

INFORMATION GEOMETRIC METHODS FOR COMPLEXITY

Domenico Felice¹, Carlo Cafaro², Stefano Mancini^{3,4}

¹*Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22-04103 Leipzig, Germany*

²*SUNY Polytechnic Institute, 12203 Albany, New York, USA*

³*School of Science and Technology, University of Camerino, 62032Camerino, Italy and*

⁴*INFN-Sezione Perugia, Via A. Pascoli, I-06123 Perugia, Italy*

Research on the use of information geometry (IG) in modern physics has witnessed significant advances recently. In this review article, we report on the utilization of IG methods to define measures of complexity in both classical and, whenever available, quantum physical settings. A paradigmatic example of a dramatic change in complexity is given by phase transitions (PTs). Hence we review both global and local aspects of PTs described in terms of the scalar curvature of the parameter manifold and the components of the metric tensor, respectively. We also report on the behavior of geodesic paths on the parameter manifold used to gain insight into the dynamics of PTs. Going further, we survey measures of complexity arising in the geometric framework. In particular, we quantify complexity of networks in terms of the Riemannian volume of the parameter space of a statistical manifold associated with a given network. We are also concerned with complexity measures that account for the interactions of a given number of parts of a system that cannot be described in terms of a smaller number of parts of the system. Finally, we investigate complexity measures of entropic motion on curved statistical manifolds that arise from a probabilistic description of physical systems in the presence of limited information. The Kullback-Leibler divergence, the distance to an exponential family and volumes of curved parameter manifolds, are examples of essential IG notions exploited in our discussion of complexity. We conclude by discussing strengths, limits, and possible future applications of IG methods to the physics of complexity.

PACS numbers: Complexity (89.70.Eg), Entropy (89.70.Cf), Information Theory (89.70.+c), Riemannian Geometry (02.40.Ky), Phase Transitions: General Studies (05.70.Fh), Probability Theory (02.50.Cw), Quantum Phase Transitions (64.70.Tg).

LEAD PARAGRAPH. Describing and, to a certain extent, understanding the concept of complexity have been investigated in a variety of research fields. Various ad hoc formal definitions and interpretations of the concept of complexity of physical systems have been proposed depending on the specific domain of interest [1]. Furthermore, since the science of complexity is growing rapidly, new complexity measures are continuously being developed. For instance, quantifying the degree of organization of a physical system often resorts to some type of entropy function stemming from the Shannon information entropy which can be shown to be itself equivalent to standard forms of entropy that appear in physics [2]. In particular, Information Geometry (IG) provides relevant tools for defining a geometric entropy for either measuring complex networks or identifying the most relevant interacting subsystems within composite complex systems at different scales of observation. Moreover, when combined with statistical inference tools, IG offers a convenient platform to characterize the complexity of statistical predictions in the presence of partial knowledge of the system being considered.

In this manuscript, starting from the observation that a change in the complexity behavior of a physical system can be geometrically detected in the proximity of a phase transition, we review information geometric aspects of both phase transitions and complexity measures in classical and quantum settings.

Given the wide range of applicability of methods of information geometry to the science of complexity, we argue that the collection of findings selected and discussed in our work may be useful to motivate other scientists to find a possible unifying information geometric complexity measure of universal applicability.

I. INTRODUCTION

Information geometry (IG) is the application of methods of modern geometry to sets of probability distributions (statistical models). IG has arisen from investigations of invariant geometrical structure within the framework of statistical inference. It provides a Riemannian metric and two dual affine connections in a statistical model [3]. Beyond the statistical inference, the geometric structure of the space of probability distributions plays a fundamental role in wider areas of information sciences, such as machine learning, signal processing, optimization as well as statistical physics and mathematical theory underlying neural networks [4]. Furthermore, the IG of statistical models can also be extended to quantum mechanics (QM), with QM regarded as a non-commutative extension of probability theory. However, the study of the so-called quantum information geometry has started only very recently and it can be essentially said to be in its infancy stage [5]. IG therefore has become a framework of great interest due to the insight it provides into information sciences and wherever information theoretical quantities like relative entropies find a natural geometric interpretation [6]. As such information geometric measures can be relevant in the science of complexity.

Phase transitions are the most impressive examples in nature of emergent phenomena. They are fairly well-understood from a theoretical standpoint and characterized by a sharp change in the complexity of the physical system that exhibits such emergent phenomena when a suitable control parameter exceeds a certain critical value. In classical thermodynamics [7, 8], a first-order phase transition is a phase transition of a pure substance if it is characterized by two phases which, when coexisting, exhibit different values of the first-order derivatives of the chemical potential with respect to temperature and pressure. The first-order derivatives of the chemical potential can be expressed in terms of the specific volume and the specific entropy. On the contrary, a second-order phase transition is a feature of pure substances that admit pairs of phases which, when coexisting, do not differ in the values of the specific volume and the specific entropy. Instead, they differ in crystallographic configuration, magnetic moment configuration (that is, ferromagnetic properties), symmetry, or other features, including the values of the second-order derivatives of the chemical potential with respect to temperature and pressure. From a quantum-mechanical perspective, first-order quantum phase transitions (QPTs, [9]) correspond to discontinuities in the first derivative of the ground state energy density. For instance, in many-body quantum systems, examples of first-order QPTs are points of degeneracy of the Hamiltonians [10]. Due to their limited physical interest, we do not take into consideration higher-order QPTs that correspond to discontinuities in the higher order derivatives of the ground state energy density.

Phase transitions were originally investigated in the framework of the Riemannian geometrization of classical Hamil-

