A geometric entropy measuring networks complexity

Roberto Franzosi
QSTAR and INO-CNR, largo Enrico Fermi 2, I-50125 Firenze, Italy
INFN-Sezione di Firenze, via G. Sansone 1, I-50019 Sesto Fiorentino, Italy

Domenico Felice
School of Science and Technology
University of Camerino, I-62032 Camerino, Italy
INFN-Sezione di Perugia, Via A. Pascoli, I-06123 Perugia, Italy

Stefano Mancini
School of Science and Technology
University of Camerino, I-62032 Camerino, Italy
INFN-Sezione di Perugia, Via A. Pascoli, I-06123 Perugia, Italy

Marco Pettini
Aix-Marseille University, Marseille, France
CNRS Centre de Physique Théorique UMR7332, 13288 Marseille, France

October 22, 2014

Abstract

A central issue of the science of complex systems is the quantitative characterization of complexity. In the present work a new entropic measure of complexity is proposed which has unprecedented advantages. Starting from the framework of the so-called information geometry we propose a new and constructive way to associate to a - in principle any - network a differentiable object (a Riemannian manifold) whose volume is used to define an entropy. The effectiveness of this new entropic measure of networks complexity is successfully proved through its capability of detecting a classical phase transition in random graphs: the emergence of the “giant component” according to the celebrated Erdős-Renyi theorem.

Keywords: Complex systems; Riemannian geometry; Entropy; Probability theory.

1 Introduction

Complex systems and phenomena are dealt with in many scientific domains. According to the domain of interest, different definitions of complexity and of the way of measuring it have been proposed and are continuously being proposed since the science of complexity is still fast growing [1, 2]. In spite of their large number, 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 [5]. In the present work we propose to quantify the degree of complexity of networks through an entropy function inspired by microcanonical entropy of statistical mechanics. A network is a discrete system, roughly speaking a set of nodes (or vertices) and of links (or edges) among them, to which we propose to associate a suitably defined continuous and differentiable system, more precisely a Riemannian manifold. In so doing, as we shall see in the following, a constructive and computable definition can be given of an entropy function as the logarithm of the Riemannian volume of the manifold associated to the network. We proceed by associating a random variable to each vertex of a network and considering their correlations as weighted edges among such vertices. The nature of these variables characterizes a network (each node can contain energy, or information, or represent some internal parameter of a neuron in a neural network, or the concentration of a biomolecule in a complex network of biochemical reactions, and so on). These variables are assumed to be random either because of the difficulty of perfectly knowing their values or because of their intrinsic random dynamical properties. Thus, as is customary for probabilistic graphs models[6], we associate a joint probability mass function to the description of a network. Being easily manageable, Gaussian joint probability mass functions play a prominent role. Actually Gaussian networks are extensively used in many applications ranging from neural networks, to wireless communication, from proteins to electronic circuits, etc. Probably the most common (quantitative) tool to deal with networks complexity is the graph entropy (see [7] for a survey on the subject). However this has been inspired by the mentioned Shannon’s information theoretic approach, namely the optimal coding of a source of information having an ambiguous alphabet. In fact emitted symbols are put in correspondence to vertices and their distinguishability to edges [8]. Here we take a different avenue and introduce a geometric quantity as a candidate to measure networks complexity. This is obtained by resorting to the afore mentioned relation between networks and joint probability mass functions and introducing in the space of the latter a Riemannian structure stemming from information geometry [9]. A first attempt in this direction was proposed in [10]; however, as it will be discussed in the following, this approach cannot be constructively applied to networks having more than a few nodes. Thus in the present work a new metric - obtained by a suitable deformation of the standard Fisher-Rao metric of information geometry - is proposed which allows to constructively lift the properties of any given network to the geometric structure of a manifold. Since a paradigm for the emergence of new phenomena in complex systems is that of phase transitions, we shall check the effectiveness of our measure by applying it to a classical transition in random graphs predicted by the Erdős-Renyi theorem [11, 12] and we shall show that our entropy clearly reveals the emergence of the so called “giant component”. An unprecedented kind of direct test of validity of the proposed entropic measure of networks complexity.
2 Results

