HAL open science

Riemannian-geometric entropy for measuring network complexity

Roberto Franzosi, Domenico Felice, Stefano Mancini, Marco Pettini

To cite this version:

Roberto Franzosi, Domenico Felice, Stefano Mancini, Marco Pettini. Riemannian-geometric entropy for measuring network complexity. Physical Review E , 2016, 93 (6), pp.062317. 10.1103/PhysRevE.93.062317 . hal-01361918

HAL Id: hal-01361918

https://hal.science/hal-01361918v1

Submitted on 7 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Riemannian-geometric entropy for measuring network complexity

Roberto Franzosi,[1,*] Domenico Felice,[2,3,†] Stefano Mancini,[2,3,‡] and Marco Pettini[4,5,§]
[1]*QSTAR and INO-CNR, largo Enrico Fermi 2, I-50125 Firenze, Italy*
[2]*School of Science and Technology, University of Camerino, I-62032 Camerino, Italy*
[3]*INFN-Sezione di Perugia, Via A. Pascoli, I-06123 Perugia, Italy*
[4]*Aix-Marseille University, Marseille, France*
[5]*CNRS Centre de Physique Théorique UMR7332, 13288 Marseille, France*
(Received 29 July 2015; revised manuscript received 30 May 2016; published 27 June 2016)

A central issue in the science of complex systems is the quantitative characterization of complexity. In the present work we address this issue by resorting to information geometry. Actually we propose a constructive way to associate with a—in principle, any—network a differentiable object (a Riemannian manifold) whose volume is used to define the entropy. The effectiveness of the latter in measuring network complexity is successfully proved through its capability of detecting a classical phase transition occurring in both random graphs and scale-free networks, as well as of characterizing small exponential random graphs, configuration models, and real networks.

DOI: 10.1103/PhysRevE.93.062317

I. INTRODUCTION

Complex systems and phenomena are dealt with in many scientific domains. According to the domain of interest, different definitions of complexity and methods of measuring it have been proposed and are continuously being proposed since the science of complexity is still growing rapidly [1,2]. The literature on this topic is so vast that any attempt at providing an exhaustive bibliography would be out of place and a very hard task here. As a consequence, instead of trying to list them all, let us note that the many ways of measuring complexity belong to a restricted number of categories. In particular, the attempts at quantifying the degree of organization of a complex system often resort to some definition of an entropy function stemming from the "archetype" represented by Shannon's information entropy [3,4]. The latter has its precursor—at least from the point of view of physics—in Boltzmann's entropy of kinetic theory. In fact, Shannon's information entropy is equivalent to the negative Boltzmann entropy, as proven by L. Brillouin [5].

Among the different statistical-mechanical approaches to networks hitherto proposed, one is the class of models with hidden variables [6]; here, the approach starts with a set of N independent nodes and a general hidden variable X; then an undirected network is generated by (i) assigning to each node i a variable X_i, independently drawn from the probability $p(X)$; and (ii) creating for each pair of vertices i and j, with respective hidden variables X_i and X_j, an undirected link with probability $p(X_i,X_j)$. So, given the independent assignment of hidden variables and links among nodes, correlated random networks are generated without either loops or multiple links, where the degree distribution and the correlation properties of the network are encoded in the two functions $p(X)$ and $p(X_i,X_j)$ $(i,j = 1,\ldots,N)$. In the present work, we consider random variables as arbitrary hidden variables [7] sitting on the nodes, and their correlations are seen as weighted links among the nodes, again. The difference from the previous approach consists of focusing the attention on the knowledge of some parameters characterizing the hidden variables. All the information about the system is retained in these parameters. In particular, given the information on the variances and covariances of the multiple hidden variables, a multivariate Gaussian probability distribution can be derived to describe the whole given network, by means of the maximum entropy principle [8]. Thus a parameter space is associated with any given network. This space encodes all the information on the structure of the associated network. Note that a similar way of associating a probability distribution with a network is that of probabilistic graph models [9]. Actually, Gaussian networks are extensively used in many applications, ranging from neural networks, to wireless communication, from proteins to electronic circuits, and so on. Then, by resorting to information geometry [10], the space of the accessible values of the parameters of a given network can be endowed with the Fisher-Rao metric, so defining a Riemannian manifold. In analogy with statistical mechanics [11], this manifold is the space of all possible states of the associated network, that is, analogous to the phase space of a physical system. By exploiting this analogy, we may define an entropy function as the logarithm of the Riemannian volume of the manifold associated with the given network.

The first step in this direction was put forward in [12]; in that paper, we found that the geometric entropy associated with the Fisher-Rao metric reflects the topological features of the network: it is an increasing function of the simplex dimension. However, as discussed in the following, this approach cannot be constructively applied to networks having more than a few nodes.

Then, in [13] a new metric—obtained by a suitable "deformation" of the standard Fisher-Rao metric of information geometry—was proposed which allows us to constructively lift the properties of any given network to the geometric structure of a manifold. In this way, we associated a differentiable system (Riemannian manifold) with a discrete system (network) through the description of the network by a set of probability distribution functions. Among the widely used probabilistic methods in the literature, the random walk method is worthy

*roberto.franzosi@ino.it
†domenico.felice@unicam.it
‡stefano.mancini@unicam.it
§pettini@cpt.univ-mrs.fr

 ©2016 American Physical Society

of mention [14]. Here the Green function, meaning the transition amplitude from one vertex to another obtained by accounting for all possible walks, gives rise to a metric [15], thus allowing for a geometric approach as well. However, the main difference is that by considering random walks, the (transition) probability is given through the adjacency matrix, while in our case probabilities are given through the Gaussian joint distribution of random variables sitting on nodes of the network. In addition, as clearly shown in [13], such a geometric entropy is able to detect the classical transition in random graphs (RGs) predicted by the Erdös-Rényi theorem [16,17].

Here we want to propose that this entropy qualifies as a network complexity measure. To this end, we further the study of the RG model in [13] and then validate our measure in complex networks.

The layout of the paper is as follows. For the sake of self-consistency and readability, the mathematical framework is reported in Secs. II and III. In particular, in Sec. II we briefly recall the relation between the Gaussian statistical model and the underlying network, putting forward the metric structure of the associated manifold. In Sec. III we present the geometric measure of complexity as the logarithm of the Riemannian volume of the manifold. Toward the end of validating this complexity measure, we check it against both theoretical and empirical networks, spreading from RGs to real networks by stepwise increasing "the degree of complexity." In Sec. IV we report a strong validation point of our complexity measure of networks, the first account of which was given in [13]. This first validation was obtained by computing our geometric entropy for the Erdös-Rényi phase transition in RGs, that is, by checking how it performs against a rigorous analytical result on networks. In Sec. V we extend our investigation to different network models. We start considering theoretical models of exponential random graphs and configuration models which are well known and studied in the literature. Next, we show that our geometric entropy is also able to detect the emergence of a giant component as predicted by the Molloy and Reed criterion for scale-free networks. Then we apply our complexity measure to real networks and compare the outcomes so obtained to the results that have been reported in the literature for these same systems using other complexity measures. Section VI is devoted to possible future developments concerning the possibility of predicting the stability of a network system. Conclusions are drawn in Sec. VII.

II. INFORMATION GEOMETRIC MODEL

Usually in mathematics in order to get information on a geometric object one endows it with a superstructure (e.g., bundles over manifolds, coverings over topological spaces). Likewise we endow a network with a statistical Riemannian manifold. This can be obtained basically via two steps: first, by interpreting a network as an undirected graph without loops on the nodes, one can associate suitable random variables with the nodes and associate the correlations among these random variables with the links among the nodes given by the adjacency matrix A; then, one can resort to methods of Information Geometry [10] to associate a statistical Riemannian manifold with the network.