tonian dynamics in Ref. [11]. Two main questions were addressed there: first, are there peculiar emerging features in the geometric properties associated with the chaotic Hamiltonian dynamics of a system in the proximity of a phase transition? For example, a suitable geometric indicator of complexity could be represented by the average curvature properties of the curved manifold underlying the dynamics. Second, how do curvature fluctuations and Lyapunov exponents behave at critical points? These questions were studied for the planar Heisenberg XY -model [11]. The cuspid-like behavior of the largest Lyapunov exponent as a function of the temperature (defined as the time average of the kinetic energy per degree of freedom) at the critical temperature was reported based on both numerical and analytical arguments. The same cuspid-like behavior was reported for the fluctuations of the Gaussian curvature at the critical temperature. Based on these findings, the so-called topological conjecture was advanced: in the vicinity of a second-order phase transition, the manifold underlying the chaotic Hamiltonian dynamics undergoes a topology change. Such a topological conjecture found further numerical support in Ref. [12] by studying the numerically-computed behavior of the variance of the scalar curvature of the manifold underlying a classical φ^4 lattice model. Furthermore, the conjecture was later analytically supported by the investigation of a mean-field model in Ref. [13]. For a theoretical discussion on the topological origin of a phase transition in a mean-field model, we refer to Ref. [14]. Within the IG framework, phase transitions are studied by analyzing the scalar curvature properties and the behavior of geodesics on the curved parameter manifold associated with the physical system being considered once a suitable metric tensor is introduced. For classical physical systems, the parameters that characterize the metric tensor are generally described by thermodynamical control parameters such as pressure, volume, and temperature. For quantum physical systems, instead, such parameters are in general the coupling constants that appear in the Hamiltonian of the system. The first investigations of the geometric structure of equilibrium thermodynamics were performed by Weinhold and reported in Refs. [15–19]. However, the origin of the application of Riemannian geometric methods into phase transition investigations is due to the inclusion of Einstein’s fluctuation theory [20, 21] in the axioms of thermodynamics. Indeed, in this case, it happens that thermodynamic systems can be described in terms of Riemannian manifolds [22]. In particular, within this geometric framework, the thermodynamic curvature is a suitable indicator of the nature (attractive or repulsive) of the interparticle interactions in the thermodynamic system. The original finding that lead to such a physical interpretation of the thermodynamical curvature was the fact that in an ideal gas there is no effective interparticle interaction and, remarkably, the curvature of the manifold that describes a classical ideal gas is identically zero [22]. This original interpretation of curvature was due to Ruppeiner, according to whom the curvature is positive when repulsive interactions are dominant, and negative when attractive interactions

dominate. It is worthwhile mentioning that there exist alternative interpretations of the sign of the thermodynamic curvature. For instance, according to Janyszek and Mrugała [23], the scalar curvature of the parameter manifold measures the stability of the physical system being considered. In particular, the larger the curvature, the more stable the system. The scientific finding that motivated such an interpretation was the fact that for a one-dimensional Ising model with short-range interactions, one observes large fluctuations (large curvature) at low temperatures and zero magnetic field in the ferromagnetic case and small fluctuations (small curvature) in the antiferromagnetic case. More specifically, it was reported in Ref. [23] that since the scalar curvature of the parameter space approaches infinity in the vicinity of critical points, the inverse of the scalar curvature was interpreted as a measure of the stability of the magnetic system being analyzed.

A major issue in complex systems, besides recognizing emergent phenomena like phase transitions, is to quantify their degree of complexity. In fact the concepts of information and distance play a key role in defining the notion of complexity of physical systems even though its origin is not completely understood [24–30]. 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 rapidly developing [24, 25]. Within the framework of classical physics, measures of complexity are understood in a more satisfactory manner. For instance, the Kolmogorov-Sinai metric entropy [31, 32], the sum of all positive Lyapunov exponents [33], is a powerful indicator of unpredictability in classical dynamical systems. It is a measure of the so-called algorithmic complexity of classical trajectories [34–37]. Other known measures of complexity include, but are not limited to, the logical depth [38], the thermodynamic depth [39], the computational complexity [40] and stochastic complexity [41]. Roughly speaking the logical and thermodynamic depths identify as complex whatever can be reached only through a difficult path. Each one of these complexity indicators captures, to a certain extent, our intuitive ideas about the meaning of complexity. Some of them just apply to computational tasks and, unfortunately, only very few of them may be generalized so that their field of applicability can include the investigation of actual physical processes. From an ideal viewpoint, a good definition of complexity should be both mathematically rigorous and have a clear intuitive interpretation so as to allow for the study of complexity-related problems in both computation theory and statistical physics. Naturally, a quantitative measure of complexity is truly useful if its range of applicability is not limited to a few unrealistic applications. Therefore, in order to properly define measures of complexity, not only should the motivations for defining such a measure be clearly stated, but also what feature the measure is intended to capture should be neatly described. Quantifying the complexity of complex networks where both the concepts of probability distributions and structure play leading roles

[42], studying the complexity of structures that characterize ensemble of objects [43], and, finally, investigating the complexity of statistical models [44] have witnessed an increasing research interest in recent years.

To the best of our knowledge, the first preliminary investigations on the use of methods of information geometry to quantify complexity appeared in 2006 in Refs. [45, 46]. In Ref. [45], Ay and collaborators exploited the information geometric framework of hierarchies of exponential families in order to propose a theory of complexity measures for finite random fields. In Ref. [46], Cafaro and collaborators investigated the complexity of geodesic paths on negatively curved statistical manifolds underlying the entropic dynamics of a two-dimensional system in the presence of incomplete knowledge on its microscopic degrees of freedom.

In this paper, we survey the application of IG methods to characterize aspects of complexity in applied science. By using IG techniques, we report on properties of both classical and quantum PTs in physical systems that exhibit critical behavior. Global and local aspects of PTs specified by means of the scalar curvature of the parameter manifold and the components of the metric tensor, respectively, will be reviewed. We shall also discuss the behavior of geodesic paths on the parameter manifold used to gain insights into the dynamics of PTs. We will then go through a geometric entropy that arises by associating smooth systems (Riemannian manifolds) to discrete systems (networks). In particular, we will show that such a geometric entropy is capable of detecting the Erdős-Rényi phase transition in uniform random graphs. Therefore, such geometric entropy will also be discussed as a candidate for measuring networks complexity [42]. Following this idea of characterizing complexity in terms of IG methods, we shall report on complexity measures displaying interactions of a fixed number of parts of a composite system that cannot be described by means of a smaller number of parts of the system [43]. Finally we shall review complexity measures of entropic motion on curved statistical manifolds that originate from a probabilistic description of physical systems in the presence of partial knowledge [44].

The layout of the manuscript is as follows. In Section II, we present the essential tools of information geometry needed to characterize the concepts of complexity and phase transitions in both classical and quantum physical systems. In Section III, we discuss the use of IG to characterize classical phase transitions of second-order that occur in a variety of physical settings, including the one-dimensional Ising model, the one-dimensional q -state Potts model, and the Ising model on dynamical planar random graphs. Methods of IG are used in Section IV to single out the Erdős-Rényi phase transition in uniform random graphs and subsequently to characterize the complexity of networks. Further methods of IG are presented in Section V to study the complexity of systems with multiple interacting units at different scales of interest, and, finally, to quantify the complexity of macroscopic predictions in the presence of limited

information in Section VI. In Section VII we review the use of IG techniques to study quantum phase transitions of second-order in the periodic XY spin chain in a transverse magnetic field and in the Dicke model of quantum optics. In particular, for the former model, we discuss both a curvature-based and a geodesic-based analysis. Finally, after briefly addressing the notion of complexity of quantum evolution, we discuss both an information geometric perspective on the phenomena of softening of chaos by quantization and a statistical complexity measure of probability distributions that characterize quantum mechanical systems exhibiting nontrivial dynamics. A summary of conclusive results together with several considerations of unresolved issues concerning the use of information geometric methods to quantify the complexity of quantum mechanical systems are reported in Section VIII. Useful tools of differential geometry appear in Appendix A.