2.1 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, and so on). Likewise we endow a network with a statistical Riemannian manifold. This can be obtained basically via two steps; first by understanding a network as an undirected graph without loops on the nodes, and account for links (weighted edges) between nodes expressed by the so-called adjacency matrix $A$ \[13\], by means of correlations. Then, considering random variables as sitting on the vertices of a network, it can be employed methods of Information Geometry \[9\] to lift the network to a statistical Riemannian manifold.

So, let us consider a set of $n$ real-valued Gaussian random variables $X_1, \ldots, X_n$, each of them with zero mean value, defined on the sample space $\Omega$ with the joint density $p(x; \theta)$ given by

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

where $x^t = (x_1, \ldots, x_n) \in \mathbb{R}$ with $t$ denoting the transposition. Furthermore, $\theta^t = (\theta^1, \ldots, \theta^m)$ are the real valued parameters characterizing the above probability distribution function, i.e. the entries of the covariance matrix $C$. Hence $m = \frac{n(n+1)}{2}$.

Next we consider a family $P$ of such probability distributions $P = \{p_\theta = p(x; \theta) | \theta^t = (\theta^1, \ldots, \theta^m) \in \Theta\}$, where $\Theta \subseteq \mathbb{R}^m$ and the mapping $\theta \to p_\theta$ is injective. Defined in such a way $P$ is an $m$-dimensional statistical model on $\mathbb{R}^n$. The open set $\Theta$ is defined as follows

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

and we call it the parameter space of the $m$ statistical model $P$.

The mapping $\varphi : P \to \mathbb{R}^n$ defined by $\varphi(p_\theta) = \theta$ allows us to consider $\varphi = [\theta^t]$ as a coordinate system for $P$. Assuming parametrizations which are $C^\infty$ we can turn $P$ into a $C^\infty$ differentiable manifold \[9\].

Let $P = \{p_\theta | \theta \in \Theta\}$ be an $m$-dimensional statistical model. Given a point $\theta$, the Fisher information matrix of $P$ in $\theta$ is the $m \times m$ matrix $G(\theta) = [g_{\mu\nu}]$, where the $\mu, \nu$ entry is defined by

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

with $\partial_\mu$ standing for $\frac{\partial}{\partial \theta^\mu}$. The matrix $G(\theta)$ is symmetric, positive semidefinite and endows the parameter space $\Theta$ with a Riemannian metric \[9\].

Given the formal definition of the Fisher-Rao metric tensor \[3\], for the case under consideration, we highlight that the integral given by Eq. \[3\] is a Gaussian one. It is well-known that, in this case, we have

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

3
where the exponential stands for a power series expansion over its argument (the differential operator) and

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

The derivative of the logarithm has the following expression

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

The latter equation and Eq. (4) show the computational complexity of the Eq. (3) when the covariance matrix has off-diagonal entries different from zero. Indeed, the well-known formulas

\[ \partial_{\mu} C^{-1}(\theta) = C^{-1}(\theta)(\partial_{\mu} C(\theta))C^{-1}(\theta) \]
\[ \partial_{\mu}(\det C(\theta)) = \det C(\theta) \text{Tr}(C(\theta) \partial_{\mu}(C(\theta))) \]

require the calculation of \( n(n+1) \) derivatives with respect to the variables \( \theta \in \Theta \) of Eq. (2) in order to work out the derivative of the logarithm in (6). Finally, to obtain the function \( f_{\mu\nu} \) in (5), we have to evaluate \( O(n^4) \) derivatives. This quickly becomes an unfeasible task with growing \( n \), even numerically.