So, let us consider a set of n real-valued random variables $X_1,\ldots,X_n$ distributed according to a multivariate Gaussian probability distribution (assumed, for the sake of simplicity, to have zero mean)

$$p(x;\theta) = \frac{1}{\sqrt{(2\pi)^n \det C}} \exp\left[-\frac{1}{2}x^t C^{-1} x\right], \tag{1}$$

where $x^t = (x_1,\ldots,x_n) \in \mathbb{R}^n$, with t denoting the transposition. Furthermore, $\theta^t = (\theta^1,\ldots\theta^m)$ are the real-valued parameters characterizing the above probability distribution function, namely, the entries in the covariance matrix C. As a consequence, $m = n(n+1)/2$.

Next consider the family $\mathcal{P}$ of such probability distributions,

$$\mathcal{P} = \{p_\theta = p(x;\theta) | \theta^t = (\theta^1,\ldots\theta^m) \in \Theta\},$$

where $\Theta \subseteq \mathbb{R}^m$. Upon requiring the mapping $\theta \to p_\theta$ to be injective, $\mathcal{P}$ becomes an m-dimensional statistical model on $\mathbb{R}^n$. The open set Θ results to be defined as

$$\Theta = \{\theta \in \mathbb{R}^m | C(\theta) > 0\}, \tag{2}$$

and we refer to this as the parameter space of the statistical model $\mathcal{P}$.

Since any element $p(x;\theta) \in \mathcal{P}$ is univocally characterized by the parameter vector θ, it follows that the mapping $\varphi : \mathcal{P} \to \Theta$ defined by $\varphi(p_\theta) = \theta$ is a coordinate chart. So, $\varphi = [\theta^i]$ can be considered a local coordinate system for $\mathcal{P}$. Then $\mathcal{P}$ can be turned into a C^∞ differentiable manifold by assuming parametrizations that are C^∞ [10].

Given an m-dimensional statistical model $\mathcal{P} = \{p_\theta | \theta \in \Theta\}$, its Fisher information matrix in θ is the $m \times m$ matrix $G(\theta) = [g_{\mu\nu}]$, whose entries are defined by

$$g_{\mu\nu}(\theta) := \int_{\mathbb{R}^n} dx\ p(x;\theta)\partial_\mu \log p(x;\theta)\partial_\nu \log p(x;\theta), \tag{3}$$

with $\partial_\mu \equiv \frac{\partial}{\partial\theta^\mu}$. The matrix $G(\theta)$ turns out to be symmetric, positive definite and provides a Riemannian metric for the parameter space Θ [10].

For our case the integral in Eq. (3) is Gaussian and can be computed as

$$\frac{1}{\sqrt{(2\pi)^n \det C}} \int dx f_{\mu\nu}(x) \exp\left[-\frac{1}{2}x^t C^{-1} x\right] = \exp\left[\frac{1}{2}\sum_{i,j=1}^{n} c_{ij}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\right] f_{\mu\nu}|_{x=0}, \tag{4}$$

where

$$f_{\mu\nu} := \partial_\mu \log[p(x;\theta)]\ \partial_\nu \log[p(x;\theta)], \tag{5}$$

and the exponential stands for a power-series expansion over its argument (the differential operator). The derivative of the logarithm reads

$$\partial_\mu \log[p(x;\theta)] = -\frac{1}{2}\left[\frac{\partial_\mu(\det C)}{\det C} + \sum_{\alpha,\beta=1}^{n} \partial_\mu\left(c^{-1}_{\alpha\beta}\right)x_\alpha x_\beta\right], \tag{6}$$

where $c^{-1}_{\alpha\beta}$ denotes the entries of the inverse of the covariance matrix C.

The computational complexity of the metric components in Eq. (3) can be readily estimated. Indeed, the well-known formulas

$$\partial_\mu C^{-1}(\theta) = C^{-1}(\theta)\big(\partial_\mu C(\theta)\big)C^{-1}(\theta),$$
$$\partial_\mu(\det C(\theta)) = \det C(\theta)\ \mathrm{Tr}(C(\theta)\,\partial_\mu(C(\theta)))$$

require the calculation of $n(n+1)$ derivatives, with respect to the variables $\theta \in \Theta$, in order to work out the derivative of the logarithm in (6). Then, to obtain the function $f_{\mu\nu}$ in (5), we have to calculate $O(n^4)$ derivatives. With increasing n this becomes a daunting task, even when afforded numerically.

An alternative to the Fisher-Rao metric

In order to overcome the difficulty of computing the components of the Fisher-Rao metric, we follow [13] and define a (pseudo)-Riemannian metric on the parameter space Θ which accounts as well for the network structure given by the adjacency matrix A. To this end we consider first a trivial network with a null adjacency matrix that is associated with a set of n independent Gaussian random variables X_i. Note that in this particular case, the covariance matrix in (1) is a diagonal matrix with entries given by $\theta^i := \mathbb{E}(X_i^2)$. Let us denote this matrix $C_0(\theta)$. So, employing Eqs. (2) and (3), a statistical Riemannian manifold $\mathcal{M} = (\Theta, g)$, with

$$\Theta = \{\theta \equiv (\theta^1, \ldots \theta^n) | \theta^i > 0\},$$
$$g = \frac{1}{2}\sum_{i=1}^{n}\left(\frac{1}{\theta^i}\right)^2 d\theta^i \otimes d\theta^i, \tag{7}$$

is associated with the bare network.

Let us remark that the entries g_{ii} of the metric g in (7), worked out in [12], depend on the entries in the matrix $C_0(\theta)$. In fact, the ii entries of the inverse matrix of $C_0(\theta)$ are given by $c_{ii}^{-1} = \frac{1}{\theta^i}$. Then, from (7) it is evident that $g_{ii} = \frac{1}{2}(c_{ii}^{-1})^2$. Inspired by this functional form of g, we associate a (pseudo)-Riemannian manifold with any network $\mathcal{X}$ with a nonvanishing adjacency matrix A by "deforming" the manifold $\mathcal{M}$ in (7) via the map $\psi_{C_0} : \mathrm{A}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$ defined by

$$\psi_{C_0(\theta)}(A) := C_0(\theta) + A. \tag{8}$$

By $\mathrm{A}(n,\mathbb{R})$ we denote the set of symmetric $n \times n$ matrices over $\mathbb{R}$ with vanishing diagonal elements that can represent any simple undirected graph. Therefore, the manifold associated with a network $\mathcal{X}$, with adjacency matrix A, is $\widetilde{\mathcal{M}} = (\widetilde{\Theta}, \widetilde{g})$. Here it is

$$\widetilde{\Theta} := \{\theta \in \Theta \mid \psi_{C_0(\theta)}(A) \text{ is nondegenerate}\} \tag{9}$$

and $\widetilde{g} = \sum_{\mu\nu} \widetilde{g}_{\mu\nu} d\theta^\mu \otimes d\theta^\nu$ with components

$$\widetilde{g}_{\mu\nu} = \tfrac{1}{2}\big(\psi_{C_0(\theta)}(A)^{-1}_{\mu\nu}\big)^2, \tag{10}$$

where $\psi_{C_0(\theta)}(A)^{-1}_{\mu\nu}$ are the entries in the inverse of the matrix $\psi_{C_0(\theta)}(A)$.

III. A MEASURE OF NETWORK COMPLEXITY

We now define a statistical measure of the complexity of a network $\mathcal{X}$ with adjacency matrix A and associated manifold $\widetilde{\mathcal{M}} = (\widetilde{\Theta}, \widetilde{g})$ as

$$\mathcal{S} := \ln \mathcal{V}(A), \tag{11}$$

where $\mathcal{V}(A)$ is the volume of $\widetilde{\mathcal{M}}$ evaluated from the element

$$\nu_g = \sqrt{|\det \widetilde{g}(\theta)|}\, d\theta^1 \wedge \ldots \wedge d\theta^n. \tag{12}$$

Note, however, that in this way $\mathcal{V}(A)$ turns out to be ill defined. In fact, the set $\widetilde{\Theta}$ in Eq. (9) is not compact because the variables θ^i are unbound from above. Furthermore, from Eq. (10), $\det \widetilde{g}(\theta)$ diverges since $\det \psi_{C_0(\theta)}(A)$ approaches 0 for some θ^i.