II. BASICS OF INFORMATION GEOMETRY

In this section, referring to [3–5], we introduce the basic notions of IG, most of which will be employed in the following parts of the manuscript.

A statistical manifold \mathcal{M} can be described as a set of parametrized probability distribution functions (pdfs), with a tensor metric g and an affine connection ∇ . Let $p(\mathbf{x}; \boldsymbol{\theta})$ denote the pdf of \mathbf{x} specified by a parameter vector $\boldsymbol{\theta} = (\theta^1, \dots, \theta^n)$. We assume the regularity conditions that the density functions $p(\mathbf{x}; \boldsymbol{\theta})$ exist on some carrier measure of the space \mathcal{X} of \mathbf{x} , that

$$p(\mathbf{x}; \boldsymbol{\theta}) > 0, \text{ and } \int_{\mathcal{X}} p(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x} = 1 \quad (1)$$

for all \mathbf{x} in the common domain \mathcal{X} . Furthermore, the mapping $\varphi : p(\mathbf{x}; \boldsymbol{\theta}) \mapsto \boldsymbol{\theta}$ is an homeomorphism between \mathcal{M} and the open set $\Theta \subset \mathbb{R}^n$ of the parameters $\boldsymbol{\theta}$. In addition, we assume that (\mathcal{M}, φ) is a differentiable structure; in such a way, \mathcal{M} is an n -dimensional C^∞ -manifold. In addition, let

$$l(\mathbf{x}; \boldsymbol{\theta}) \stackrel{\text{def}}{=} \ln p(\mathbf{x}; \boldsymbol{\theta}). \quad (2)$$

Then, we may regard every point $\boldsymbol{\theta}$ of Θ as carrying a function $l(\mathbf{x}; \boldsymbol{\theta})$ of \mathbf{x} . In this way, we can have a useful representation of vector fields in terms of the basis $\{\partial_i l(\mathbf{x}; \boldsymbol{\theta})\}$ with $\partial_i \stackrel{\text{def}}{=} \frac{\partial}{\partial \theta^i}$. This representation has the advantage that a vector field has zero expectation value with respect to the pdf $p(\mathbf{x}; \boldsymbol{\theta})$,

$$\mathbb{E}_p [\partial_i l(\mathbf{x}, \boldsymbol{\theta})] \stackrel{\text{def}}{=} \int_{\mathcal{X}} p(\mathbf{x}; \boldsymbol{\theta}) \frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta^i} d\mathbf{x} = \frac{\partial}{\partial \theta^i} \int_{\mathcal{X}} p(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x} = 0. \quad (3)$$

In addition to the connection ∇ , the statistical manifold \mathcal{M} also carries a dual connection ∇^* which is defined as [5]

$$g(\nabla_X^* Y, Z) \stackrel{\text{def}}{=} Xg(Y, Z) - g(\nabla_X Y, Z), \quad (4)$$

with X , Y , and Z denoting vector fields on \mathcal{M} . A typical example of pdfs is provided by the so-called exponential family [47]. For such a family, the notion of distinguishability between pdfs is quantified by means of the so-called Fisher–Rao information metric tensor [47]

$$g_{ij}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \mathbb{E}_p [\partial_i l(\mathbf{x}; \boldsymbol{\theta}) \partial_j l(\mathbf{x}; \boldsymbol{\theta})]. \quad (5)$$

Furthermore, the affine connection ∇ related to g_{ij} in Eq. (5) is known as the exponential connection and can be defined in terms of the following Christoffel coefficients

$$\Gamma_{ijk} \stackrel{\text{def}}{=} g(\nabla_{\partial_i} \partial_j, \partial_k) = \mathbb{E}_p [\partial_i \partial_j l(\mathbf{x}, \boldsymbol{\theta}) \partial_k l(\mathbf{x}, \boldsymbol{\theta})]. \quad (6)$$

Associated to the exponential connection ∇ , the dual connection ∇^* is given by

$$\Gamma_{ijk}^* \stackrel{\text{def}}{=} g(\nabla_{\partial_i}^* \partial_j, \partial_k) = \mathbb{E}_p [(\partial_i \partial_j l(\mathbf{x}, \boldsymbol{\theta}) + \partial_i l(\mathbf{x}, \boldsymbol{\theta}) \partial_j l(\mathbf{x}, \boldsymbol{\theta})) \partial_k l(\mathbf{x}, \boldsymbol{\theta})]. \quad (7)$$

Indeed, a statistical structure on the manifold \mathcal{M} is given in terms of a dualistic structure (g, ∇, ∇^*) . When an exponential family is considered, it is well-known that the Riemann curvature tensor \mathcal{R} with respect to the exponential affine connection ∇ [Appendix A] is zero implying that such a statistical manifold is ∇ -flat. In this particular case, even the Riemann curvature tensor \mathcal{R}^* with respect to the dual affine connection ∇^* equals zero. Hence, it is evident that the exponential family is a dually flat statistical manifold [5]. Given a statistical geometric structure $(\mathcal{M}, g, \nabla, \nabla^*)$, the information about the statistical geometric structure is retained in a two points differentiable real-valued function \mathcal{D} [48]. This is a divergence function $\mathcal{D} : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}^+$ for the manifold \mathcal{M} which is a distance-like measure of the separation between two points $p, q \in \mathcal{M}$ such that

$$\mathcal{D}[p, q] \geq 0, \text{ and } \mathcal{D}[p, q] = 0 \Leftrightarrow p = q. \quad (8)$$