2.2 An alternative to the Fisher-Rao metric

In order to overcome the difficulty of explicitly computing the components of the Fisher-Rao metric, we proceed by defining a new Riemannian metric on the parameter space \( \Theta \) which account as well for the network structure given by the adjacency matrix \( A \).

To this end we consider first a trivial network with null adjacency matrix that is associated with a set of \( n \) independent Gaussian random variables \( X_i \). In this particular case, making use of Eqs. (2) and (3), to the network it corresponds the statistical Riemannian manifold \( \mathcal{M} = (\Theta, g) \)

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

where the \( \theta^i \)'s are the variances \( \theta^i = E[X_iX_i] \) of the random variables \( X_i \) that we assumed to have vanishing mean values. Recall that, in this case, the covariance matrix \( C(\theta) \) of the model (1) is a diagonal one. Let us remark that the entries \( g_{ii} \) of the metric \( g \) in (7) are related with the entries of the diagonal covariance matrix \( C(\theta) \) when the random variables are independent. In fact from (7) it can be easily checked that \( g_{ii} = \frac{1}{2} (c^{-1}_{ii})^2 \), where \( c^{-1}_{ii} \) is the \( ii \) entry of the inverse of the diagonal covariance matrix \( C(\theta) \). Inspired by this functional form of \( g \), we propose to associate a Riemannian manifold to any network \( X \) with non vanishing adjacency matrix \( A \) by deforming the manifold \( \mathcal{M} \) in (7) via the map \( \psi_C : \text{A}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R}) \) defined by

\[ \psi_C(\theta)(A) := C(\theta) + A, \] (8)
where $C(\theta)$ is the diagonal covariance matrix. With $A(n, \mathbb{R})$ we denote the set of the symmetric $n \times n$ matrices over $\mathbb{R}$ with vanishing diagonal elements that can represent any simple undirected graph. Hence the manifold associated to a network $\mathcal{X}$ with adjacency matrix $A$ is $\tilde{\mathcal{M}} = (\tilde{\Theta}, \tilde{g})$ with

$$\tilde{\Theta} := \{ \theta \in \Theta : \psi_{C(\theta)}(A) > 0 \} \quad (9)$$

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

$$\tilde{g}_{\mu \nu} = \frac{1}{2} \left( \psi_{C(\theta)}(A)^{-1}_{\mu \nu} \right)^2 \quad (10)$$

where $\psi_{C(\theta)}(A)^{-1}_{\mu \nu}$ is the $\mu \nu$ entry of the inverse of the matrix $\psi_{C(\theta)}(A)$.

### 2.3 A measure of networks complexity

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

$$S := \ln V(A) \quad (11)$$

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

$$\nu_g = \sqrt{\det \tilde{g}(\theta)} \ d\theta^1 \wedge \ldots \wedge d\theta^n \quad (12)$$

Notice, however, that in such a way $V(A)$ is ill-defined. Thus we regularize it as follows

$$V(A) := \int_{\tilde{\Theta}} \Upsilon(\psi_{C(\theta)}(A)) \nu_g \quad (13)$$

where $\Upsilon(\psi_{C(\theta)}(A))$ is any suitable ”infrared” and ”ultraviolet” regularizing function; $\psi_{C(\theta)}(A)$ and $\nu_g$ are given in (8) and (12), respectively. The need for regularization is twofold: the set $\tilde{\Theta}$ in Eq.(9) is not compact because the variables $\theta^i$ are unbound from above; furthermore, from Eq.(10), $\det \tilde{g}(\theta)$ diverges since $\det \psi_{C(\theta)}(A)$ approaches zero for some $\theta^i$. Thus the regularizing function $\Upsilon(\psi_{C(\theta)}(A))$ provides a kind of compactification of the parameter space and exclude the contributions of $\theta^i$ making $\det \tilde{g}(\theta)$ divergent.

