems, the phase space may be highly anisotropic. Our terminology is borrowed from the mathematical literature [11], where anisotropic generalizations of the classical L_p, or Lebesgue spaces, are studied. Even more general Banach spaces appear in the study of partial differential equations. Because in general the full phase space of a dynamical system is a direct product of subspaces, with different physical interpretation (like generalized coordinate and conjugated momentum, see below, or the coordinates of driving and driven system, see Sec. VI B), it is natural to expect that the integrability, singularity properties of the distribution function, or even more complex mathematical structures attached to these subspaces are different. Consequently we need to develop a formalism that does not miss exactly this additional information, namely, this direct product structure of the phase space, or the different behavior of the distribution function with respect to its different variables. The anisotropy does not mean the breaking continuous symmetry group; rather it suggests the very different physical nature and mathematical properties of different subspaces of the whole phase space. The first example is the following. In the framework of equilibrium statistical

*gsonnino@ulb.ac.be

†gyorgy.steinbrecher@gmail.com

physics, when the PDF is $f(\mathbf{p},\mathbf{q}) = \exp[-H(\mathbf{p},\mathbf{q})/kT]/Z$ and

$$H(\mathbf{p},\mathbf{q}) = \frac{1}{2m}\sum_{\iota} p_i^2 + \sum_{\iota<j} V(x_i - x_j)$$

the phase space is anisotropic.

Clearly in the case when $V(x)$ vanishes at infinity and is singular in origin the anisotropy is clear; $f(\mathbf{p},\mathbf{q})$ has very different differentiability properties in p_i respectively in x_i, and very different asymptotic decay at large p_i respectively in x_i. A second example, from nonequilibrium statistical mechanics is the following. We consider two small identical particles of mass m, immersed in a fluid governed by the Stokes law, with friction coefficient $\tilde{\gamma}$, and subject to a central two-body force $-\nabla V(r)$. In this case, it is easy to check that the problem is equivalent to the one of a single particle, of reduced mass $\mu = m/2$, immersed in a bath with friction coefficients $\gamma = \tilde{\gamma}/2$. After a few calculations, we see that in this case the statistical properties of the center of mass $R = (r_1 + r_2)/2$ are completely decoupled by those of the relative motion $r \equiv r_1 - r_2$. It is also easily checked that the velocity of the center of mass, i.e., $\dot{R} = (\dot{r}_1 + \dot{r}_2)/2$, obeys the Ornstein Uhlenbeck Fokker-Planck equation, whereas the statistical properties of the relative motion are described by the following Kramer equation:

$$\frac{\partial\rho(\mathbf{r},\mathbf{v},t)}{\partial t} = \left\{-\frac{\partial}{\partial\mathbf{r}}\mathbf{v} + \frac{1}{\mu}\frac{\partial}{\partial\mathbf{v}}[\gamma\mathbf{v} + \nabla V(r)] + \frac{\gamma K_B T}{\mu^2}\frac{\partial^2}{\partial v^2}\right\}\rho(\mathbf{r},\mathbf{v},t)$$

with K_B denoting Boltzmann's constant. Moreover, T and $\mathbf{v}$ stand for temperature and $\dot{\mathbf{r}}$, respectively. The steady-state solution is of the Boltzmann distribution type [12]:

$$\rho_{\text{stat}}(r,v) = \rho_0 \exp[-V(r)/(K_B T)]\exp[-\mu v^2/(2K_B T)]. \tag{1}$$

We are interested in diffusion of particles in the presence of a logarithmically growing potential. This potential has attracted much interest since it serves as a model of several physical systems. For example, charged particles near a long and uniformly charged polymers are subject to a logarithmic potential. Other examples of systems having potentials showing a logarithm and power-law potentials are in condensation in polyelectrolyte solutions [13], nanoparticles with an arbitrary two-dimensional force field [14], and vortex dynamics in the two-dimensional model [15]. For a logarithmically growing potential $V(x) \asymp V_0 \log(x/a)$, for $x \gg a$ and $V_0 > 0$, we get an equilibrium distribution with power-law tail $\rho_{\text{stat}}(x) \sim x^{-V_0/(K_B T)}$ [16]. Reference [17] reports the study of an overdamped motion of a Brownian particle in the logarithmic-harmonic potential. In Ref. [18] can be found the study of the trajectories of a Brownian particle moving in a confining asymptotically logarithmic potential, obeying the overdamped Langevin equation with potential

$$V(r) = g\log(1 + r^2), \tag{2}$$

where the parameter $g > 0$ specifies the strength of the attractive potential. Notice that this situation may be realized in experiments [19,20]. In this case, the stationary solution reads

$$\rho_{\text{stat}}(r,v) = \rho_0 \frac{1}{(1 + r^2)^{g/(K_B T)}} \exp[-\mu v^2/(2K_B T)]. \tag{3}$$

Equation (3) clearly shows that the steady-state solution has a short tail in velocity and long tail in distance r. We shall show that this anisotropy of the PDF in the $\{\mathbf{r},\mathbf{v}\}$ phase space can be retrieved from the MaxEnt principle applied to the new class of entropy here introduced, subject to natural scale-invariant restrictions (see below). This anisotropy of the phase space is manifest in magnetically confined plasmas and, in general, in the case of integrable systems subject or not to perturbations. Another example is provided by the evolution of dynamical systems under the effect of noise where we consider the extended phase space of the system plus source of noise. Also in this quite general example, the extended phase space is anisotropic. Our main task is to propose a generalization of the Rényi entropy (GRE) that, still preserving the additivity, is able to treat these anisotropic situations. We prove that the GRE provides a new set of Liapunov functionals (more exactly H theorems) for a large class of Fokker-Planck equations describing particles transport in a random physical environment. In this case the GRE is always monotonic. We encounter this situation when we study the dynamics of charged particles in random electric and magnetic fields [21].