Given a coordinate system $[\theta^i]$, one can represent a pair of points $(p, q) \in \mathcal{M} \times \mathcal{M}$ as a pair of coordinates $([\theta^i(p)], [\theta^i(q)])$. In particular, the metric tensor g and the connection ∇ provide a third order approximation for \mathcal{D} [5],

$$\begin{aligned} \mathcal{D}[p, q] &= \frac{1}{2} g_{ij} \Delta \theta^i \Delta \theta^j + \frac{1}{6} h_{ijk} \Delta \theta^i \Delta \theta^j \Delta \theta^k + o(\|\Delta \theta\|^3) \\ &= \frac{1}{2} g_{ij} \Delta \theta^i \Delta \theta^j - \frac{1}{6} h_{ijk}^* \Delta \theta^i \Delta \theta^j \Delta \theta^k + o(\|\Delta \theta\|^3), \end{aligned} \quad (9)$$

where $\Delta\theta^i \stackrel{\text{def}}{=} \theta^i(p) - \theta^i(q)$ and $o(\|\Delta\theta\|^3)$ is a term that vanishes at a rate faster than $\|\Delta\theta\|^3$ as p tends to q . The coefficients h_{ijk} and h_{ijk}^* define a 3-covariant symmetric tensor which, in terms of the metric tensor g and the connection coefficients Γ_{ijk} of the affine connection ∇ , is given by

$$h_{ijk} = \partial_i g_{jk} + \Gamma_{jk,i}, \text{ and, } h_{ijk}^* = \partial_i g_{jk} + \Gamma_{jk,i}^*, \quad (10)$$

respectively, where $\partial_i \stackrel{\text{def}}{=} \frac{\partial}{\partial\theta^i}$. Conversely, given a divergence function \mathcal{D} for $(\mathcal{M}, g, \nabla, \nabla^*)$ and the standard coordinate systems $([\theta^i], [\theta'^i])$, the statistical geometric structure is recovered as follows [5],

$$g_{ij} = -\partial_i \partial'_j \mathcal{D}[p, q]_{|p=q}, \text{ and } \Gamma_{ij,k} = -\partial_i \partial_j \partial'_k \mathcal{D}[p, q]_{|p=q}, \quad (11)$$

where $\partial'_i \stackrel{\text{def}}{=} \frac{\partial}{\partial\theta'^i}$ is the partial derivative with respect to the second variable. Despite the infinitely many ways in which one can define a divergence function that generates a given geometry, it is often useful to define a divergence function which is, in some sense, canonical. A canonical divergence satisfies desirable properties such as the generalized Pythagorean theorem and the geodesic projection theorem [5]. In the particular case of a dually flat statistical manifold, a canonical divergence (a Bregman-type of divergence) was introduced by Amari and Nagaoka in Ref. [5]. An exponential model is a statistical manifold that consists of probability distribution functions $e(\mathbf{x}; \boldsymbol{\theta})$ defined as

$$e(\mathbf{x}; \boldsymbol{\theta}) \stackrel{\text{def}}{=} \exp \left[C(\mathbf{x}) + \sum_{i=1}^n \theta^i F_i(\mathbf{x}) - \psi(\boldsymbol{\theta}) \right], \quad (12)$$

where $\{C, F_1, \dots, F_n\}$ are functions on the sample space \mathcal{X} and ψ is a function on the space Θ of the parameters $\boldsymbol{\theta}$. From the normalization condition $\int_{\mathcal{X}} e(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x} = 1$, we obtain

$$\psi(\boldsymbol{\theta}) = \ln \left\{ \int_{\mathcal{X}} \exp \left[C(\mathbf{x}) + \sum_{i=1}^n \theta^i F_i(\mathbf{x}) - \psi(\boldsymbol{\theta}) \right] d\mathbf{x} \right\}. \quad (13)$$

We refer to this function ψ in Eq. (13) as the potential function of the exponential family. Indeed, the family $\mathbf{E} = \{e(\mathbf{x}; \boldsymbol{\theta})\}$ is a statistical manifold with the metric tensor g as in Eq. (5) and the connections ∇ and ∇^* as in Eqs. (6),(7), respectively. We shall refer to such an information geometric structure as $(\mathbf{E}, g, \overset{e}{\nabla}, \overset{e}{\nabla}^*)$. From the definition of an exponential family given in Eq. (12), we have

$$\partial_i l(\mathbf{x}; \boldsymbol{\theta}) = F_i(\mathbf{x}) - \partial_i \psi(\boldsymbol{\theta}), \text{ and } \partial_i \partial_j l(\mathbf{x}; \boldsymbol{\theta}) = -\partial_i \partial_j \psi(\boldsymbol{\theta}). \quad (14)$$

Furthermore, note that from Eqs. (3) and (14), it is possible to rewrite the components of the Fisher-Rao information metric $g_{ij}(\boldsymbol{\theta})$ as follows,

$$g_{ij}(\boldsymbol{\theta}) = -\mathbb{E}_p[\partial_i \partial_j l(\mathbf{x}; \boldsymbol{\theta})] = \partial_i \partial_j \psi(\boldsymbol{\theta}). \quad (15)$$

In addition, from Eqs. (3) and (14), we determine that the connection coefficients of the affine connection $\overset{e}{\nabla}$ are given by

$$\overset{e}{\Gamma}_{ij,k} = -\partial_i \partial_j \psi(\boldsymbol{\theta}) \mathbb{E}_{\boldsymbol{\theta}}[\partial_k l(\mathbf{x}; \boldsymbol{\theta})] = 0. \quad (16)$$

Eq. (16) implies that the exponential family \mathbf{E} is a flat manifold with respect to the affine connection $\overset{e}{\nabla}$ and $[\theta^i]$ is called an affine coordinate system of $\overset{e}{\nabla}$ for \mathbf{E} . Associated with $[\theta^i]$, a dual coordinate system $[\eta_i]$ can be defined as $\eta_i \stackrel{\text{def}}{=} \frac{\partial \psi}{\partial \theta^i}$. The coordinates $[\theta^i]$ and $[\eta_i]$ are dual with respect to the Fisher-Rao information metric g in the sense that the following relation holds true,

$$g(\partial_i, \partial^j) = \delta_i^j, \quad (17)$$

where $\partial^i \stackrel{\text{def}}{=} \frac{\partial}{\partial \eta_j}$. It is straightforward from the definition of η_j that Eq. (15) is equivalent to

$$g_{ij} = \frac{\partial \eta_j}{\partial \theta^i}, \text{ and } g^{ij} = \frac{\partial \theta^i}{\partial \eta_j}, \quad (18)$$

where g^{ij} are the components of the inverse of the Fisher-Rao information metric g . Therefore, together with the $\overset{e}{\nabla}$ -flatness, the manifold \mathbf{E} is also $\overset{e}{\nabla}^*$ -flat and $[\eta_i]$ becomes as an affine coordinate system of $\overset{e}{\nabla}^*$ for \mathbf{E} . For this reason, one says that the exponential manifold $(\mathbf{E}, g, \overset{e}{\nabla}, \overset{e}{\nabla}^*)$ endowed with the Fisher-Rao information metric g and exponential connection $\overset{e}{\nabla}$ is a dually flat statistical manifold. To complete the investigation of dualistic geometric structure of \mathbf{E} , we can consider the dual potential function φ of ψ being defined as,

$$\varphi(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \theta^i \eta_i - \psi(\boldsymbol{\theta}). \quad (19)$$