The definition (11) is inspired by the microcanonical definition of entropy $S$ in statistical mechanics, that is $S := k_B \ln \Omega(E)$, where $\Omega(E)$ is the phase space volume bounded by the hypersurface of constant energy $E$. After integration on the momenta one finds $S = k_B \ln \varpi \int_{M_E} [E - V(q_1, \ldots, q_N)]^{N/2} dq^1 \wedge \ldots \wedge dq^n$, where $\varpi$ is a constant stemming from the integration on the momenta, $M_E$ is the configuration space subset bounded by the equipotential level set $E = V(q_1, \ldots, q_N)$, and $q_1, \ldots, q_N$ are the configurational coordinates. Now, the term $[E - V(q_1, \ldots, q_N)]^{N/2}$ is just $\sqrt{\det g_J}$, with $g_J$ the Jacobi kinetic energy metric whose associated geodesic flow coincides with the underlying Hamiltonian flow [14]. In the end the microcanonical entropy is $S = k_B \ln \int_{M_E} \sqrt{\det g_J} \ dq^1 \wedge \ldots \wedge dq^n + k_B \ln \varpi$, that is proportional to the logarithm of the volume of the Riemannian manifold associated with the underlying dynamics.
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 can wonder how to assess its effectiveness. Our proposal is to check a complexity measure against a system which makes a clear jump of complexity as some parameter is varied. In physics a paradigmatic situation is given by phase transitions (a snowflake is intuitively more complex than a drop of water). Applied to networks this leads us to consider the classical Erdős-Rényi phase transition in random graphs [11, 12].

The reason for this choice is twofold. First, real complex networks are so large that any attempt at completely describing them is hopeless, thus local descriptions of probabilistic kind are usually adopted. This naturally leads to considering random graphs as satisfactory models for complex networks. Second, the fundamental model of random graphs introduced by Erdős and Rényi [11] displays a fascinating phase transition phenomenon which is analytically predicted [12]; thus this provides a solid ground to check the effectiveness of the proposed complexity measure. Let us remark that many real networks are scale-free, a property which is not shared by the Erdős-Rényi random graphs, however, these can be easily generalized making them also scale-free [15].

2.4 The Erdős-Rényi phase transition

There are two basic models of random graphs: the binomial random graph \( G(n, p) \) and the uniform random graph \( G(n, k) \). The first one is defined by choosing at random, with binomial probability distribution \( B(n, p) \), a graph from the collection of all the graphs having \( n \) vertices, and where \( p \in \mathbb{R}, 0 \leq p \leq 1 \). Otherwise, the model \( G(n, k) \) 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 non negative integer. We can think of \( G(n, k) \) as a process evolving by adding the edges one at a time. When \( k \) has the same order of magnitude of \( n \), the evolution of \( G(n, k) \) from \( k = 0 \) to \( k = \left( \begin{array}{c} n \\ 2 \end{array} \right) \) yields, according to Erdős-Rényi theorem [11], a phase transition, revealing itself in a rapid growth with \( k \) of the size of the largest component (number of vertices fully connected by edges). Specifically, the structure of \( G(n, k) \) 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 \( G(n, k) \) rapidly grows, asymptotically almost surely (a.a.s.), from \( \log n \) to \( n \), if \( k \) has the same order of magnitude of \( n \). In fact, if \( k < n/2 \), as the process evolves, the components of \( G(n, k) \) [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 [12]. Nonetheless, at the same point of the process, the largest components become so large that it is likely for a new edge to connect two of them. Thus, fairly quickly, all the largest components of \( G(n, k) \) merge into one giant component, much larger than any of the remaining ones [12]. It is worth noticing that this process represents the mean-field case of percolation [16].