Thus, as commonly happens [18], we regularize the volume as

$$\mathcal{V}(A) := \int_{\widetilde{\Theta}} \Upsilon(\psi_{C_0(\theta)}(A))\ \nu_g, \tag{13}$$

where $\Upsilon(\psi_{C_0(\theta)}(A))$ is any suitable "infrared" and "ultraviolet" regularizing function, i.e., providing a kind of compactification of the parameter space and excluding the contributions of θ^i making $\det \widetilde{g}(\theta)$ divergent. Theoretically, a regularizing function $\Upsilon(\psi_{C_0(\theta)}(A))$ might be devised by taking into account particular structures of the integration set, (9), and the functional relation, (10). In practice, a very suitable function has been built in Ref. [12]; it reads as

$$\Upsilon(C(\theta)) = \log[1 + \det(C(\theta))^n]\ e^{-\mathrm{tr}C(\theta)} \tag{14}$$

when the covariance matrix $C(\theta)$ has only 1 or 0 as off-diagonal entries. Here, the logarithm hales contributions of θ^i that make $\det C$ diverge at the lower bound of the parameter space, while the exponential fixes the problem of noncompact integration space. In this paper, we tackle networks with weighted links among the nodes. Still the functional type of the regularizing function is as in (14).

The definition, (11), is inspired by the microcanonical definition of entropy in statistical mechanics, which is proportional to the logarithm of the volume of the Riemannian manifold associated with the underlying dynamics [11]. Of course we need to validate the proposed measure of network complexity defined in Eq. (11). Though in principle any measure of complexity is admissible, we may wonder how to assess its effectiveness. The first step is to check a complexity measure against a system which makes a clear jump in complexity as some parameter is varied. In physics a paradigmatic situation is offered by phase transitions (a snowflake is intuitively more complex than a drop of water). Applied to networks this leads us to consider the classical Erdös-Rényi phase transition in RGs [16,17]. Then, moving on from RGs, more complex networks can be considered and the proposed measure of complexity compared with other known measures. These are the subjects of the following sections.

IV. THE ERDÖS-RÉNYI PHASE TRANSITION

One of the basic models of RGs is the *uniform random graph* $\mathbb{G}(n,k)$. This is devised by choosing with uniform probability a graph from the set of all the graphs having n vertices and k edges, with k a nonnegative integer. We can think of $\mathbb{G}(n,k)$ as a process that evolves by adding the edges one at a time. When k has the same order of magnitude as n, the evolution of $\mathbb{G}(n,k)$ from $k = 0$ to $k = \binom{n}{2}$ yields, according to

Erdös-Rényi theorem [16], a *phase transition*, revealing itself in the rapid growth with k of the size of the largest component (number of vertices fully connected by edges). Specifically, the structure of $\mathbb{G}(n,k)$ when the expected degree of each of its vertices is close to 1, i.e., $k \sim n/2$, shows a jump: the order of magnitude of the size of the largest component of $\mathbb{G}(n,k)$ inreases rapidly, asymptotically almost surely (a.a.s.), from $\log n$ to n, if k has the same order of magnitude as n. In fact, if $k < n/2$, as the process evolves, the components of $\mathbb{G}(n,k)$ [the largest of them being a.a.s. of size $O(\log n)$] merge mainly by attaching small trees; thus they grow slowly and quite smoothly [17]. Nonetheless, at the same point in the process, the largest components become so large that it is likely that a new edge will connect two of them. Thus, fairly quickly, all the largest components of $\mathbb{G}(n,k)$ merge into one giant component, much larger than any of the remaining ones [17]. It is noteworthy that this process represents the mean-field case of percolation [19].

We numerically compute $\mathcal{S}(k)$, the geometric entropy in Eq. (11) vs k for a fixed n, in order to investigate its sensitivity to the appearance of the giant component during the evolution of the RG model $\mathbb{G}(n,k)$. It is noteworthy that a Gibbs entropy has been defined for the statistical set of random graphs [20] as

$$S := \ln \frac{1}{n!}\binom{\binom{n}{2}}{k}. \tag{15}$$

Following up, the research lineof characterizing other classes of RGs, like scale-free and fixed-degree sequences, has been pursued [21]. However, the entropy (15) (as a function of k/n) is not able to detect the Erdos-Rényi phase transition.

In practice we have considered four numbers of vertices: $n = 25$, 50, 100, and 200. Note that the magnitude of n is not important; what matters is the n dependence, the so-called finite-size scaling, of the relevant observables. The magnitude of n simply determines the dimension of the manifold $\widetilde{\mathcal{M}}$. For any fixed n we have considered the number of links, k, to be $k = 0,1,\ldots,n(n-1)/2$. Then, for any pair (n,k) we have randomly generated a set of k entries (i,j), with $i < j$, of the nonvanishing adjacency matrix elements A_{ij}.

In this way, since the covariance matrix C is functionally assigned, we have gotten the $\psi_C(A)$ of Eq. (8) and, finally, the metric $\widetilde{g}$ of Eq. (10). Next, having determined $\widetilde{\mathcal{M}} = (\widetilde{\Theta},\widetilde{g})$, we computed the volume $\mathcal{V}(A)$ in Eq. (13) and the entropy $\mathcal{S}$ in Eq. (11). The volume regularization is performed in two steps. The first step is to restrict the manifold support $\widetilde{\Theta} \subset \mathbb{R}^n$ to a hypercube. Inside it we generate a Markov chain to perform a Monte Carlo estimation of the average $\langle\sqrt{\det \widetilde{g}}\rangle = \int \sqrt{\det \widetilde{g}}\, d\theta^1 \wedge \ldots \wedge d\theta^n / \int d\theta^1 \wedge \ldots \wedge d\theta^n$. The number of considered random configurations ranges between 10^4 and 10^6. As the second step of the regularization we have excluded those points where the value of $\sqrt{\det \widetilde{g}}$ exceeds 10^{308} (the numerical overflow limit of the computers used). Then, for any given pair (n,k) this computational procedure is repeated 10^3 times, each time considering a different randomly generated realization of the adjacency matrix A. Thus, the final values of the entropy $\mathcal{S}$ are obtained as averages over 10^3 different manifolds $\widetilde{\mathcal{M}}$, namely,

$$\widetilde{\mathcal{S}}(k) := \frac{1}{n}\langle(\mathcal{S}(k) - \mathcal{S}(0))\rangle = \frac{1}{n}\left\langle \ln \frac{\int \sqrt{\det \widetilde{g}}\, d\theta^1 \wedge \ldots \wedge d\theta^n}{\int \sqrt{\det g}\, d\theta^1 \wedge \ldots \wedge d\theta^n}\right\rangle, \tag{16}$$

where g is the metric corresponding to the null adjacency matrix.