It immediately follows from (19) that $\theta^i = \partial^i \varphi$ and, as a consequence, from Eq. (18) we find

$$g^{ij} = \partial^i \partial^j \varphi. \quad (20)$$

At this point, given the dual flat manifold $(\mathbf{E}, g, \overset{e}{\nabla}, \overset{e}{\nabla}^*)$, the coordinate systems $[\theta^i]$ and $[\eta_i]$, and the potentials ψ and φ , the canonical divergence \mathcal{D} becomes

$$\mathcal{D}[p_{\boldsymbol{\theta}}, p_{\boldsymbol{\theta}'}] = \psi(p_{\boldsymbol{\theta}}) - \varphi(p_{\boldsymbol{\theta}'}) - \theta^i \eta_i. \quad (21)$$

In the case of an exponential family, $\mathcal{D}[p_{\boldsymbol{\theta}}, p_{\boldsymbol{\theta}'}]$ in Eq. (21) becomes the well-known Kullback–Leibler divergence,

$$D_{\text{KL}}[p_{\boldsymbol{\theta}}, p_{\boldsymbol{\theta}'}] \stackrel{\text{def}}{=} \int_{\mathcal{X}} p(\mathbf{x}; \boldsymbol{\theta}') \ln \left(\frac{p(\mathbf{x}; \boldsymbol{\theta}')}{p(\mathbf{x}; \boldsymbol{\theta})} \right) d\mathbf{x}. \quad (22)$$

It is straightforward to check that $\mathcal{D}_{\text{KL}}[p_{\boldsymbol{\theta}}, p_{\boldsymbol{\theta}'}] \geq 0$ and $\mathcal{D}_{\text{KL}}[p_{\boldsymbol{\theta}}, p_{\boldsymbol{\theta}'}] = 0$ if and only if $p_{\boldsymbol{\theta}} = p_{\boldsymbol{\theta}'}$. In addition, by means of simple algebraic computations, it follows that

$$\frac{\partial^2 \mathcal{D}_{\text{KL}}}{\partial \theta^i \partial \theta^j} \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}'} = g_{ij}(\boldsymbol{\theta}), \quad -\frac{\partial^3 \mathcal{D}_{\text{KL}}}{\partial \theta^i \partial \theta^j \partial \theta^k} \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}'} = \nabla_{ijk}^e, \quad \text{and} \quad -\frac{\partial^3 \mathcal{D}_{\text{KL}}}{\partial \theta^i \partial \theta^j \partial \theta^k} \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}'} = \nabla_{ijk}^{*e}. \quad (23)$$

As previously mentioned, we refer to Refs. [3–5] for additional mathematical details on IG.

III. INFORMATION GEOMETRY AND PHASE TRANSITIONS

In this section, we report on the use of IG methods to characterize classical phase transitions of second-order that emerge in a variety of physical scenarios.

The information theoretic analysis of second-order phase transitions in classical systems relies on the definition of a Riemannian metric on the equilibrium thermodynamic state space of a system. In Ref. [49], Ruppeiner conjectured that the scalar curvature \mathcal{R} [Appendix A] arising from such a metric is related to the correlation volume of the system, i.e.,

$$\mathcal{R} \sim \xi^d, \quad (24)$$

where ξ is the correlation length and d is the system dimension. This association follows from the idea that to a greater “distance” between two equilibrium thermodynamic states there corresponds a smaller probability that these can be related by a fluctuation. We remark that from a physics standpoint, the sign of \mathcal{R} is linked to the nature of the interparticle interactions in composite thermodynamical systems [50]. Specifically, the sign of the scalar curvature regarded as a measure of the strength of interaction can be zero, positive or negative when such interactions are absent (that is, flat geometry for an ideal gas), repulsive or attractive, respectively. The “distance” between two probability distributions in parametric statistics can be evaluated, as mentioned in Section II, starting from the Fisher-Rao information metric tensor of the system. In statistical mechanics, the probability distributions $p(\mathbf{x}; \boldsymbol{\theta})$ of major interest are those exhibiting the so-called Gibbs form,

$$p(\mathbf{x}; \boldsymbol{\theta}) \stackrel{\text{def}}{=} \exp \left[- \sum_{i=1}^n \theta^i \mathcal{H}_i(\mathbf{x}) - \ln \mathcal{Z}(\boldsymbol{\theta}) \right], \quad (25)$$

where \mathbf{x} characterizes the state of the system (for instance, spins), $\mathcal{H}_i(\mathbf{x})$ are the various Hamiltonian terms, $\mathcal{Z}(\boldsymbol{\theta})$ is the normalizing partition function and, finally, θ^i are the statistical parameters (for instance, the inverse temperature β , the external magnetic field intensity h , and so on). The n -dimensional manifold \mathcal{M} of statistical parameters is endowed

with the Fisher-Rao information metric that measures the distance between different probabilistic configurations. The components of the Fisher-Rao information metric tensor can be expressed as,

$$g_{ij} \stackrel{\text{def}}{=} \partial_i \partial_j f, \quad (26)$$

where $f \stackrel{\text{def}}{=} \ln \mathcal{Z}$ is the reduced free energy per site. For one-dimensional spin models, \mathcal{M} is a two-dimensional manifold with local coordinates $(\theta^1, \theta^2) = (\beta, h)$ and the scalar curvature \mathcal{R} is given by [49],

$$\mathcal{R} \stackrel{\text{def}}{=} -\frac{1}{2(\det g)^2} \det \begin{pmatrix} \partial_\beta^2 f & \partial_\beta \partial_h f & \partial_h^2 f \\ \partial_\beta^3 f & \partial_\beta^2 \partial_h f & \partial_\beta \partial_h^2 f \\ \partial_\beta^2 \partial_h f & \partial_\beta \partial_h^2 f & \partial_h^3 f \end{pmatrix}. \quad (27)$$

It is worth noticing that, unlike many statistical mechanical quantities, the curvature \mathcal{R} in Eq. (27) depends on third order derivatives of f . However, the consideration of \mathcal{R} offers the possibility of determining critical exponents in a non-standard way. Indeed, the hypothesis that for a second-order phase transition the curvature depends on the correlation volume combined with standard scaling assumptions and the hyperscaling relation, $\nu d = 2 - \alpha$ [51], leads to

$$\mathcal{R} \sim |\beta - \beta_c|^{\alpha-2}, \quad (28)$$

where α is the exponent characterizing the scaling of the specific heat while β_c denotes the inverse of the critical temperature.