To give an example of the rapid growth of the size of the largest component of \( G(n, k) \), for \( G(n, 0.49n) \) there is no component with size of (asymptotically almost surely) order of magnitude larger than \( \log n \), whereas for \( G(n, 0.51n) \) a giant component already exists with a number of fully connected vertices of (asymptotically
The occurrence of this sudden structural change is explained by means of the Galton-Watson branching process \cite{12} detailing step by step the structure of the largest component of $G(n,p)$. The analysis of the behavior of $G(n,k)$ is similar, but the fact that the edges of $G(n,p)$ appear independently from each other makes the application of the branching process simpler. In particular, the process of generating the component containing a given vertex can be closely approximated by the stochastic process $Y = (Y_n)_{n \in \mathbb{N}}$ defined by the recurrence formula $Y_0 = 1$, $Y_{n+1} = \sum_{i=0}^{Y_n} \xi_i$, where the $\xi_i$s are independent random variables, each representing the appearance of a new link with binomial distribution $B(n,p)$. The random variables $Y_i$ represent the number of new vertices that are added to the component in the $i$-th step. Let $Z = \sum_{i \geq 0} Y_i$ be the total number of vertices added, the probability $\rho$ of extinction of the process is given by $\rho := \mathbb{P}(Z < \infty) = \lim_{n \to \infty} \mathbb{P}(Y_n = 0)$. Since the $\xi_i$s are binomial random variables it follows that $\rho$ is a random Poisson distributed variable with mean value $c = np$. In this approximation one expects that the probability that a vertex is contained in a small component is given by the probability that the process dies out. This happens with probability 1 for $c < 1$. On the other hand, if $c > 1$ then with some positive probability the process continues indefinitely and thus we may expect that about $O(n)$ vertices of $G(n,p)$ belong to one giant component. The equivalence of $G(n,p)$ and $G(n,k)$ \cite{12} implies that to the transition value $c = 1$ it corresponds the transition value $k = n/2$. Hence also for $G(n,k)$ we have the appearance of the giant component when $k$ crosses the transition value.

2.5 Numerical results

In the present work, we numerically compute $\mathcal{S} := \mathcal{S}(k)$ to investigate the sensitivity of the complexity measure in Eq. (11) to the appearance of the giant component during the evolution of the random graph model $G(n,k)$, neglecting finer details as, for instance, the topological structure of the graph at any fixed $k$.

To this aim, we consider four different numbers of vertices: $n = 25, 50, 100, 200$. These determine the dimensions of the associated manifolds $\mathcal{M}$. Then, for a given $n$, we choose the number of links $k$, with $k = 0, 1, \ldots, n(n-1)/2$. Next, for a given pair $(n,k)$ we generate at random a set of $k$ entries $(i,j)$, with $i < j$, of the non-vanishing adjacency matrix elements $A_{i,j}$.

Hence, since the covariance matrix $C$ is functionally assigned, we get $\psi(C)(A)$ of Eq. (8) and finally the metric $\tilde{g}$ of Eq. (10). Now, having determined $\mathcal{M} = (\hat{\Theta}, \tilde{g})$, we can compute the volume $V(A)$ in Eq. (13) and the entropy $S$ of Eq. (11). In numerical computations the volume regularization is performed in two steps, the first one is by restricting the manifold support $\hat{\Theta} \subset \mathbb{R}^n$ to an hypercube. Inside $\hat{\Theta}$ we generate a Markov chain, to perform a Monte Carlo estimate of the average $\langle \sqrt{\det g} \rangle = \int \sqrt{\det g} \ d\theta^1 \wedge \ldots \wedge d\theta^n / \int d\theta^1 \wedge \ldots \wedge d\theta^n$. The number of random configurations considered varies between $10^4$ and $10^6$; the second step of the regularization procedure of the volume is obtained by excluding those points where the value of $\sqrt{\det \tilde{g}}$ exceeds $10^{308}$ (the numerical overflow limit of the computers used). For any given pair $(n,k)$ this computational scheme is repeated $10^4$ times, each time considering a different randomly generated realisation of the adjacency matrix $A$. 