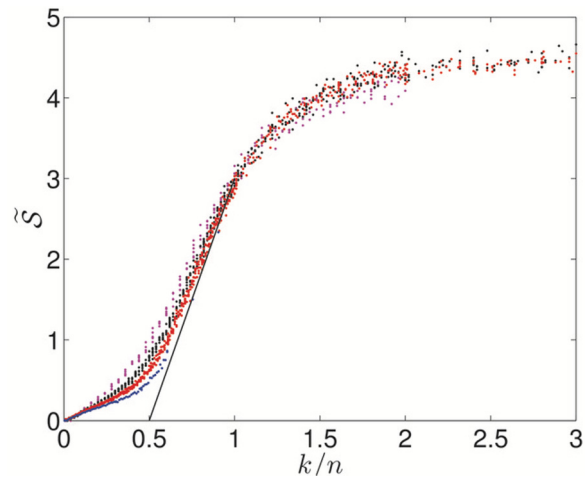


FIG. 1. Values of the entropy $\widetilde{\mathcal{S}}(k/n)$ of $\mathbb{G}(25,k)$ (magenta points), $\mathbb{G}(50,k)$ (black points), $\mathbb{G}(100,k)$ (red points), and $\mathbb{G}(200,k)$ (blue points) networks as a function of the number k of randomly chosen links of weights equal to $r = 0.2$. The solid black line is a guide for the eye coming from a linear fitting of a linear-logarithmic presentation of the data.

In Fig. 1 we report the behavior of $\widetilde{\mathcal{S}}(k/n)$ in the case of equal weights $A_{ij} = r$ for all k nonvanishing links. This is what, in the context of statistical mechanics, is known as a *collapse plot* of the results obtained at different n values. It shows a typical phenomenon arising in numerical investigations of second-order phase transitions: likewise, the finite-size effect observed for the order parameter, which asymptotically would be a sharp bifurcation, is rounded at finite n. However, the larger n is, the more pronounced the "knee" of $\widetilde{\mathcal{S}}(k/n)$ becomes. This is in agreement with an n-asymptotic bifurcation at $k/n = 0.5$ (solid black line), where the Erdös-Rényi phase transition takes place.

At present, this beautiful and unambiguous result (presented also in [13]) lends credit to our proposed measure of network complexity. To reinforce it we think a stability check is in order. Then, in Fig. 2 we report the outcomes for $\mathbb{G}(50,k)$ having chosen at random the values of the nonvanishing entries A_{ij} of the adjacency matrix, that is, $A_{ij} = 0.2 + \omega$, where ω is a random variable of zero mean and variance equal to 0.1. Of course negative values of the A_{ij} are excluded. Comparison with the results obtained with $A_{ij} = 0.2$ confirms the robustness of the entropy defined in Eq. (11).

V. BEYOND RANDOM GRAPHS

Here we go beyond the RG model and apply the proposed complexity measure defined in Eq. (11) to complex networks

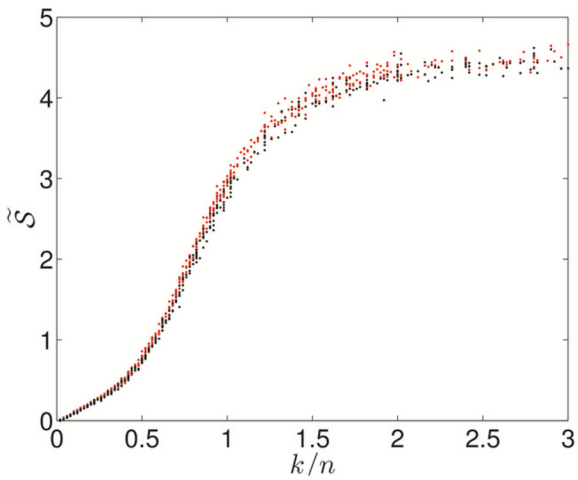


FIG. 2. Values of the entropy $\widetilde{S}(k/n)$ of $\mathbb{G}(50,k)$ networks as a function of the number k of randomly chosen links of weight equal to $r = 0.2$ (red points), $r = 0.2 + \omega$, with ω a random Gaussian of zero mean and 0.1 variance (black points).

with the aim of comparing our results with others already known.

A. Small exponential random graphs

The general idea that a system is complex when it does not coincide with the "summation" of all its parts has been formalized in [22] within the framework of information geometry. With this approach, a hierarchy of exponential families is provided, which is widely studied in information geometry [23], modeling networks of progressively increasing order of the interactions between their parts. The model, known as exponential random graphs (ERGs), is the distribution over a specified set of graphs that maximizes the Gibbs entropy under some suitable constraints; more precisely, let us suppose to be given a collection $\{\xi_i\}_{i=1,\ldots,r}$ of graph observables (such as nodes and edges) measured in empirical observations of some real-world network or of any other network as well. In addition, assume that we have an estimate $\langle\xi_i\rangle$ of the expectation value of each observable. Consider now a graph G and $P(G)$, the probability of that graph within a given ensemble $\mathcal{G}$. Then we wonder what is the best choice for $P(G)$ so that the expectation value of each of our graph observables $\{\xi_i\}$ within that distribution is equal to its observed value. The answer is obtained by maximizing Gibbs' entropy

$$S = -\sum_{G\in\mathcal{G}} P(G)\ln P(G)$$

under the constraints $\sum_{G\in\mathcal{G}} P(G)\xi_i(G) = \langle\xi_i\rangle$ and $\sum_{G\in\mathcal{G}} P(G) = 1$, where $\xi_i(G)$ is the value of ξ_i for the graph G. This procedure leads to

$$P(G) = \frac{e^{-\mathcal{H}(G)}}{Z}, \tag{17}$$

where $\mathcal{H}(G) = \sum_i \lambda_i\xi_i(G)$ is the graph Hamiltonian and $Z = \sum_{G\in\mathcal{G}} e^{-\mathcal{H}(G)}$ is the partition function. Here, λ_is are Lagrangian multipliers. This equation defines the exponential random graph model [24].

This model has been employed to quantify the degree of interaction of all the parts of a given system [25]. Still in Ref. [25], simple ERGs are considered in order to describe "typical" graphs, i.e., the graphs that are most probable in the ensemble defined by this model and that correspond to the lowest "energy" characterizing the model. In particular, in Ref. [25] the authors consider the simple ERG model with six nodes, where only the interactions between triangles and 3-chains are taken into account, that is, only a subset of the family of all graphs with six nodes is considered. Then the convex hull of all the possible expectation values of the probabilities of the triangles and of the 3-chains is derived. Those graphs that correspond to the minimal energy are found to lie on the lower boundary of the mentioned convex hull.

The geometric entropy proposed in the present work has been computed to provide a "pointwise" description of the complexity of single members of a given family of graphs. The outcomes of these computations allow us to rank the members of a given family of graphs according to their degree of complexity, of course on the basis of the proposed method of measuring it. The results are summarized in Table I. They suggest that going up along the lower boundary of the previously mentioned convex hull (which moves along a line representing a given family), the degree of complexity increases.

TABLE I. The value of $\widetilde{S}$ for exponential random graphs (ERGs) corresponding to the minimal energy.

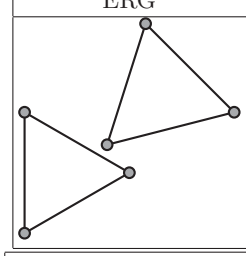
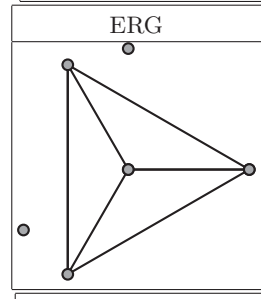
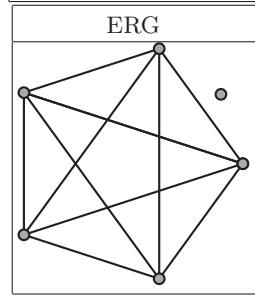
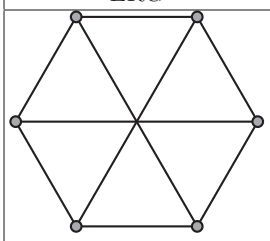
ERG	$\widetilde{S}$
	0.568
	1.006
	1.303

TABLE II. The value of $\widetilde{S}$ for an exponential random graph (ERG) corresponding to the maximal energy.