A. The one-dimensional Ising model

The scaling behavior in Eq. (28) near criticality of the scalar curvature \mathcal{R} of the two-dimensional parameter manifold was analyzed in Ref. [23] for the one-dimensional (1D) Ising model defined by the Hamiltonian,

$$\mathcal{H}_{1D-\text{Ising}} \stackrel{\text{def}}{=} -J \sum_{k=1}^N s_k s_{k+1} - h \sum_{k=1}^N s_k, \quad (29)$$

where the spins $\{s_k\}_{k=1,\dots,N}$ are such that $s_k = +1$ (spin up) or $s_k = -1$ (spin down), and $s_{N+k} = s_k$. Moreover, J is the coupling constant for the nearest-neighbor pairs in the lattice and h denotes the external magnetic field multiplied by the magnetic dipole moment of one spin. The equilibrium probability distribution function ρ of the 1D Ising model

is given by,

$$\rho \stackrel{\text{def}}{=} \frac{\exp(-\beta \mathcal{H})}{\mathcal{Z}(\beta, \beta h)}, \quad (30)$$

where \mathcal{H} is defined in Eq. (29), $\beta \stackrel{\text{def}}{=} T^{-1}$ (the Boltzmann constant is set equal to one in the manuscript), and \mathcal{Z} denotes the partition function,

$$\mathcal{Z}(\beta, \beta h) \stackrel{\text{def}}{=} e^{N\beta J} \left\{ \cosh(\beta h) + [\cosh^2(\beta h) - 2e^{-2\beta J} \sinh(2\beta J)]^{\frac{1}{2}} \right\}^N. \quad (31)$$

The information geometry for this model was implemented in the limit of large N in Ref. [23]. In particular, its curvature $\mathcal{R}_{1D\text{-Ising}}$ is given by

$$\mathcal{R}_{1D\text{-Ising}} = 1 + [\sinh^2(\beta h) + e^{-4\beta J}]^{-\frac{1}{2}} \cosh(\beta h). \quad (32)$$

It is interesting to note that $\mathcal{R}_{1D\text{-Ising}}$ in Eq. (32) is a positive function of βh and βJ .

In particular, the scalar curvature does not depend on the orientation of the external magnetic field h since it is a symmetric function of the latter, being $\mathcal{R}_{1D\text{-Ising}}(h) = \mathcal{R}_{1D\text{-Ising}}(-h)$. The 1D Ising model can be thought of as having a zero-temperature phase transition. Therefore, near the criticality regime for the 1D Ising model, in the limit of $h = 0$ and β approaching infinity, from Eq. (32) one obtains (setting $J = 1$)

$$\mathcal{R}_{1D\text{-Ising}} \sim e^{2\beta}, \quad (33)$$

corresponding to the expected $\alpha = 1$ and $\nu = 1$ with $\mathcal{R}_{1D\text{-Ising}} \sim \xi^d = |\xi|^{\frac{2-\alpha}{\nu}}$ and $\xi \stackrel{\text{def}}{=} -\ln[\tanh(\beta)] \sim e^{2\beta}$. [52]

B. The one-dimensional q -state Potts model

In analogy to the one-dimensional Ising model, it is possible to obtain the expression of the scalar curvature \mathcal{R} for the curved parameter manifold that characterizes the one-dimensional q -state Potts model with $q = 2$ for the Ising model [52, 53]. The partition function of the 1D q -state Potts model is defined as,

$$\mathcal{Z}_N(y, z) \stackrel{\text{def}}{=} \sum_{\{s\}} \exp \left\{ \beta \sum_{k=1}^N \left[\delta(s_k, s_{k+1}) - \frac{1}{q} \right] + h \sum_{k=1}^N \left[\delta(s_k, 1) - \frac{1}{q} \right] \right\}, \quad (34)$$

with the spins $s_k \in \{1, \dots, q\}$. In this case, the scalar curvature $\mathcal{R}_{\text{Potts}}$ assumes the functional form,

$$\mathcal{R}_{\text{Potts}}(q, e^\beta, e^h) \stackrel{\text{def}}{=} A(q, e^\beta, e^h) + \frac{B(q, e^\beta, e^h)}{\sqrt{\eta(q, e^\beta, e^h)}}, \quad (35)$$

where $\eta(q, e^\beta, e^h)$ is given by,

$$\eta(q, e^\beta, e^h) \stackrel{\text{def}}{=} [e^\beta (1 - e^h) + q - 2]^2 + (q - 1) e^h. \quad (36)$$

The quantities A and B in Eq. (35) are smooth functions of e^β and e^h and, importantly, do not diverge for finite values of temperature and/or fields. Observe that Eq. (35) reduces to Eq. (33) for $q = 2$. [52] Unlike the 1D Ising model, the scalar curvature in Eq. (35) is no longer positive definite and the $h \rightarrow -h$ symmetry is no longer present. Finally, despite the absence of a true phase transition, in the limit of $h = 0$ and β approaching infinity, Eq. (35) yields the divergence of the scalar curvature $\mathcal{R}_{\text{Potts}}$ at zero temperature,

$$\mathcal{R}_{\text{Potts}} \sim e^\beta. \quad (37)$$

For further details on the 1D q -state Potts model, we refer to Refs. [52, 54, 55].

C. The Ising model on dynamical planar random graphs

Within the framework of classical PTs, the examples of 1D Ising model and 1D q -state Potts model exhibit unsatisfactory characteristics: the former has no real phase transition, while the latter is mean field in nature. These facts motivated the search for soluble models beyond mean-field theory and with a genuine finite-temperature phase transition. One such case is represented by the so-called Ising model on planar random graphs [51]. In Refs. [56, 57], the first investigations of the Ising model on an ensemble of Φ^4 (4-regular) or Φ^3 (3-regular) planar random graphs were presented. These investigations were motivated by string-theoretic considerations where the continuum limit of the theory represents matter coupled to 2D quantum gravity [56, 57].