7
Thus the final values of the entropy $S(k/n)$ are obtained as averages on these $10^3$ different manifolds $\tilde{\mathcal{M}}$. In the following figures we report the Monte Carlo numerical estimates of

$$
\tilde{S}(k) := \frac{1}{n} \langle (S(k) - S_0) \rangle = \frac{1}{n} \left\langle \ln \frac{\int \sqrt{\det \bar{g}} \, d\theta^1 \wedge \ldots \wedge d\theta^n}{\int \sqrt{\det g} \, d\theta^1 \wedge \ldots \wedge d\theta^n} \rightangle
$$

(repeated for different values of $k$; here $\langle \cdot \rangle$ stands for the above mentioned average over the different realisations of the adjacency matrix $A$, and $g$ is the metric corresponding to the null adjacency matrix.

For all the four cases reported in Figure 1 we have equal weights $A_{ij} = r$ for all the $k$ non-vanishing links.

![Figure 1: Detection of the ”giant component” transition in random graphs. Complexity of $G(25, k)$ (magenta points), $G(50, k)$ (black points), $G(100, k)$ (red points) and $G(200, k)$ (blue points) networks as a function of the number $k$ of randomly chosen links of weights equal to $r = 0.2$. The black solid line is a guide to the eye coming from a linear fitting of a linear-logarithmic presentation of the data.](image-url)
Figure 2: Sensitivity of the complexity measure on the model parameters. Complexity of $G(50,k)$ networks as a function of the number $k$ of randomly chosen links of weights equal to $r = 0.1$ (green points), $r = 0.2$ (red points) and $r = 0.4$ (black points).

The reason for displaying $\tilde{S}$ of Eq. (14) instead of $S$ in (11) and versus $k/n$ instead of $k$ is that one obtains what in the context of statistical mechanics is called a collapse plot of the results obtained at different $n$-values. The corresponding points crowd on a common pattern for large $k$ whereas for $k/n$ ranging from 0 to approximately 1 the patterns obtained show a phenomenon which is familiar in the context of numerical investigations of second order phase transitions: as in the case of finite-size effects observed for the order parameter, what asymptotically would be a sharp bifurcation is rounded at finite $n$; however, the larger $n$ the more pronounced the "knee" of $\tilde{S}(k/n)$ in the range $(0,1)$. This is clearly in excellent agreement with an $n$-asymptotic bifurcation at $k/n = 0.5$ (marked by the solid line) where the Erdös-Rényi phase transition takes place. At present, this beautiful and unambiguous result lends credit to our proposed measure of networks complexity. Of course some basic checks of the stability of our result are in order.
In Figure 2 we report the outcomes for $G(50,k)$ having set all the non-vanishing entries $A_{ij}$ of the adjacency matrix again equal to a constant value $r$. For $r = 0.1$ a considerable softening of the shape of $\tilde{S}(k/n)$ is observed; this is of course an expected result since for $r \to 0$ the transition must disappear. For $r = 0.2$ and $r = 0.4$ the shapes of $\tilde{S}(k/n)$ look almost the same, the only interesting difference being a slightly more pronounced "knee" in the $r = 0.4$ case, which, going in the opposite direction, is coherent with the previous ones.

In Figure 3 we report the outcomes of $G(50,k)$ having chosen at random the values of the non-vanishing entries $A_{ij}$ of the adjacency matrix, that is, $A_{ij} = 0.2 + \omega$ where $\omega$ is a random variable of zero mean and variance equal to 0.1. Of course negative values of the $A_{ij}$ are excluded. The comparison with the results obtained with $A_{ij} = 0.2$ confirms the robustness of the entropy defined in Eq.(11).

Another interesting property of this entropy consists in its versatility, that is, it
can be easily adapted to more refined descriptions of networks/graphs.