ERG	$\widetilde{S}$
	2.332

Moreover, the result in Table II shows that our entropic measure is capable of distinguishing among different families of networks. In fact, while the graphs in Table I represent typical graphs on the minimal energy boundary, the graph in Table II is a typical graph on the maximal energy boundary. Note that the results in Table I indicate that the network with two triangles (2-simplices) is less complex than the network with one tetrahedron (3-simplex) plus two points (0-simplices), which is less complex than the network with one 4-simplex plus one point (0-simplex); in other words, network complexity is nontrivially influenced by network topology (homology). The first account of this fact is given in Ref. [12].

B. Configuration model

Real networks usually differ from Erdös-Rényi random graphs in their degree distribution [1]. Given an undirected network with adjacency matrix $A = (A_{ij})$, the degree of a node i is just the sum of the ith row's entries, $d_i := \sum_i A_{ij}$. It represents the number of connections that node i has. The degree distribution $P(d)$ of a network is then defined to be the fraction of nodes in a network with degree d. The degree distribution clearly captures information on the structure of the network. For example, in binomial Erdös-Rényi random graphs, usually indicated as simple (not complex) networks, one finds that most nodes in the network have similar degrees; this model, in which each of the n nodes is connected with independent probability p, has a binomial distribution of degrees d, namely, $P(d) = \binom{n-1}{d} p^d (1-p)^{n-1-d}$ [17]. However, real-world networks usually have very different degree distributions. That is, most of the nodes have a relatively small degree (low connectivity), while a few of them have a very large degree (i.e., are connected to many other nodes). These large-degree nodes are often referred to as hubs [1].

The first step toward testing the effectiveness of our geometric entropy in quantifying the complexity of real networks is to compare networks where each node has the same given degree d to networks containing hubs. When each node of a network has the same degree d, the network is called a d-regular graph [17]. One of the most widely used methods to generate these special networks is the *configuration model* [26]. This is specified in terms of a sequence of degrees; for a network of n nodes we have a desired degree sequence $(d_1, \ldots, d_n)$, which specifies the degree d_i of each node i, for $i = 1, \ldots, n$.

TABLE III. The value of $\widetilde{S}$ for random d-regular graphs.

n	d	$\widetilde{S}$
50	2	1.0265
50	6	3.8498

The average vertex degree $\langle d_i \rangle$ is the ratio between the total number of links in a given network and the number of nodes. It represents the first level of characterization of the topological complexity [27]. We consider it a benchmark to strengthen the validation of our proposal. We have numerically computed the entropy $\widetilde{S}$ given by Eq. (16) for networks of number of nodes $n = 50$, constructed as *random* d-regular graphs of two values of d, that is, $d = 2$ and $d = 6$. A random d-regular graph is a random graph with a uniform distribution over all d-regular graphs. The computed value of the geometric entropy $\widetilde{S}$ increases with d, as reported in Table III. This result is very good because it is in agreement with the obvious fact that the larger d, the more complex the network.

The next step toward real networks consists of considering random graphs, again with a number of nodes $n = 50$, and with a given sequence $d = (d_1, d_2, \ldots, d_n)$ of nonincreasing degrees: $d_1 \geqslant d_2 \geqslant d_3 \geqslant \ldots \geqslant d_n$. In so doing we proceed with the validation of the geometric entropy $\widetilde{S}$ in Eq. (16) by considering networks with one or more hubs. In the previous notation, a network with hubs is identified by one or more values in the sequence $d = (d_1, d_2, \ldots, d_n)$ which are larger than the others. In Table IV, the numerically obtained values of $\widetilde{S}$ are reported for networks with hubs of degree $d = 8$, $d = 10$, and $d = 14$, respectively, while the other nodes have degree $d = 2$. It is found that the complexity of a network increases with the number of hubs in it. Moreover, Table IV shows that a network with degrees $(8, 2, \ldots, 2)$ is less complex than a network with degrees $(10, 2, \ldots, 2)$, which is less complex than a network with degrees $(14, 2, \ldots, 2)$; in addition, the network with degrees $(8, 8, 2, \ldots, 2)$ is less complex than the network with degrees $(10, 10, 2, \ldots, 2)$. Again, this confirms that the geometric entropy $\widetilde{S}$ in Eq. (16) leads to an overall consistent scenario.

In a way, homogeneity in the interaction structures entails topological equivalence of almost all the nodes; on the contrary, it was found that most real networks show degree distribution $P(d) = A\, d^{-\gamma}$, where A is a positive *real* constant and the exponent γ varies in the range $2 \leqslant \gamma \leqslant 3$ [28]. These networks are called *scale-free networks*, because power laws have the property of self-similarity at all scales [29,30]. Before tackling the study of scale-free networks, we now deeply focus on the issue of heterogeneity. In particular, we consider networks of the same average degree but of different degrees of heterogeneity. Understanding a regular graph as the less heterogeneous network, we see heterogeneity as a measure of how far from $P(d_0) = 1$, $P(d) = 1 (d_0 \neq d)$ a network is. This entails that the most heterogeneous network corresponds to $P(1) = P(2) = \ldots P(n-1)$ [31]. In particular, we consider networks with $n = 50$ nodes and the same average degree $d = \frac{53}{50}$; moreover, we assume the following degree sequences: $(8, 2, \ldots, 2)$, $(7, 3, 2, \ldots, 2)$, $(6, 4, 2, \ldots, 2)$, $(5, 5, 2, \ldots, 2)$, and $(6, 3, 3, 2 \ldots, 2)$. Next, the degree of heterogeneity of such

TABLE IV. The value of $\widetilde{S}$ for networks with hubs.

d	$\widetilde{S}$	d	$\widetilde{S}$	d	$\widetilde{S}$
(2, . . . ,2)	1.0265	(2, . . . ,2)	1.0265	(2, . . . ,2)	1.0265
(8,2, . . . ,2)	1.6140	(10,2, . . . ,2)	1.9156	(14,2, . . . ,2)	2.7631
(8,8,2, . . . ,2)	2.1263	(10,10,2, . . . ,2)	2.3878		
(8,8,8,2, . . . ,2)	2.2120				
(8,8,8,8,2, . . . ,2)	2.8473				
(8,8,8,8,8,2, . . . ,2)	3.2298				

networks is measured through the definition given in [32], that is,

$$h = \sum_{i,j\in E} \left(\frac{1}{\sqrt{d_i}} - \frac{1}{\sqrt{d_j}} \right)^2, \tag{18}$$

where E is the set of edges of the network. This heterogeneity measure vanishes for regular graphs, while, as the difference in the degrees of adjacent nodes increases, it also increases. Another advantage of h is that it can be expressed in terms of the Laplacian matrix [33], which is widely employed in the study of network complexity. Therefore, to perform a benchmarking we compute the geometric entropy $\widetilde{S}$ in (16) and compare it to the heterogeneity h in (18) for the above-mentioned networks. In Table V we can see that the ordering of these networks according to their decreasing degree of heterogeneity as measured by $\widetilde{S}$ is the same as the ordering produced by the h measure of heterogeneity.

C. Scale-free networks

Many real networks show a power-law degree distribution and very often they are modelled by scale-free networks; for instance, this is the case for some social networks [34] and biological networks [35], among others. These graphs with a power-law degree distribution can be obtained as special cases of RGs with a given degree distribution, as discussed in the previous section.

Power laws play a particular role in statistical physics, especially because of their connections to fractals [36] and phase transitions [37]. In this section, we refer to scale-free networks as the class of networks with any power law of the degree distribution and focus our investigation on the issue of phase transitions.

Beyond the two well-known RG models (uniform and Poissonian), another kind has been considered, which is suitable for describing power-law-degree families. This is described by two parameters, α and γ, which define the size and the density of a network; hence, given the number of nodes N with degree d, this model, denoted $G_{\alpha,\gamma}$ [28], assigns a uniform probability to all graphs with $N = e^{\alpha} d^{-\gamma}$.