In Ref. [51], the scaling behavior of the scalar curvature for the Ising model on an ensemble of planar random graphs (that is, a graph in which each possible edge is present or not with a certain probability) was investigated from an information geometric perspective. The partition function \mathcal{Z} of this model is given by,

$$\mathcal{Z} \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} \left[\frac{-4ge^{-2\beta}}{(1 - e^{-4\beta})^2} \right]^n \mathcal{Z}_n, \quad (38)$$

where

$$\mathcal{Z}_n \stackrel{\text{def}}{=} \sum_{\{G^n\}} \sum_{\{s\}} \exp \left(\beta \sum_{\langle i,j \rangle} G_{ij}^n s_i s_j + h \sum_i s_i \right), \quad (39)$$

g is the coupling constant, G_{ij}^n is the connectivity matrix for an n -vertex planar graph, and G^n denotes a graph with n -vertices. A finding of interest uncovered in this investigation is that the general scaling behavior of the scalar curvature $\mathcal{R}_{\text{Ising-PRG}}$ is not recovered,

$$\mathcal{R}_{\text{Ising-PRG}} \sim \xi^d \sim |\beta - \beta_c|^{\alpha-2}, \quad (40)$$

where, as pointed out earlier, ξ is the correlation length, d is the dimensionality of the system, α is the critical exponent, and β_c denotes the critical value of β . Indeed, setting $h = 0$, omitting technical details that can be found in Ref. [54], the leading asymptotic term of the scalar curvature $\mathcal{R}_{\text{Ising-PRG}}$ becomes

$$\mathcal{R}_{\text{Ising-PRG}} \sim \frac{225}{176} |\beta - \beta_c|^{-2}. \quad (41)$$

However, while $\alpha = 1$ for the standard 1D Ising model, the standard critical exponent for this model is $\alpha = -1$ [56, 57]. Therefore, there is a discrepancy between Eq. (40) with $\alpha = -1$ and Eq. (41). Such discrepancy was ascribed to the effect of the negativity of the critical exponent on the components of the metric tensor and, as a consequence, on the scalar curvature. For further details, we refer to Ref. [51].

For the sake of completeness, we point out here that the spherical model of a ferromagnet [58] was also studied from an information geometric perspective [59]. In analogy to Ref. [51], the scalar curvature of the this model exhibits a scaling behavior of the form $\mathcal{R}_{\text{Ising-PRG}} \sim |\beta - \beta_c|^{-2}$, although the critical exponent α equals -1 in this case [58]. For further details, we refer to ref. [59]. Finally, we refer to Ref. [60] for an intriguing investigation on the physical interpretation of the scalar curvature of parameter manifolds corresponding to finite Ising models.

D. The Erdős-Rényi phase transition in uniform random graphs

The Erdős-Rényi theorem [61] proves the existence of a phase transition undergone by random graphs. This is a classical example of an analytically known major variation of the degree of complexity of a network. It reveals in a rapid growth of the largest components merging suddenly in a giant component, much larger than any of the remaining ones. In the next section we will address this paradigmatic phase transition by making use of a geometric entropy [62].

IV. GEOMETRIC ENTROPY AND COMPLEX NETWORKS

In this section, we report on a Riemannian geometric entropy that is above all an attempt to lift heterogeneous systems (complex networks) to homogeneous systems (differential manifolds) [63]. More precisely, this modelling scheme consists in considering random variables as sitting on the vertices of a network and their correlations as weighted edges. Specifically, consider a set of n random variables X_1, \dots, X_n distributed according to a multivariate

Gaussian probability distribution $p(\mathbf{x}; \boldsymbol{\theta})$ with zero mean value,

$$p(\mathbf{x}; \boldsymbol{\theta}) \stackrel{\text{def}}{=} \frac{\exp \left[-\frac{1}{2} \mathbf{x}^\top \Sigma^{-1} \mathbf{x} \right]}{[(2\pi)^n \det \Sigma]^{1/2}} \quad (42)$$

where $\mathbf{x}^\top = (x_1, \dots, x_n) \in \mathbb{R}^n$, with \top denoting the transposition. In addition, $\boldsymbol{\theta} = (\theta^1, \dots, \theta^m)$ are the real-valued parameters that characterize the above probability distribution function in Eq. (42). Such parameters are the entries in the covariance matrix Σ . In this case, the Gaussian model in Eq. (42) happens to be an m -dimensional statistical model with $m = \frac{n(n+1)}{2}$. The parameter space Θ is the space of the variances and covariances of the above mentioned random variables; it is given by all $\boldsymbol{\theta} \in \mathbb{R}^n$ such that $\Sigma(\boldsymbol{\theta}) > 0$ (that is, Σ is a positive definite matrix). Now, the space Θ is endowed with the Fisher-Rao information metric tensor $g = g_{ij} d\theta^i \otimes d\theta^j$ whose components are defined in (5).

In this way, a configuration space is associated to each network. Such a space consists of a subset of the linear vector space \mathbb{R}^m given by the m -parameters that characterize the joint probability distribution of the random variables. In turn, this configuration space is endowed with the Fisher-Rao information metric. Then, in analogy to classical statistical mechanics, a Riemannian geometric entropy is defined as the logarithm of the volume of this manifold

$$\mathcal{S} \stackrel{\text{def}}{=} \ln \mathcal{V}_{\boldsymbol{\theta}}, \quad (43)$$

where $\mathcal{V}_{\boldsymbol{\theta}}$ is the Riemannian volume of the parameter space Θ .

The fact that this quantity was proposed in the context of networks inspired its analysis in the Erdős-Rényi phase transition in random graph [61]. A random graphs model $\mathbb{G}(n, k)$ is devised by choosing with uniform probability a graph from the set of all graphs having n vertices and k edges [64]. We can think of a type of dynamics by adding the edges one at a time. When k has the same order of magnitude as n , the evolution from $k = 0$ to $k = \binom{n}{k}$ yields, according to the Erdős-Rényi theorem [61], a phase transition, revealing itself in a rapid growth with k of the size of the largest component. Specifically, the structure of graphs 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 graphs rapidly grows from $\ln n$ to n , if k has the same order of magnitude of n .

In Ref. [62], the Riemannian geometric entropy $\mathcal{S}(k)$, defined in Eq. (43), has been numerically evaluated as a function of k for a fixed n , in order to investigate its sensitivity to the appearance of the giant component during the evolution of the random graph model $\mathbb{G}(n, k)$. Two main results were uncovered. First, in agreement with typical features that arise in numerical investigations of second order phase transitions, what asymptotically would be a sharp

bifurcation is rounded at finite n , in analogy to finite-size effects on the order parameter. Second, considering

$$\tilde{\mathcal{S}}(k) \stackrel{\text{def}}{=} \frac{1}{n} \langle \mathcal{S}(k) - \mathcal{S}(0) \rangle, \quad (44)$$

where $\langle \cdot \rangle$ denotes a Monte Carlo estimation of the average, it was observed in Ref. [62] that the larger n is, the more pronounced the “knee” of $\tilde{\mathcal{S}}(k/n)$ becomes. Interestingly, this finding is in agreement with an n -asymptotic bifurcation at $k/n = 0.5$ where the phase transition takes place.