For example, we can consider a refined modeling of graphs where the entries of the $n \times n$ adjacency matrix $A$ are given by terms of the form $r_{ij}\theta^i\theta^j$. Here the $r_{ij}$s ($i, j = 1, \ldots, n$) are the weights of the links between the nodes of the network described by $A$. Furthermore, the $\theta$'s ($i = 1, \ldots, n$) are local coordinates on the manifold $\tilde{M}$ of Eqs. (9) and (10), representing the variances of the random variables on the nodes of the network.

This kind of model has an interesting property: the closer a given variable $\theta^i$ to zero, the weaker the weights of all the links associated to the $i$-th node. In such a way, this second model, allows one to describe a more general class of networks. In fact, consider for example the flow of some quantity across a network, the vanishing of the flow on a given node $i$ implies that all the other nodes connected to it become effectively independent of it. In view of this argument, we consider the entropy $\tilde{S}$ in Eq. (11) against the more general model just described above. In Figure 4 it is shown $\tilde{S}$ versus $k/n$ where $n = 50$ is the dimension of the manifold associated to the network, and $k = 0, \ldots, n(n-1)/2$ is the number of non-vanishing $r_{ij}$ (all of them are chosen equal to 1), for $i, j = 1, \ldots, n$ and $i < j$. Also in this case our entropy detects the phase-transition predicted by the Erdös-Rényi theorem occurring at $k/n = 0.5$.

The pattern of $\tilde{S} = \tilde{S}(k/n)$ reported in Figure 4 shows a more pronounced ”knee” at the asymptotic transition value $k/n = 0.5$ with respect to what is found for the same $n$ value and is reported in Figures 2 and 3.

This result provides further evidence of the robustness of the proposed way of measuring the degree of complexity of a network.

3 Outlook on future developments

The geometric framework so far proposed paves the way to interesting developments. In fact, the two models considered above are adequate to describe those networks for which the weights of the links are given once and for all. Now, a relevant generalization made possible by the Riemannian geometric framework consists in considering the time evolution of a network. In order to do this, in addition to the $\theta$'s of the diagonal covariance matrix $C$ also the entries $\sigma_{ij} := A_{ij}$ of the adjacency matrix $A$ must be considered as local coordinates of the statistical manifold $\tilde{M}$ of Eqs. (9) and (10). In this way the dimension of $\tilde{M}$ increases from $n$ to $n+n(n-1)/2$.

Denoting with $\zeta^i = (\psi_C(A))_{lm}$ the $n(n+1)/2$ local coordinates of $\tilde{M}$, 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 $(\tilde{M}, \tilde{g})$, that is, through the geodesic flow given by the following set of equations

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

(15)

where the $\Gamma^i_{jk}$ are the standard Christoffel connection coefficients

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

(16)
Figure 4: Robustness of the detection of the "giant component" transition in random graphs. Complexity of $G(50, k)$ networks as a function of the number $k$ of randomly chosen non-vanishing adjacency matrix elements $A_{ij}$ of the form $r_{ij}\theta^i\theta^j$ with $r_{ij} = 0.2$.

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

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

(17)

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. (15) 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 $\tilde{S}(t)$ of its complexity. The fitting of empirical data concerning the true evolution of a real network by means of the model dynamics given by Eq.s (15) and (17) could allow to get relevant information about the laws that drive the network evolution (conservation, extremalization, optimisation of some quantities
Another prospective and remarkable application of the differential geometrical approach put forward in the present work, and notably related with the dynamical equations (15), concerns the study of the stability properties of a network. In fact, by setting $\zeta^i(t) \rightarrow \zeta^i(t) + \varphi^i(t)$, where $\varphi^i(t)$ are small functional perturbations, after substitution into Eq. (15) and using (17) one can work out the tangent dynamics equations in the form of a system of first order linear differential equations

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

These equations, numerically integrated together Eqs. (15) and (17), are the natural tool to investigate the stability of either stationary or non stationary states of a network, for example - for a stationary state - to investigate a network stability/instability under addition or deletion of one or more links and so on.