The connectivity properties of the model $G_{\alpha,\gamma}$, as a function of the power γ, have been shown to hold almost surely for sufficiently large graphs. Likewise, as we have previously seen, the transitional properties of the uniform random graph model $\mathbb{G}(n,k)$ hold almost surely asymptotically in n. For the model $G_{\alpha,\gamma}$ a critical value of γ exists for the emergence of a giant component. It has been proved that for values of γ smaller than $\gamma_c = 3.4785$ ($\gamma < \gamma_c$), a giant component always exists. On the other hand, when $\gamma > \gamma_c$ the graphs $G_{\alpha,\gamma}$ almost surely have no giant component [38].

In order to check whether $\widetilde{S}$ in Eq. (16) detects this property of scale-free networks, we have proceeded as follows. We considered networks of $n = 250$ nodes [39] for which, without loss of generality, we set $\alpha = 0$. For each value of γ, we selected 10 different realizations of the networks, each realization having the same value of k/n larger than the threshold value (0.5) for the appearance of a giant component. Actually, because of the practical difficulty of getting different realizations of a scale-free network with exactly the same value of k/n at different γ values, we accepted a spread of values in the range 0.7–0.85.

In Fig. 3 we report the behavior of the geometric entropy $\widetilde{S}$ of the power-law random graphs $G_{\alpha,\gamma}(n,k)$ when $\alpha = 0$,

TABLE V. The value of $\widetilde{S}$ for networks with the same average degree and different degrees of heterogeneity.

d	$\widetilde{S}$	h
(8,2, . . . ,2)	1.6140	1
(7,3,2, . . . ,2)	1.3070	0.8088
(6,4,2, . . . ,2)	1.0941	0.7074
(5,5,2, . . . ,2)	1.0924	0.6754
(6,3,3,2, . . . ,2)	1.0357	0.4970

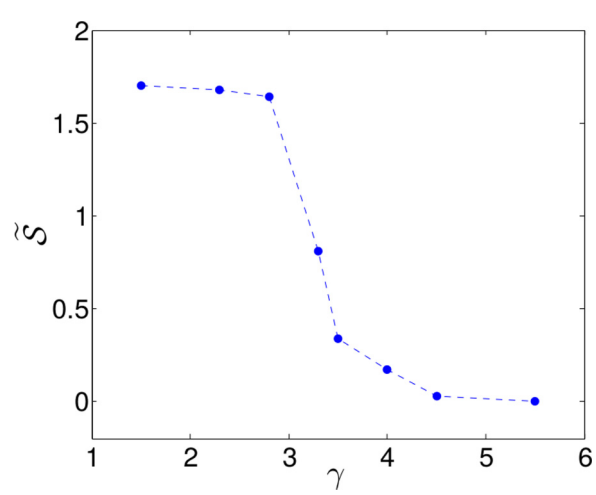


FIG. 3. Values of the entropy $\widetilde{S}$ of power-law $G_{0,\gamma}(250,k)$ networks as a function of the exponent γ; k values varied—according to the realization of the RG and, independently, of γ—approximately in the range 180–210.

$n = 250$, $r = 0.2$ (the weights of the links), and the exponent γ is in the range $1.5 < \gamma < 5.5$. The pattern of $\widetilde{S}(\gamma)$ displays the typical phenomenon found in numerical investigations of second-order phase transitions: what asymptotically would be a sharp bifurcation is rounded here because of a finite-size effect. This is in excellent agreement with the n-asymptotic bifurcation at $\gamma = \gamma_c = 3.478\,75$ predicted by the Molloy Reed criterion for the emergence of a giant component in the power-law RGs $G_{\alpha,\gamma}$ [38].

D. Real networks

Real-world graphs are usually more complex than random graphs [29,30]. In contrast to Erdös-Rényi random graphs real-world graphs have some typical features, such as a power-law degree distribution, correlation of node degrees [40], modularity, and structures [41]. Many complexity measures have been proposed to describe real networks capturing one or another of their typical features. Given undirected graphs $\mathbb{G}(n,k)$ with n nodes and k edges, some measures maximize the complexity of graphs with nearly the complete number of edges; other measures indicate as highly complex graphs real networks with modular structures at different levels, which are expected only for a medium number of edges [42]. The latter complexity measures have been basically defined by the specification of which subnetworks are considered and when two subnetworks are different. They are outlined through (i) the one-edge-deleted subnetwork with respect to the number of spanning trees, (ii) the one-edge-deleted subnetwork with respect to different spectra of the Laplacian matrix (degree matrix minus adjacency matrix) and the signless Laplacian matrix (degree matrix plus adjacency matrix), and (iii) the two-edges-deleted subgraph complexity with respect to different spectra of the Laplacian and signless Laplacian matrices. However, only the first one, denoted $C_{1e,\mathrm{ST}}$ in [42], has been applied to 33 real networks, as the other measures required too great a computational effort. Thus, the measure $C_{1e,\mathrm{ST}}$ has been compared to some product measures and some entropic measures. Entropic measures quantify the diversity of different topological features and our measure $\widetilde{S}$ of complexity may be ascribed to this class. Here we compare our measure to $C_{1e,\mathrm{ST}}$ and discuss the two entropic measures employed in [42]: the spanning tree sensitivity (STS) and the spanning tree sensitivity differences (STSDs). These are based on the idea that complex graphs have very diverse edge sensitivities with respect to the removal of different edges, while in very simple graphs all edges play the same role and the graph has the same edge sensitivity with respect to the removal of different edges.

In Table VI the outcomes of the numerical computation of $\widetilde{S}$ are listed for some of the networks considered in [42], for which also the corresponding $C_{1e,\mathrm{ST}}$ values are listed. These networks are the coauthorship network of scientists working on network theory (Net Science), the coappearance network of characters in the novel *Les Miserables*, the network (Dolphins) of frequent associations between dolphins, and the adjacency network (Word Net) of common adjectives and nouns in the novel *David Copperfield*. Though our geometric entropy has proven its own meaningfulness above, it is very interesting to note that the ordering of these networks according to their complexity, which is established by $\widetilde{S}$, is the same as the ordering produced by the $C_{1e,\,\mathrm{ST}}$ measure of complexity. It is worth mentioning that the network Les Miserables is the only weighted network among those considered here, and its adjacency matrix has a relatively small number of high-weight edges. The corresponding entropy value, $\widetilde{S} = 1.670$, reported in Table VI, increases to $\widetilde{S} = 2.644$ upon setting all the weights of the edges of the network equal to 1. Loosely speaking, this amounts to increasing the effective network connectivity, and this is correctly detected by a corresponding increase in $\widetilde{S}$. Finally, let us note that the relative variations of $\widetilde{S}$ are much larger than those of $C_{1e,\mathrm{ST}}$ (and of other parameters defined in Ref. [42] for the same networks). This means that $\widetilde{S}$ has a greater "resolving power" in comparative measurement of the complexity of different networks.

TABLE VI. The value of $\widetilde{S}$ for real networks: n is the number of nodes; k is the number of links [45].

Network	n	k	$\widetilde{S}$	$C_{1e,\mathrm{ST}}$
Net Science	413	948	1.376	0.01
Les Miserables	77	254	1.670	0.10
Dolphins	62	159	2.852	0.14
Word Net	112	425	3.010	0.15

The two entropic measures, STS and STSD, are assessed to be strongly correlated—in terms of Pearson correlations—with the $C_{1e,\,\mathrm{ST}}$ measure of complexity. All these measures are able to discriminate different networks with the same number of nodes and of links. However, if, on the one hand, the STS and STSD perform better in characterizing the complexity of nontree graphs, on the other hand, $C_{1e,\,\mathrm{ST}}$ best discriminates different complexities of different trees. Furthermore, the STS and STSD are still slightly different. Indeed, one of them (STS) identifies more homogeneity with more complexity; on the contrary, the STSD quantifies as more complex a network with more heterogeneity [42].