III. THE RÉNYI ENTROPY AND GEOMETRY OF FUNCTION SPACES

Our generalization of RE results from the reinterpretation of the RE and MaxEnt principles in terms of geometrical concepts in the functional spaces of all PDFs in a given phase space. Let us now proceed by introducing a rigorous definition of the GRE and illustrating its main properties. The definition of GRE starts from the reformulation of RE in the geometric term of norm (or pseudonorm). Starting from the initial axiomatic definition in the case of a discrete probability field, Rényi proved that, for a fixed value of parameter q, appearing in his set of axioms, in particular the "axiom 5′" (see Ref. [3]), the form $S_{R,q}(p_i) = \frac{q}{1-q}\log(\sum_{i=1}^N p_i^q)^{1/q}$ is unique, up to a multiplicative constant. In general, i.e., including the continuum case, when the probability is defined in the terms of some (possibly preferred invariant) measure $dm(\mathbf{x})$ with PDF $\rho(\mathbf{x})$ in the space Ω, the previous definition of RE can be extended, as

$$S_{R,q}[\rho] = \frac{q}{1-q}\log\left\{\int_\Omega [\rho(\mathbf{x})]^q \; dm(\mathbf{x})\right\}^{\frac{1}{q}}. \tag{4}$$

In the particular case when Ω is a set with N elements and m is the counting measure the original form is obtained, and the permutation symmetry from the original Axiom 1 (see Ref. [3]) became the invariance under transformations that preserve the measure m.

In the limit $q \to 1$ we obtain the Shannon-Boltzmann entropy

$$\lim_{q\to 1} S_{R,q}[\rho] = -\int_\Omega \rho(\mathbf{x})\log[\rho(\mathbf{x})]\, dm(\mathbf{x}) := S_{S,B}[\rho]. \tag{5}$$

There are several examples when $S_{S,B}[\rho]$ is infinite while for some q the Rényi (and Tsallis) entropies are finite

Example 1. Let consider $\Omega = (0, +\infty)$:

$$dm_1(x) = dx, \quad dm_2(y) = dy, \tag{6}$$

$$\rho_1(x) = \frac{1 + x^{-\beta}}{(1+x)[\log(2+x)]^{1+\varepsilon}} K_1, \tag{7}$$

$$\rho_2(y) = \frac{1}{y(1+y^{\gamma})[\log(2+1/y)]^{1+\varepsilon}} K_2, \tag{8}$$

where $0 < \varepsilon < 1$, $0 < \beta < 1$, $\gamma > 0$, and $K_{1,2}$ are normalizing factors. It is easy to check that in the $q \to 1$ limit, the Shannon-Boltzmann entropy $S_{S,B}[\rho]$ is infinite but there are ranges of the parameter q such that the Rényi entropy is finite:

$$|S_{R,q}[\rho_1]| < \infty; \quad 1 < q < 1/\beta, \tag{9}$$

$$|S_{R,q}[\rho_2]| < \infty; \quad \frac{1}{1+\gamma} < q < 1. \tag{10}$$

The logarithmic terms are necessary for integrability; see the Appendix. Observe that $S_{R,q}[\rho]$ for $q > 1$ can geometrically be reinterpreted in terms of the norm $\|\rho\|_q$:

$$S_{R,q}[\rho] = \frac{q}{1-q} \log \|\rho\|_q; \qquad \|\rho\|_q = \left[\int_{\Omega} [\rho(\mathbf{x})]^q dm(\mathbf{x}) \right]^{\frac{1}{q}}; \quad q > 1 \tag{11}$$

with the norm $\|\rho\|_q$ defining the distance in the standard Lebesgue $L_q(\Omega, dm)$ spaces [22,23]. However, in the case $0 < q < 1$ we are no longer able to interpret $\|\rho\|_q$ in Eq. (11), as a distance, because it does not satisfy the triangular inequality. This problem may easily be overcome by observing that the functional

$$N_q[\rho] = \int_{\Omega} [\rho(\mathbf{x})]^q dm(\mathbf{x})$$

may be interpreted as a distance [22,24,25], so, for $0 < q < 1$, the definition of entropy reads

$$S_{R,q}[\rho] = \frac{1}{1-q} \log N_q[\rho]$$

with $N_q[\rho]$ playing the role of "distance," but in the more complicated space $L_{q<1}(\Omega, dm)$ [22]. We remark that $N_q[\rho]$ and spaces $L_{q<1}$ were used for studying the steady-state distributions of linear stochastic differential equations [25] or of the stable distributions with a heavy tail [24]. However, it should be mentioned that, due to the complexity of the formalism, mathematicians prefer to transfer $f(x) \in L_{q<1}$ in the standard L_1 space by $f(x) \to |f(x)|^q \in L_1$ [22]. The physical counterpart of this transformation is the Tsallis averaging rule [26]: if $\rho(x)$ is a PDF, then for averaging it use the *escort distribution* $[\rho(x)]^q / \langle [\rho(x)]^q \rangle$, not $\rho(x)$. Under a natural set of restrictions (consisting of normalization, fixing the expectation values, positivity) on $\rho(\mathbf{x})$, the MaxEnt principle, applied to $S_{R,q}(\rho)$ with $q < 1$, generates distribution functions with a heavy tail. The set of the PDFs satisfying the restrictions (this set of PDFs will be denoted with $\mathcal{K}$) is always convex. According to this interpretation, geometrically we have that, for $0 < q < 1$, the MaxEnt PDF is the PDF $\in \mathcal{K}$ corresponding to the maximal distance from the origin $\rho \equiv 0$, and reversely, for $q > 1$, it is the PDF $\in \mathcal{K}$, closest to the origin $\rho \equiv 0$. Despite that the corresponding equations for the Lagrange multipliers could be quite complicated, from general arguments on convex analysis we have that the convexity of the functional $\|\rho\|_q$ with respect to the variable ρ (or, similarly, the concavity of the functional $N_q[\rho]$ with respect to ρ) ensures the uniqueness of the solution of the MaxEnt problem [27]. These properties, as well as the extensivity of RE, will be preserved in our definition of the GRE. However, Axiom 1 in the original work of Rényi, i.e., the symmetry of RE [3], will not be preserved in the GRE. In fact, the invariance under symmetry expresses the maximal lack of specific information. In the case of a PDF depending on many variables imposing symmetry under a general measure-preserving transformation means to forget, for pure mathematical formal reasons, the specific physical meaning of the variables. Another problem that appears in the case of a PDF depending on many variables appears if we consider now the joint distribution function obtained from PDFs $\rho_1(x), \rho_2(y)$ from Example 1 [Eqs. (7) and (8)]:

$$\rho_{1,2}(x,y) = \rho_1(x)\rho_2(y). \tag{12}$$

Due to additivity of the Rényi entropy, we have

$$S_{R,q}[\rho_{1,2}] = S_{R,q}[\rho_1] + S_{R,q}[\rho_2]. \tag{13}$$

Remark 2. The problem is that, according to the Eqs. (9) and (10), the domain of q when $S_{R,q}[\rho_{1,2}]$ is finite is empty.

Probability density functions that have similar asymptotic behavior can appear in processes involving an extreme heavy tail [28]. In contrast to the Rényi entropy, we will see that there exists much finite GRE.

The complicated aspect of the GRE is compensated by the fact that it is possible to obtain more complex MaxEnt distribution functions starting from simple, natural restrictions, in contrast with the Shannon and Rényi entropies where the MaxEnt PDFs come from a tautological transcription of the restrictions. It is worth mentioning that more complex generalizations of the GRE are possible, and study in this respect is in progress [29].

IV. DEFINITION OF THE GRE

One possible way to take into account, in the definition of the generalization of the entropy, the anisotropy, and the inhomogeneity of, or more generally, the symmetry of, the phase space, is already apparent in Eq. (4). Suppose the phase space measure can be written as

$$dm(\mathbf{x}) = n(\mathbf{x}) \prod_{i=1}^{N} dx_i,$$

where $n(\mathbf{x})$ is a locally integrable, not necessary continuous, pseudoscalar density. When $n(\mathbf{x})$ has several singular points or zeros at $\mathbf{x} = \mathbf{a}_k$, $1 \leqslant k \leqslant m$, then there is no global measure-preserving group,

$$\mathbf{x} \to \mathbf{y} = \mathbf{g}(\mathbf{x}, s), \quad n(\mathbf{x}) = n(\mathbf{g}(\mathbf{x}, s)) \det \frac{\partial \mathbf{g}(\mathbf{x}, s)}{\partial \mathbf{x}},$$

whose action is isomorphic with the action of the classical geometrical transformation groups. So we have a way to include the *geometric* anisotropy as well as the inhomogeneity of the phase space.

Another way to introduce the new generalized, additive entropy in the anisotropic phase space is by starting from the definition of the (anisotropic) norm of functions of many variables. According to the authors' knowledge there is no possibility to extend this functional analytic method for inhomogenous case, in order to obtain additive, inhomogenous generalized entropy.

To this end, we introduce the norm of functions in the generalized L_p spaces [11]. Suppose for simplicity that the variable components $\mathbf{x}$ from phase space Ω can be split like $\mathbf{x} = \{x_1, \ldots, x_p, x_{p+1}, \ldots, x_n\}$ or $\mathbf{x} = \{\mathbf{y}, \mathbf{z}\}$ with $\mathbf{y} = \{x_1, \ldots, x_p\}$ and $\mathbf{z} = \{x_{p+1}, \ldots, x_n\}$. The measure is also factorized: $dm(\mathbf{x}) = dm(\mathbf{y}, \mathbf{z}) = dm_y(\mathbf{y}) dm_z(\mathbf{z})$.

More general splitting, or grouping, fits in this scheme, but we restrict ourselves here to the two subsets. In analogy with Eq. (11), we define for $p_y, p_z > 1$ the anisotropic norm as [11]

$$\|f\|_{p_y,p_z} = \left(\int_{\Omega_y} dm_y(\mathbf{y}) \left\{\left[\int_{\Omega_z} dm_z(\mathbf{z}) |f(\mathbf{y},\mathbf{z})|^{p_z}\right]^{1/p_z}\right\}^{p_y}\right)^{1/p_y}. \tag{14}$$

The functional $\|f\|_{p_y,p_z}$ is *convex with respect to f*. The corresponding new entropy is defined as

$$S^{(1)}_{p_y,p_z}[\rho] = \frac{p_y}{1-p_z} \log \|\rho\|_{p_y,p_z}; \quad p_y, p_z > 1. \tag{15}$$

Similarly, for $0 < q_y < 1$, $0 < q_z < 1$ we have a *concave functional* $N_{q_y,q_z}(f)$ that, in analogy with the standard Rényi case, may also be interpreted as the *distance* in the corresponding functional space and the corresponding entropy:

$$N_{q_y,q_z}(f) = \int_{\Omega_y} dm_y(\mathbf{y}) \left[\int_{\Omega_z} dm_z(\mathbf{z}) |f(\mathbf{y},\mathbf{z})|^{q_z}\right]^{q_y}, \tag{16}$$

$$S^{(2)}_{q_y,q_z}[\rho] = \frac{1}{1-q_z} \log N_{q_y,q_z}(\rho); \quad 0 < q_y, q_z < 1. \tag{17}$$