4 DISCUSSION

Summarizing, the present work contributes the fascinating subject of quantifying the degree of complexity of systems that are commonly defined as "complex". There is a large number of such definitions that are already available. Perhaps this history begins with Kolmogorov’s definition of algorithmic complexity [17, 18] which, in spite of its theoretical beauty, is hardly applicable in practice. Since then the many definitions put forward were adapted to the specific systems/problems tackled. However, the number of categories in 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 seducing way to proceed. In fact, in physics, entropy is just a measure of disorder and conversely negentropy - as defined by L.Brillouin a long time 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 a vast literature addressing the quantification of complexity on the basis of Shannon’s information entropy which, on the other hand, has its template in Boltzmann’s entropy of kinetic theory. However, what was still missing was a general definition of an entropic measure of complexity accounting for both 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 to statistical mechanics, however, instead of being modeled on Boltzmann entropy is rather modeled on the microcanonical ensemble definition of entropy. 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 the so-called Information Geometry and partly is 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 pro-
posed entropic-geometric measure of complexity. In order to do this we needed, so to
speak, a paradigmatic example of a major change of 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 - asso-
ciated with a sharp qualitative and quantitative change of 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 Erdős-Rényi theorem: a paradigmatic - and at present
unique - example of an analytically known major variation of the degree of com-
plexity of a good model of complex networks. This kind of check is unprecedented
and very successful, in fact, the entropic-geometric measure of complexity proposed
here displays both a pattern and its size-dependence which are typically found for
the order parameter of a second-order phase transition in physics.

Finally, the differential-geometric framework proposed opens some fascinating
perspectives of application to data analysis of real systems as well as to the study
of the time evolution of complex systems.

References

[1] Albert, R., and Barabasi, A.-L., Statistical Mechanics of Complex Networks,
Rev. Mod. Phys. 74, 47 (2002).

[2] Dorogovtsev, S. N., Goltsev, A. V., and Mendes, J. F. F., Critical Phenomena
in Complex Networks, Rev. Mod. Phys. 80, 1275 (2008).

[3] Shannon, C. E., A Mathematical Theory of Communication, Bell System Tech-
nical Journal 27, 379 (1948).

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

[5] Brillouin, L., Science and Information Theory, (Academic Press, 1962),
(reprinted: Courier Dover Publications, 2004).

[6] Keshav, S. Mathematical Foundations of Computer Networking, (Addison Wes-
ley, 2012).

[7] Simonyi, G., Graph Entropy: A Survey in: Combinatorial Optimization, (W.
Cook, L. Lovasz, and P. Seymour eds.), DIMACS Series in Discrete Mathemat-
ics and Theoretical Computer Science 20, 399 (1995).

[8] Körner, J., Transaction of the 6th Prague Conference on Information Theory,
p.411 (1973).

[9] Amari, S., and Nagaoka, H. Methods of Information Geometry, (Oxford Uni-
versity Press, 2000).
[10] Felice, D., Mancini, S., and Pettini, M. Quantifying Networks Complexity from Information Geometry Viewpoint, *J. Math. Phys.* **55**, 043505 (2014).

[11] Erdös, P., and Renyi, A. On the evolution of random graphs, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **5**, 17 (1960).

[12] Janson, S., Luczak, T., Rucinski, A. *Random Graphs*, (John Wiley and Sons Inc., 2000).

[13] Distel, R. *Graph Theory*, (Springer-Verlag, 2000).

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

[15] Britton, T., Deijfen, M., and Martin-Löf, A. Generating random graphs with prescribed degree distribution, *J. Stat. Phys.* **124**, 1377 (2006).

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

[17] Kolmogorov, A.N. Three approaches to the quantitative definition of information, *Probl. Inf. Transm. (USSR)* **1**, 4 (1965).

[18] Kolmogorov, A.N. Logical basis for information theory and probability theory, *IEEE Trans. Inf. Theory* **IT14**, 662 (1968).