Beyond the properties of assigning the highest complexity values only to networks with a medium number of edges and discriminating, at best, different networks with the same number of nodes and links, a measure of complexity should assign higher values to real networks than to their randomized counterparts [1]. In Table VII the outcomes of the numerical computation of $\widetilde{S}$ are reported for the above-mentioned *real* networks, which are now randomized; that is, for the same number n of nodes and number k of links of each network, a random graph is generated. Then, the geometric entropy $\widetilde{S}$ is computed for each of these RGs. The remarkable result is that,

TABLE VII. The value of $\widetilde{S}$ for real networks and their randomized versions: n is the number of nodes; k is the number of links.

Network	n	k	$\widetilde{S}$	$\widetilde{S}$ of RG
Net Science	413	948	1.376	0.4454
Les Miserables	77	254	1.670	1.6655
Dolphins	62	159	2.852	1.7246
Word Net	112	425	3.010	1.4537

for each network, its corresponding RG has a lower degree of complexity according to $\widetilde{S}$.

A comparison of our measure of complexity $\widetilde{S}$ with the two above-mentioned entropic measures shows that it orders the real networks in Table VI in the same way as does the STS. On the other hand, it does not match the STSD as far as the Dolphins and the Word Net networks are concerned. However, our measure assesses heterogeneous networks as more complex than homogeneous ones, exactly as the STSD measure of complexity. Furthermore, while RGs always have higher STS values than the corresponding real networks, the STSD measure sometimes assesses real networks as more complex than their corresponding randomized versions [42].

Summarizing, we have compared our measure of complexity $\widetilde{S}$ to the $C_{1e,\mathrm{ST}}$ measure, which is used to quantify the complexity of a given network according to the many different subgraphs it may contain. The two measures order the real networks listed in Table VI in the same way. Furthermore, this ordering of these real networks coincides with the STS ordering, whereas it does not match the STSD values as far as the Dolphins and Word Net networks are concerned. However, the fact that the randomized versions of all the real networks considered in [42] correspond to higher STS values with respect to the original networks constitutes a difficulty with this measure. At variance, the absence of this difficulty when our geometric entropy is used, at least for the considered cases, lends further support to the validity and consistency of the measure $\widetilde{S}$. All in all, in spite of the minor discrepancies between the STSD and $\widetilde{S}$ in the ordering of two real networks, both measures attribute a higher degree of complexity to networks with a higher degree of heterogeneity. Once again, all these factors confirm that the complexity measure $\widetilde{S}$ in Eq. (16) leads to an overall consistent scenario.

VI. OUTLOOK ON FUTURE DEVELOPMENTS

The geometric framework so far put forward paves the way to interesting developments. A relevant generalization made possible by the Riemannian-geometric framework consists in considering the time evolution of a network. In order to do this, one should drop the simplifying assumptions of the present work by adding to the θ^i's of the diagonal covariance matrix C also the entries $\sigma^{ij} := A_{ij}$ in the adjacency matrix A as local coordinates of the statistical manifold $\widetilde{\mathcal{M}}$ in Eqs. (9) and (10). In this way the dimension of $\widetilde{\mathcal{M}}$ increases from n to $n(n+1)/2$.

Denoting by $\zeta^i = (\psi_C(A))_{lm}$ the $n(n+1)/2$ local coordinates of $\widetilde{\mathcal{M}}$, where $i = \sum_{r=0}^{l-2}(n-r) + m - l + 1$, there is a natural way of tackling the dynamical evolution of the network associated with $(\widetilde{\mathcal{M}}, \widetilde{g})$, that is, through the geodesic flow given by the following set of equations

$$\frac{d^2\zeta^i}{ds^2} + \sum_{jk} \Gamma^i_{jk} \frac{d\zeta^j}{ds}\frac{d\zeta^k}{ds} = 0, \quad i,j,k = 1,\ldots,n(n+1)/2, \tag{19}$$

where the Γ^i_{jk} are the standard Christoffel connection coefficients

$$\Gamma^i_{jk} = \frac{1}{2}\sum_l \widetilde{g}^{il}(\partial_j \widetilde{g}_{lk} + \partial_k \widetilde{g}_{jl} - \partial_l \widetilde{g}_{jk}). \tag{20}$$

The physical time parametrization of the arc length s is derived by means of the metric tensor as

$$\frac{ds}{dt} = \sqrt{\sum_{ij} \widetilde{g}_{ij}\dot{\zeta}^i\dot{\zeta}^j}, \tag{21}$$

where the $\dot{\zeta}^i$ are the variation rates of the local coordinates expressed with respect to the physical time t. Let us remark that the dynamical evolution described by Eq. (19) encompasses also the time evolution of the weights of the links of a network, including their appearance and disappearance, thus *a priori* allowing the computation of the time variation of its complexity $\widetilde{S}(t)$. The fitting of empirical data concerning the true evolution of a real network by means of the model dynamics given by Eqs. (19) and (21) could allow us to obtain relevant information about the laws that drive network evolution (conservation, extremalization, optimization of some quantities, and so on).

Another prospective and remarkable application of the differential geometrical approach put forward in the present work, and notably related to the dynamical equations, (19), concerns the study of the stability properties of a network. In fact, by setting $\zeta^i(t) \to \zeta^i(t) + \varphi^i(t)$, where $\varphi^i(t)$ are small functional perturbations, after substitution into Eq. (19) and use of (21), one can work out the *tangent dynamics equations* in the form of a system of first-order linear differential equations [11]:

$$\frac{d\varphi^i}{dt} = \Phi^i(\{\zeta^j\}). \tag{22}$$

These equations, numerically integrating Eqs. (19) and (21), are the natural tool for investigating the stability of either stationary or nonstationary states of a network, for example—for a stationary state—investigating a network stability or instability under the addition or deletion of one or more links, and so on.

VII. CONCLUSION

Summarizing, the present work contributes to the fascinating subject of quantifying the degree of complexity of systems that are commonly defined as "complex." A large number of such definitions are already available. Perhaps the history begins with Kolmogorov's definition of algorithmic complexity [43,44], which, in spite of its theoretical beauty, is hardly applicable in practice. Since then the many definitions put forward have been adapted to the specific systems and problems tackled. However, the number of categories into which all these definitions can be gathered is rather limited. Of course, borrowing the concept of physical entropy from statistical mechanics is the most inspiring and intriguing way to proceed. In fact, in physics, entropy is just a measure of disorder and, conversely, negentropy—as defined by L. Brillouin long ago—is a measure of the degree of order in a system and is also the physical equivalent of Shannon's information entropy—again, as shown by L. Brillouin [5]. Whence the vast literature addressing the quantification of complexity on the basis of Shannon's information entropy, which, on the other hand, was inspired by Boltzmann's entropy of kinetic theory. However, what was still missing was a general definition of an entropic measure of complexity