The distance between PDFs ρ_1, ρ_2 is $d(\rho_1,\rho_2) := \|\rho_1 - \rho_2\|_{p_y,p_z}$ for $p_y, p_z > 1$, and $d(\rho_1,\rho_2) := N_{q_y,q_z}(\rho_1 - \rho_2)$ for $0 < q_y, q_z < 1$. Notice that, as shown in Ref. [11], for $p_y, p_z > 1$, the function $d(\rho_1,\rho_2)$ preserves the triangle inequality:[1]

$$d(\rho_1,\rho_3) \leqslant d(\rho_1,\rho_2) + d(\rho_2,\rho_3).$$

The norm $\|\rho\|_{p_y,p_z}$ is a convex functional, whereas $N_{q_y,q_z}[\rho]$ is concave functional.[2] These properties give the intuitive geometric interpretation of MaxEnt problem subject to linear constraints, in the framework of convex analysis [27]. It follows that the GRE is related to the geometry of generalized Lebesgue space L_{p_y,p_z} consisting of the set of functions $f(y,z)$ such that $N_{q_y,q_z}[f]$ or $\|f\|_{p_y,p_z}$ is finite (like the Rényi entropy in the Lebesgue space L_p). Convexity properties imply the uniqueness of the solution of MaxEnt problem with restrictions, despite that the equations for Lagrange multipliers could be very complex [30]. Instead of working with two different definitions of entropy (for two separate cases $0 < q_y, q_z < 1$ and $q_y, q_z > 1$) we prefer to compact the definitions into only one expression. To this end, we observe that, for fixed $\rho \geqslant 0$, both functions $\|\rho\|^{p_y}_{p_y,p_z}$ and $N_{q_y,q_z}(\rho)$ are analytic in the variables p_y, p_z, q_y, q_z (at least near the positive real axis) so we can do a unique analytic continuation in their formula outside their initial domains in the following manner:

$$N_{p_y/p_z,p_z}[\rho] = \|\rho\|^{p_y}_{p_y,p_z}, \tag{18}$$

$$S^{(2)}_{p_y/p_z,p_z}[\rho] = S^{(1)}_{p_y,p_z}[\rho]. \tag{19}$$

For compactness of the formulas, we use for all $q_y, q_z > 0$, and $q_z \neq 1$:

$$N_{q_y,q_z}[\rho] = \int_{\Omega_y} dm_y(\mathbf{y}) \left[\int_{\Omega_z} dm_z(\mathbf{z}) \, |f(\mathbf{y},\mathbf{z})|^{q_z}\right]^{q_y}, \tag{20}$$

$$S^{(2)}_{q_y,q_z}[\rho] = \frac{1}{1-q_z} \log N_{q_y,q_z}(\rho). \tag{21}$$

Example 3. In the case of PDF defined in Eqs. (7), (8), and (12), by using Eqs. (20) and (21) and the definition of the Rényi entropy [Eq. (4)] we obtain

$$S^{(2)}_{q_y,q_z}[\rho_1(y)\rho_2(z)] = \frac{1-q_y q_z}{1-q_z} S_{R,q_y q_z}[\rho_1] + q_y S_{R,q_z}[\rho_2].$$

According to Eqs. (9) and (10) the GRE is finite in the domain

$$\frac{1}{1+\gamma} < q_z < 1, \quad 1 < q_y q_z < 1/\beta.$$

In a similar way

$$S^{(2)}_{q_y,q_z}[\rho_2(y)\rho_1(z)] = \frac{1-q_y q_z}{1-q_z} S_{R,q_y q_z}[\rho_2] \quad + q_y S_{R,q_z}[\rho_1],$$

which is well defined and finite in the domain

$$1 < q_y < 1/\beta, \quad \frac{1}{1+\gamma} < q_y q_z < 1.$$

Remark that Axiom 1, the symmetry invariance under permutations for Rényi entropy [3], appears in a more general form: the invariance under transformations that acts independently in the spaces Ω_y and Ω_z that preserves the measures m_y respectively m_z. In the example given below, passive advection diffusion of a tracer in a turbulent field, the variables y and measure m_y are related to the statistical properties of a macroscopic, external, given turbulent velocity field, while z and the measure m_z give a statistical description of the effects of molecular diffusion. In this case is no symmetry transformation that mixes these very different types of variables; rather, it is meaningful to relate this asymmetry of

[1] A similar method illustrated by Rudin in his textbook [22] may be adopted for showing the validity of the triangle inequality also for $0 < q_y, q_z < 1$.

[2] That is, for $0 \leqslant \alpha \leqslant 1$ we have for $\rho(\mathbf{y},\mathbf{z}) = \alpha\rho_1(\mathbf{y},\mathbf{z}) + (1-\alpha)\rho_2(\mathbf{y},\mathbf{z})$:

$$\|\rho\|_{p_y,p_z} \leqslant \alpha \, \|\rho_1\|_{p_y,p_z} + (1-\alpha) \, \|\rho_2\|_{p_y,p_z},$$

$$N_{q_y,q_z}[\rho] \geqslant \alpha N_{q_y,q_z}[\rho] + (1-\alpha) N_{q_y,q_z}[\rho].$$

the GRE to the hierarchical relation between multiple scales, or causality effects, between spaces Ω_y, Ω_z.

V. PROPERTIES OF THE GRE

Notice that in the limit case $q_y \to 1$ we obtain the standard Rényi entropy

$$S^{(2)}_{1,q_z}[\rho] = \frac{1}{1-q_z}\log\int_{\Omega_y} dm_y(\mathbf{y})\int_{\Omega_z} dm_z(\mathbf{z})|\rho(\mathbf{y},\mathbf{z})|^{q_z},$$

and for $p_z \to 1$ the Shannon entropy,

$$\lim_{p_z\to 1}\lim_{p_y\to 1} S^{(2)}_{p_y,p_z}\{\rho\} = -\int_{\Omega_y} dm_y(\mathbf{y}) \times \int_{\Omega_z} dm_z(\mathbf{z})\rho(\mathbf{y},\mathbf{z})\log\rho(\mathbf{y},\mathbf{z}).$$

We would like to underline that, if we perform the following scaling of the variables $\mathbf{y}\to\alpha\mathbf{y}$, $\mathbf{z}\to\beta\mathbf{z}$, and the measures transform like as $dm_y(\mathbf{y})\to\alpha^{d_1}dm_y(\mathbf{y})$,$dm_z(\mathbf{z})\to\beta^{d_2}\,dm_z(\mathbf{z})$, then the previously defined entropies change by constant. In this context, the variation of the GRE is invariant under scaling, exactly as in the case of the Shannon entropy. In addition, notice that the GRE is extensive, like the Rényi entropy, because the norm $\|\rho\|_{p_y,p_z}$ and the functional $N_{q_y,q_z}[\rho]$ are multiplicative, in analogy with properties of the norm in the L_p space.

A. The MaxEnt principle

The probability distribution functions may be obtained by the MaxEnt principle. Here we shall determine the PDF by generalizing the calculations made for the case of the Shannon entropy, subject to the most general scale-invariant restrictions [31]. To this end, we maximize the GRE, $S^{(1,2)}_{p_y,p_z}[\rho]$, subject to the constraints

$$\int_{\Omega_y} dm_y(\mathbf{y})\int_{\Omega_z} dm_z(\mathbf{z})\rho(\mathbf{y},\mathbf{z})f_k(\mathbf{y},\mathbf{z}) = c_k;\quad 1\leqslant k\leqslant M, \tag{22}$$

$$\rho(\mathbf{y},\mathbf{z})\geqslant 0;\quad f_0(\mathbf{y},\mathbf{z})=1;\quad c_0=1 \tag{23}$$

This means to find the extrema of $N_{q_y,q_z}[\rho]$. From the Kuhn-Tucker theorem for maximization [32], we get

$$\frac{\delta}{\delta\rho(\mathbf{y},\mathbf{z})}\left\{N_{p_y,p_z}[\rho] + \int_{\Omega_y\times\Omega_z} dm_y(\mathbf{y})\,dm_z(\mathbf{z})\rho(\mathbf{y},\mathbf{z}) \times\left[\mu(\mathbf{y},\mathbf{z}) - \sum_{k=0}^{N}\lambda_k f_k(\mathbf{y},\mathbf{z})\right]\right\} = 0,$$
$$\times\mu(\mathbf{y},\mathbf{z})\geqslant 0;\quad \mu(\mathbf{y},\mathbf{z})\rho(\mathbf{y},\mathbf{z}) = \mathbf{0}, \tag{24}$$

where $\mu(y,z)$ and λ_κ and are the multipliers corresponding to the positivity inequality and the linear restrictions, respectively. Here we consider only the case $0<p_y,p_z<1$. We introduce the notations

$$g(\lambda,\mathbf{y},\mathbf{z}) := \frac{1}{p_y p_z}\sum_{k=0}^{N}\lambda_k f_k(\mathbf{y},\mathbf{z}),$$
$$a := \frac{1-p_y}{1-p_y p_z};\quad b := \frac{p_z}{1-p_z};$$
$$h(\lambda,\mathbf{y}) := \int_{\Omega_z} dm_z(\mathbf{z}')|g(\lambda,\mathbf{y},\mathbf{z}')|^{-b}.$$

By straightforward calculations, we get

$$\rho(\lambda,\mathbf{y},\mathbf{z}) = g(\lambda,\mathbf{y},\mathbf{z})^{1/(p_z-1)}|h(\lambda,\mathbf{y})|^{-a}. \tag{25}$$

VI. APPLICATIONS

A. MaxEnt with scale-invariant constraints

(This is an extension of the results from Ref. [31].)
Consider the particular case

$$\int_{\mathbb{R}^2} dy\,dz\rho(y,z)y^2 = c_1;\quad \int_{\mathbb{R}^2} dy\,dz\rho(y,z)z^2 = c_2. \tag{26}$$

From Eq. (25), we obtain (up to a multiplicative constant)

$$\rho(\lambda,y,z) = \frac{[1+(\lambda_1 y)^2]^m}{[1+(\lambda_1 y)^2+(\lambda_2 z)^2]^{\frac{1}{1-p_z}}};\quad m = -\frac{1-3p_z}{2(1-p_z)}\frac{1-p_y}{1-p_y p_z}, \tag{27}$$

which corresponds to a PDF with different tail exponents in the variables y,z. If to Eqs. (26) we add the supplementary restriction

$$\int_{\mathbb{R}^2} dy\,dz\rho(y,z)y^2z^2 = c_3, \tag{28}$$

it is possible to find a combination of the Lagrange multipliers such that, up to a multiplicative constant, we get

$$g(\lambda,\mathbf{y},\mathbf{z}) = (1+a_1y^2)(1+a_2z^2),$$

so that ($c_{1,2}->a_{1,2}$)

$$\rho(y,z) = K\frac{1}{(1+a_1y^2)^{\kappa_y}}\frac{1}{(1+a_2z^2)^{\kappa_z}};\quad \kappa_y = 1/(1-p_yp_z);\quad \kappa_z = 1/(1-p_z). \tag{29}$$

Hence, by putting $c_2 = (1-p_z)/\sigma^2$, and in the limit case $p_z\to 1$ (with the rest of the parameters kept constant), we get

$$\rho(y,z) = K\frac{1}{(1+c_1y^2)^{\kappa_y}}\exp(-z^2/\sigma^2), \tag{30}$$

which is exactly the form of the stationary solution analyzed in our example Eq. (3).

B. H theorems (Liapunov functionals)

Let us now consider the problem of variation of the GRE in a dynamical system, whose microscopic statistical features are described by the Fokker-Planck equation, when additional random effects, due to turbulence at macroscopic scale, are taken into account by a random variable ω. In general, the diffusion term describes the effect of the interactions at the atomic scale. The space of the additional random variable ω will be denoted simply by Ω. Ω corresponds to the previous Ω_y, but now it describes the effect of turbulent environment. The previous space Ω_z is the usual phase space of the dynamical system, with coordinates $\mathbf{z} = \{z_1,\ldots,z_m\}$. The typical example is the passive advection diffusion of tracer by a velocity field with turbulent components and molecular diffusion. Consider the case when the evolution is modeled by the more general advection diffusion stochastic differential equation (SDE) driven by the white noise $\zeta_i(t)$

$$dz_i/dt = V_i(\mathbf{z},\omega) + \zeta_i(t); \quad 1 \leqslant i \leqslant m,$$
$$\langle\zeta_i(t_1)\zeta_j(t_2)\rangle = 2D_{i,j}(\mathbf{z},\omega)\delta(t_1 - t_2), \tag{31}$$

where $V_i(\mathbf{z},\omega)$ and $D_{i,j}(\mathbf{z},\omega)$ satisfy the conditions

$$\frac{\partial V_i(\mathbf{z},\omega)}{\partial z_i} = 0; \quad \frac{\partial D_{i,j}(\mathbf{z},\omega)}{\partial z_i} = 0, \tag{32}$$

where the convention of summation on repeated indices is adopted. The corresponding Fokker-Planck equation, for a fixed ω, is

$$\frac{\partial\rho(t,\omega,\mathbf{z})}{\partial t} = -\frac{\partial}{\partial z_i}(V_i\rho) + \frac{\partial^2}{\partial z_i\partial z_j}(D_{i,j}\rho). \tag{33}$$