accounting both for the structure of any given network and for its statistical complexity, that is, for the complexity of the probability distributions of the entities constituting the network. The new definition put forward in the present work embraces both these aspects. It is still inspired by statistical mechanics, however, instead of being modeled on the Boltzmann entropy, it is modeled on the microcanonical ensemble definition of entropy, with the phase-space volume being replaced by the volume of a "state manifold" (that is, a Riemannian manifold whose points correspond to all the possible states of a given network). The state manifold is defined through a suitable definition of a metric which is partly borrowed from so-called information geometry and partly an original proposal put forward in the present work. The result is a constructive way of associating a differentiable and handy mathematical object to any simple undirected and weighted graph or network. Another novelty consists in having directly tested, by means of numerical computations, the validity and effectiveness of the proposed entropic-geometric measure of complexity. In order to do this we needed, so to speak, a paradigmatic example of a major change in complexity. A possible natural choice is suggested by the observation that phase transitions are the most impressive examples in nature of emergent phenomena—theoretically well understood—associated with a sharp qualitative and quantitative change in the complexity of a physical system when a control parameter exceeds a critical value. This kind of phenomenon exists also in complex networks. In fact, random graphs undergo a well-known phase transition as proved by the Erdös-Rényi theorem: a paradigmatic—and at present unique—example of an analytically known major variation of the degree of complexity of a network. This kind of check is unprecedented and very successful; in fact, the entropic-geometric measure of complexity proposed here displays both a pattern and a size dependence of that pattern which are typically found for the order parameter of a second-order phase transition in physics. Then, since RGs undergoing the Erdös-Rényi transition are not considered genuinely complex networks, the proposed entropic-geometric measure of complexity has been applied to small exponential random graphs, to different versions of random d-regular graphs with and without hubs generated according to the configuration model. Moreover, the ability to detect the transitions predicted by the Molloy Reed criterion in power-law RGs is an important confirmation of the consistency of the geometric entropy measure of complexity proposed here. Finally, this has been applied to some real networks and compared to three other measures of complexity: one describing the modularity structures; one, homogeneity; and one, the heterogeneity of real networks. The outcomes of these applications compose a consistent scenario validating the meaningfulness and effectiveness of the proposed measure of complexity.

Finally, the differential-geometric framework put forward opens some fascinating perspectives of application to the study of the time evolution of complex systems.

ACKNOWLEDGMENTS

We are indebted to R. Quax for providing us with data on power-law random graphs. We also thank M. Rasetti for useful discussions. This work was supported by the Seventh Framework Programme for Research of the European Commission under FET-Proactive Grant TOPDRIM (FP7-ICT-318121).

[1] R. Albert and A.-L. Barabasi, Statistical mechanics of complex networks, Rev. Mod. Phys. **74**, 47 (2002).

[2] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Critical phenomena in complex networks, Rev. Mod. Phys. **80**, 1275 (2008).

[3] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[4] K. Anand, G. Bianconi, and S. Severini, Shannon and von Neumann entropy of random networks with heterogeneous expected degree, Phys. Rev. E **83**, 036109 (2011), and references quoted therein.

[5] L. Brillouin, *Science and Information Theory* (Academic Press, New York, 1962); reprinted by Dover, Mineola, NY (2004).

[6] M. Boguñá and R. Pastor-Satorras, Class of correlated random networks with hidden variables, Phys. Rev. E **68**, 036112 (2003).

[7] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz, Scale-Free Networks from Varying Vertex Intrinsic Fitness, Phys. Rev. Lett. **89**, 258702 (2002).

[8] C. Cafaro, Works on an information geometrodynamical approach to chaos, Chaos Solitons Fractals **41**, 886 (2009).

[9] S. Keshav, *Mathematical Foundations of Computer Networking* (Addison Wesley, Reading, MA, 2012).

[10] S. Amari and H. Nagaoka, *Methods of Information Geometry* (Oxford University Press, New York, 2000).

[11] M. Pettini, *Geometry and Topology in Hamiltonian Dynamics and Statistical Mechanics* (Springer-Verlag, Berlin, 2007).

[12] D. Felice, S. Mancini, and M. Pettini, Quantifying networks complexity from information geometry viewpoint, J. Math. Phys. **55**, 043505 (2014).

[13] R. Franzosi, D. Felice, S. Mancini, and M. Pettini, A geometric entropy detecting the Erdös-Rènyi phase transition, Europhys. Lett. **111**, 20001 (2015).

[14] J. D. Noh and H. Rieger, Random Walks on Complex Networks, Phys. Rev. Lett. **92**, 118701 (2004).

[15] S. Blachére, P. Haïssinsky, and P. Mathieu, Asymptotic entropy and Green speed for random walks on countable groups, Ann. Prob. **36**, 1134 (2008).

[16] P. Erdös and A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci., Ser. A **5**, 17 (1960).

[17] S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs* (John Wiley and Sons, New York, 2000).

[18] G. Leibbrandt, Introduction to the technique of dimensional regularization, Rev. Mod. Phys. **47**, 849 (1975).

[19] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Network Robustness and Fragility: Percolation on Random Graphs, Phys. Rev. Lett. **85**, 5468 (2000).

[20] L. Bogacz, Z. Burda, and B. Wacław, Homogeneous complex networks, Phys. A: Stat. Mech. Appl. **366**, 587 (2006).

[21] G. Bianconi, The entropy of randomized network ensembles, Europhys. Lett. **81**, 28005 (2008).

[22] N. Ay, E. Olbrich, N. Bertschinger, and J. Jost, A geometric approach to complexity, Chaos **21**, 037103 (2011).

[23] S. I. Amari, Information theory, IEEE Trans. Inf. Theor. **47**, 1701 (2001).
[24] J. Park and M. E. J. Newman, Statistical mechanics of networks, Phys. Rev. E **70**, 066117 (2004).
[25] E. Olbrich, T. Khale, N. Bertschinger, N. Ay, and J. Jost, Quantifying structure in networks, Eur. Phys. J. B **77**, 239 (2010).
[26] E. A. Bender and E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences, J. Combin. Theor. Ser. A **24**, 296 (1978).
[27] D. Bonchev and G. A. Buck, Quantitative measure of network complexity, in *Complexity in Chemistry, Biology and Ecology*, edited by D. Bonchev and D. H. Rouvray (Springer, New York, 2005).
[28] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. **424**, 175 (2006).
[29] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature **393**, 440 (1998).
[30] A.-L. Barabasi and R. Albert, Emergence of scaling in random networks, Science **286**, 509 (1999).
[31] J. Wu, Y.-J. Tan, H.-Z. Deng, and D.-Z. Zhu, A new measure of heterogeneity of complex networks based on degree sequence, in *Unifying Themes in Complex Systems* (Springer-Verlag, Berlin, 2008), pp. 66–73.
[32] E. Estrada, Quantifying network heterogeneity, Phys. Rev. E **82**, 066102 (2010).
[33] D. S. Bernstein, *Matrix Mathematics. Theory, Facts and Formulas* (Princeton University Press, Princeton, NJ, 2009).
[34] R. Quax, A. Apolloni, and P. M. A. Sloot, The diminishing role of hubs in dynamical processes on complex networks, J. Roy. Soc. Interf. **10**, 20130568 (2013).
[35] D. Breitkreutza, L. Hlatky, E. Rietman, and J. A. Tuszynski, Molecular signaling network complexity is correlated with cancer patient survivability, Proc. Natl. Acad. Sci. USA **109**, 9209 (2012).
[36] K. J. Falconer, *Fractal Geometry: Mathematical Foundations and Applications* (Wiley, New York, 1990).
[37] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, New York, 1971).
[38] W. Aiello, F. Chung, and L. Lu, A random graphs model for power law graphs, Exp. Math. **10**, 53 (2001).
[39] Data for power-law random graphs were provided by R. Quax; see Ref. [34] for details.
[40] M. E. J. Newman, Assortative Mixing in Networks, Phys. Rev. Lett. **89**, 208701 (2002).
[41] M. E. J. Newman, Modularity and community structure in networks, Proc. Natl. Acad. Sci. USA **103**, 8577 (2006).
[42] J. Kim and T. Wilhelm, What is a complex graph? Physica A **387**, 2637 (2008).
[43] A. N. Kolmogorov, Three approaches to the quantitative definition of information, Probl. Inf. Transm. (USSR) **1**, 4 (1965).
[44] A. N. Kolmogorov, Logical basis for information theory and probability theory, IEEE Trans. Inf. Theor. **IT14**, 662 (1968).
[45] Data taken from http://www-personal.umich.edu/∼mejn/netdata/