Notice that the first condition of Eqs. (32) is satisfied, e.g., by the most general Hamiltonian system, with $m/2$ degrees of freedom (Liouville theorem). This general model contains some important particular cases. When $m = 3$ this corresponds to the passive tracer transport by advection and molecular diffusion, in a turbulent flow, whose statistical properties are encoded in the probability measure $dP(\omega)$. It also describes the stochastic magnetic field line dynamics in a tokamak [21]. For $m = 2$ it may describe the transversal motion (transversal to a constant magnetic field $\mathbf{B}$) of the charged particles in the drift approximation, and subject to a random electric field $-\nabla\phi(\mathbf{z},t,\omega)$ and collisions modeled by white noise $\zeta_i(t)$ [33]:

$$dz_i/dt = \frac{1}{B}e_{i,j}\frac{\partial\phi}{\partial z_j} + \zeta_i(t); \quad i = \overline{1,2},$$
$$\langle\zeta_i(t_1)\zeta_j(t_2)\rangle = \delta_{i,j}\sigma\delta(t_1 - t_2).$$

Here $e_{i,j}$ is the Levi-Civita symbol and σ describes the effects of the collisions.

We prove now the following important theorem, which is a sort of H theorem describing the tendency of the GRE to increase in time. In general, we define the Liapunov function $L(t) := \int_\Omega dP(\omega)[\int_\Gamma d^m\mathbf{x}(\rho(t,\omega,\mathbf{x}))^{p_z}]^{p_y}$ where $\rho(t,\omega,\mathbf{z})$ is the solution of the Fokker-Planck equation [Eq. (33)] for a fixed ω. Then we have the following.

Proposition. Under the conditions of Eqs. (32), $\frac{d}{dt}L(t)(p_z - 1) > 0$, and the corresponding GRE $S^{(2)}_{q_y,q_z}[\rho] = \frac{1}{1-q_z}\log L(t)$ is nondecreasing in time.

Proof. We start from the definition of the GRE and differentiate the expression with respect to time. Then we use the Fokker-Planck equation for the time derivative of ρ, and after integration by part in the z coordinate, and by taking into account Eqs. (32), we obtain

$$\frac{d}{dt}L(t) = \int_\Omega dP(\omega)M(\omega,t)$$
$$\times \int_\Gamma d^m\mathbf{z}(\rho(t,\omega,\mathbf{z}))^{p_z-2}\frac{\partial\rho}{\partial z_i}\frac{\partial\rho}{\partial z_j}D_{i,j},$$
$$M(\omega,t) = -p_y p_z(p_z - 1)\left[\int_\Gamma d^m\mathbf{z}'(\rho(t,\omega,\mathbf{z}'))^{p_z}\right]^{p_y-1}$$

with $\frac{\partial\rho}{\partial z_i}\frac{\partial\rho}{\partial z_j}D_{i,j} \geqslant 0$ from the second law of thermodynamics.

VII. CONCLUSIONS

In conclusion, we have introduced a generalization of the Rényi entropy (GRE) that, still preserving the additivity, is able to treat dynamical systems in a highly anisotropic phase space. This is the case of magnetically confined plasmas or of integrable systems subject to perturbations. The anisotropy of the PDF in the phase space can be retrieved from the MaxEnt principle applied to the GRE, subject to natural scale-invariant restrictions. In these situations, the PDFs may show different tail exponents in the variables; this property belongs only to the GRE and not to the standard Rényi entropy. We have also seen that the Rényi and Shannon entropies are reobtained by the GRE as limit cases. Even though the extensivity of the RE is preserved in the GRE, the symmetry of the RE (Axiom 1 of the RE), is not completely preserved in our generalized version. The symmetry group of Rényi entropy, i.e., the measure-preserving transformations of $\Omega_y \times \Omega_z$, splits into a direct product of measure-preserving transformations of Ω_y, respectively. This is the "price" we have to pay to treat dynamical systems in highly anisotropic phase space.

The functionals that appear in the definition of the GRE may be interpreted as the distance in the corresponding functional space and in a wide range of the parameters p_y, p_z have useful concavity, respectively convexity, properties. Again, when the evolution of the system is modeled by the general advection diffusion SDE driven by white noise, we proved the validity of a sort of H theorem, which results in being satisfied by the GRE when the velocity flows and the diffusion coefficients are *divergenceless* in the phase space of the dynamical system [see Eq. (30)].

This work gives several perspectives. Through the thermodynamical field theory (TFT) [34], it is possible to estimate the PDF when the nonlinear contributions cannot be neglected [35]. The next task should be to establish the relation between the reference, stationary PDF, derived by the MaxEnt principle applied to GRE, subject to scale-invariant restrictions, with the ones found by the TFT.

ACKNOWLEDGMENTS

G.S. is very grateful to Malek Mansour, Université Libre de Bruxelles, for his scientific suggestions and for his help in the development of this work. Gy.S. acknowledges J. Misquich for useful discussions. We also acknowledge M. Van Schoor and D. Van Eester of the Royal Military School.

APPENDIX

1. Proof of Eq. (9)

Denote

$$u_1(x) = \frac{1+x^{-\beta}}{(1+x)[\log(2+x)]^{1+\varepsilon}}.$$

From Eqs. (4) and (7) we have to study the $\delta \to 0$ and $A \to \infty$ limits in

$$I_{q,\delta,A} = \int_{\delta}^{A} [u_1(x)]^q\, dx. \tag{A1}$$

We have $I_{q,\delta,A} = J^{(1)}_{q,\delta,A} + J^{(2)}_{q,\delta,A}$ where

$$J^{(1)}_{q,\delta} = \int_{\delta}^{1} [u_1(x)]^q\, dx, \tag{A2}$$

$$J^{(2)}_{q,A} = \int_{1}^{A} [u_1(x)]^q\, dx. \tag{A3}$$

Remark that

$$\lim_{x\to 0} \frac{[u_1(x)]^q}{x^{-\beta q}} = \frac{1}{[\log(2)]^{q(1+\varepsilon)}}.$$

It follows that for $q > 1$ and $0 < \beta < 1$ the convergence/divergence in $J^{(1)}_{q,\delta}$ for $\delta \to 0$ is the same as in the integral $\int_{\delta}^{1} x^{-\beta q}\, dx$, while $J^{(2)}_{q,A}$ has finite limit for $A \to \infty$, which proves Eq. (9).

Remark 4. Moreover, we observe that also in the limit $q \searrow 1$ the normalization integral $I_{1,\varepsilon,A}$ has a finite limit, due to the logarithmic term.

Indeed, $J^{(1)}_{q,\delta}$ from Eq. (A2) is finite also for $q = 1$ in the limit $\delta \to 0$, because $\beta < 1$. The behavior for large A in the integral $J^{(2)}_{q,A}$ is the same as for the integral $\int_1^A \frac{1}{(2+x)\ [\log(2+x)]^{1+\varepsilon}}\, dx$, which is finite.

2. Proof of Eq. (10)

Denote

$$u_2(y) = \frac{1}{y(1+y^{\gamma})[\log(2+1/y)]^{1+\varepsilon}}.$$

Similarly,from Eqs. (4) and (8) we have to study the $\delta \to 0$ and $A \to \infty$ limits in

$$F_{q,\delta,A} = \int_{\delta}^{A} [u_2(y)]^q\, dy. \tag{A4}$$

We decompose $F_{q,\delta,A} = G^{(1)}_{q,\delta} + G^{(2)}_{q,A}$ where

$$G^{(1)}_{q,\delta} = \int_{\delta}^{1} [u_2(y)]^q\, dy, \tag{A5}$$

$$G^{(2)}_{q,A} = \int_{1}^{A} [u_2(y)]^q\, dy. \tag{A6}$$

For $\gamma > 0$ the integral $G^{(1)}_{q,\delta}$ behaves for $\delta \to 0$ like the integral $\int_{\delta}^{1} y^{-q}\, dy$, so converge for $0 < q < 1$ and diverge for $q > 1$. The behavior for $A \to \infty$ of $G^{(2)}_{q,A}$ is the same as for $\int_1^A y^{-q(1+\gamma)}\, dy$: it is finite for $q(1+\gamma) > 1$ and it is divergent for $q(1+\gamma) < 1$, which proves Eq. (10).

Remark 5. Similarly to Remark 4 we observe that also in the limit $q \nearrow 1$ the normalization integral $F_{1,\varepsilon,A}$ has finite limit, due to the logarithmic term.

For proof, we observe that $G^{(2)}_{q,A}$ for $q = 1$ and $A \to \infty$ is finite. The $\delta \to 0$ behavior of $G^{(1)}_{q,\varepsilon}$ for $q = 1$ is similar to the integral $\int_{\delta}^{1} \frac{1}{y[\log(2+1/y)]^{1+\varepsilon}}\, dy$, or that of integral $\int_{\delta}^{0.5} \frac{1}{y[\log(1/y)]^{1+\varepsilon}}\, dy$, which is finite.

3. Proof of example with infinite Boltzmann-Shannon entropy

We prove that the Boltzmann-Shannon entropy for distributions from Eqs. (7) and (8) is infinite. First, we observe that according to Remarks 4 and 5, both distributions from Eqs. (7) and (8) are integrable and the resulting convergence or divergence is not influenced by the normalizing constants $K_{1,2}$. From Eqs. (5) and (7) results

$$S_{S,B}[\rho_1] = k_1 + \lim_{\delta\to 0} H^{(1)}_{\delta} + \lim_{A\to\infty} H^{(2)}_{A}, \tag{A7}$$

where in the following we denote by $k_{1,\ldots}$ some finite constants and

$$H^{(1)}_{\delta} = -\int_{\delta}^{1} u_1(x) \log u_1(x)\, dx, \tag{A8}$$

$$H^{(2)}_{A} = -\int_{1}^{A} u_1(x) \log u_1(x)\, dx. \tag{A9}$$

Remark that the dominant term in the integrand of $H^{(1)}_{\delta}$ is of the form $x^{-\beta} \log x^{-\beta}$, so $\lim_{\delta\to 0} H^{(1)}_{\delta}$ is finite.

Denote

$$v(x) = \frac{1}{(2+x)\ [\log(2+x)]^{\varepsilon}}, \tag{A10}$$

and remark that

$$\lim_{x\to\infty} \frac{u_1(x) \log u_1(x)}{v(x)} = 1, \tag{A11}$$

$$\int_1^A v(x)\, dx = O([\log A]^{1-\varepsilon}). \tag{A12}$$

Then from Eqs. (A7) and (A9)–(A12) we obtain

$$H^{(2)}_{A} = O([\log A]^{1-\varepsilon}). \tag{A13}$$

In conclusion, from Eqs. (A7) results that the Shannon-Boltzmann entropy of the distribution $\rho_1(x)$ from Eq. (7) is infinite. In order to prove that the Shannon-Boltzmann entropy of the distribution $\rho_2(y)$ is infinite, we use Eqs. (5) and (8). Similarly to Eq. (A7) results

$$S_{S,B}[\rho_2] = k_3 + \lim_{\delta\to 0} L^{(1)}_{\delta} + \lim_{A\to\infty} L^{(2)}_{A}, \tag{A14}$$

where

$$L^{(1)}_{\delta} = -\int_{\delta}^{1} u_2(y) \log u_2(y)\, dy, \tag{A15}$$

$$L^{(2)}_{A} = -\int_{1}^{A} u_2(y) \log u_2(y)\, dy. \tag{A16}$$

Remark that

$$\Big|\lim_{A\to\infty} L_A^{(2)}\Big| < \infty \tag{A17}$$

because $\gamma > 0$. Denote

$$s(y) = \frac{1}{y[\log(1/y)]^{\varepsilon}}. \tag{A18}$$

It is easy to see

$$\lim_{y\to 0} \frac{u_2(y)\log u_2(y)}{s(y)} = -1. \tag{A19}$$

On the other hand, by the substitution $y = \exp(-t)$ we obtain

$$\int_{\delta}^{1/2} s(y)\,dy = \frac{1}{1-\varepsilon}\left[\left(\log\frac{1}{\delta}\right)^{1-\varepsilon} - (\log 2)^{1-\varepsilon}\right]. \tag{A20}$$

From Eqs. (A15), (A19), and (A20) results that

$$L_{\delta}^{(1)} = -O\left(\log\frac{1}{\delta}\right)^{1-\varepsilon}, \tag{A21}$$

which proves that $S_{S,B}[\rho_2] = -\infty$.

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948)
[2] T. Maszczyk and X. Duch, Lect. Notes Comput. Sci. **5097**, 643 (2008).
[3] A. Rényi, in *Proceedings of the Fourth Berkeley Symposium on Mathematical, Statistical Probability, Berkeley and Los Angeles* (University of California Press, Oakland, CA, 1961), Vol. 1, pp. 547–561.
[4] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[5] F. Rieke, D. Warland, R. Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1997).
[6] J. P. Huelsenbeck, F. Ronquist, R. Nielsen, and J. P. Bollback, Science **294**, 2310 (2001).
[7] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. (Springer Science, New York, 2002).
[8] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[9] C. H. Bennett, M. Li, and B. Ma, Sci. Am. **288**, 76 (2003).
[10] F. Franchini, A. R. Its, and V. E. Korepin, J. Phys. A **41**, 25302 (2008); A. R. Its and V. E. Korepin, Theor. Math. Phys. **164**, 1136 (2010).
[11] O. V. Besov, V. P. Il'in, and S. M. Nikol'skii, *Integral Representations of Functions and Embedding Theorems* (Ed. Nauka, Moscow, 1975), pp. 9–40 (in Russian).
[12] H. Risken, *The Fokker-Planck Equation. Methods of Solution and Applications* (Springer, Berlin, 1996).
[13] G. S. Manning, J. Chem. Phys. **51**, 924 (1969).
[14] A. E. Cohen, Phys. Rev. Lett. **94**, 118102 (2005).
[15] A. J. Bray, Phys. Rev. E **62**, 103 (2000).
[16] S. Marksteiner, K. Ellinger, and P. Zoller, Phys. Rev. A **53**, 3409 (1996).
[17] A. Ryabov, M. Dierl, P. Chvosta, M. Einax, and P. Maass, J. Phys. A **46**, 075002 (2013).
[18] A. Dechant, E. Lutz, D. A. Kessler, and E. Barkai, Phys. Rev. E **85**, 051124 (2012).
[19] D. A. Kessler and E. Barkai, Phys. Rev. Lett. **105**, 120602 (2010).
[20] V. Blickle and C. Beckinger, Nat. Phys. **8**, 103 (2012).
[21] M. Vlad, F. Spineanu, J. H. Misguich, and R. Balescu, Phys. Rev. E **53**, 5302 (1996); **54**, 791 (1996); M. Negrea, I. Petrisor, and R. Balescu, *ibid.* **70**, 046409 (2004); M. Vlad, F. Spineanu, J. H. Misguich, and R. Balescu, *ibid.* **67**, 026406 (2003).
[22] W. Rudin, *Real and Complex Analysis*, 3rd ed. (McGraw Hill, New York, 1987), p. 74.
[23] M. Reed and B. Simon, *Functional Analysis, Methods of Modern Mathematical Physics*, Vol. 1 (Academic Press, New York, 1981).
[24] H. Luschgy and G. Pagès, Electr. Comm. Prob. **13**, 422 (2008).
[25] G. Steinbrecher, X. Garbet and B. Weyssow, arXiv:1007.0952v1 (2010).
[26] C. Tsallis, R. S. Mendes, and A. R. Plastino, Physica A **261**, 534 (1998).
[27] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2004).
[28] G. Steinbrecher and B. Weyssow, Phys. Rev. Lett. **92**, 125003 (2004); T. L. Rhodes, R. A. Moyer, R. Groebner, E. J. Doyle, R. Lehmer, W. A. Peebles, and C. L. Rettig, Phys. Lett. A **253**, 181 (1999).
[29] G. Sonnino and G. Steinbrecher (unpublished).
[30] G. Comte, *Entropy and Quantitative Transversality*, *Encyclopedia of Mathematical Physics* (Academic Press, New York, 2006), pp. 237–242.
[31] G. Sonnino, G. Steinbrecher, A. Cardinali, A. Sonnino, and M. Tlidi, Phys. Rev. E **87**, 014104 (2013).
[32] S. Stephen and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2004), p. 244.
[33] F. Spineanu and M. Vlad, J. Plasma Phys. **54**, 333 (1995); M. Vlad, F. Spineanu, J. H. Misguich, and R. Balescu, Phys. Rev. E **61**, 3023 (2000); M. Vlad, F. Spineanu, and S. Benkadda, Phys. Rev. Lett. **96**, 085001 (2006); Phys. Plasma **15**, 032306 (2008).
[34] G. Sonnino, Phys. Rev. E **79**, 051126 (2009).
[35] G. Sonnino, Eur. Phys. J. D **62**, 81 (2011).