This unprecedented result suggests to propose the Riemannian geometric entropy as a measure of networks complexity. To this end let us recall that the “state manifold” $\mathcal{M} \stackrel{\text{def}}{=} (\Theta, g)$ associated with a given network of n nodes has, in principle, dimension $m = \frac{n(n+1)}{2}$. Then, in order to simplify the computation of components of the Fisher-Rao information metric, we define a new Riemannian metric which accounts also for the network structure given by the adjacency matrix [62]. We first consider a trivial network with null adjacency matrix and interpret n independent Gaussian variables X_1, \dots, X_n as sitting on its vertices. It is clear that the joint probability distribution is now given by (42) with covariance matrix as $\Sigma_0 = \text{diag}[\theta^n, \dots, \theta^n]$, where $\theta^i = \mathbb{E}[X_i^2]$ for all $i : 1, \dots, n$. In this case, the Riemannian manifold $\mathcal{M}_0 \stackrel{\text{def}}{=} (\Theta_0, g_0)$ associated with the bare network is given by [63]

$$\begin{aligned} \Theta_0 &\stackrel{\text{def}}{=} \{\boldsymbol{\theta} = (\theta^1, \dots, \theta^n) \mid \theta^i > 0\}, \\ g_0 &\stackrel{\text{def}}{=} \frac{1}{2} \sum_{i=1}^n \left(\frac{1}{\theta^i} \right)^2 d\theta^i \otimes d\theta^i. \end{aligned} \quad (45)$$

From (45), it is clear that a functional relation holds between the components of the metric g_0 and the entries of the matrix Σ_0 . This is given by $g_{ii} = \frac{1}{2} (\sigma_{ii}^{-1})^2$ where σ_{ii}^{-1} is the ii -th entry of the inverse matrix of Σ_0 , i.e. $\sigma_{ii}^{-1} = \frac{1}{\theta^i}$. Inspired by this functional relation, we associate a Riemannian manifold to any network \mathcal{X} with non vanishing adjacency matrix A . To this aim, consider the map $\psi_{\boldsymbol{\theta}}(A) : \mathbf{A}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R})$ defined by the relation

$$\psi_{\boldsymbol{\theta}}(A) \stackrel{\text{def}}{=} \Sigma_0(\boldsymbol{\theta}) + A, \quad (46)$$

where $\mathbf{A}(n, \mathbb{R})$ denotes the set of symmetric $n \times n$ matrices over \mathbb{R} with vanishing diagonal elements that can represent any simple undirected graph. Here, $\text{GL}(n, \mathbb{R})$ is the general linear group of matrices. Then, we deform the manifold \mathcal{M} via $\psi_{\boldsymbol{\theta}}(A)$. Hence the manifold associated to a network \mathcal{X} with adjacency matrix A is $\tilde{\mathcal{M}} \stackrel{\text{def}}{=} (\tilde{\Theta}, \tilde{g})$ where,

$$\tilde{\Theta} \stackrel{\text{def}}{=} \{\boldsymbol{\theta} \in \mathbb{R} \mid \psi_{\boldsymbol{\theta}}(A) \text{ is positive definite}\}, \quad (47)$$

and $\tilde{g} \stackrel{\text{def}}{=} \tilde{g}_{ij} d\theta^i \otimes d\theta^j$. The components of \tilde{g} are given by,

$$\tilde{g}_{ij} \stackrel{\text{def}}{=} \frac{1}{2} (\psi_{\boldsymbol{\theta}}^{-1}(A)_{ij})^2, \quad (48)$$

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

In this way, a differentiable system (Riemannian manifold) is associated to a discrete system (network) through the description of the network by a set of probability distribution functions. Now, having a phase space for describing the network structure, we supply the microcanonical definition of entropy in statistical mechanics in order to introduce a geometric entropy of a network \mathcal{X} with adjacency matrix A . It is given by

$$\mathcal{S} \stackrel{\text{def}}{=} \ln \mathcal{V}(A), \quad (49)$$

where $\mathcal{V}(A)$ is the volume of $\widetilde{\mathcal{M}}$ computed from the volume element $\nu_{\widetilde{g}}$ defined as,

$$\nu_{\widetilde{g}} \stackrel{\text{def}}{=} \sqrt{\det \widetilde{g}} d\theta^1 \wedge \dots \wedge d\theta^n, \quad (50)$$

where “ \wedge ” in Eq. (50) denotes the wedge product of n 1-forms. The volume $\mathcal{V}(A)$ in Eq. (49) however, is ill-defined. This happens because $\widetilde{\Theta}$ is not compact since the variables θ^i are unbound from above. In addition, $\det [\widetilde{g}(\boldsymbol{\theta})]$ can be zero for some θ^i s and this causes the volume element to diverge as it is clear from Eq. (48). Thus, as is customary in the literature [65], we regularize $\mathcal{V}(A)$ as follows

$$\mathcal{V}(A) = \int_{\widetilde{\Theta}} \mathcal{R}(\psi_{\boldsymbol{\theta}}(A)) \nu_{\widetilde{g}}, \quad (51)$$

where $\mathcal{R}(\psi_{\boldsymbol{\theta}}(A))$ is any suitable regularizing function. This procedure is essentially the compactification of the parameter space where we exclude those sets of θ^i values which make $\det [\widetilde{g}(\boldsymbol{\theta})]$ divergent.

Theoretically, a regularizing function $\mathcal{R}(\psi_{\boldsymbol{\theta}}(A))$ could be devised by taking into account the probabilistic nature of the model description. Then, in order to make a natural choice for $\mathcal{R}(\psi_{\boldsymbol{\theta}}(A))$, consider the Shannon entropy H_{Shannon} of a multivariate normal distribution as the one in Eq.(42),

$$H_{\text{Shannon}} = \ln \sqrt{(2\pi e)^n \det \Sigma(\boldsymbol{\theta})}. \quad (52)$$

Interpreting the Shannon entropy as a measure of the lack of information about a system, to a smaller variance of a Gaussian pdf there corresponds a smaller entropy. In particular, in the limiting case in which $\det \Sigma(\theta)$ approaches zero for some θ^i , we have the minimum entropy and, consequently, the maximum information about a system. In such a case, it is clearly meaningless to consider a complexity measure. Similarly, if θ^i approaches infinity, the Shannon entropy diverges and there is a lack of information. For these reasons, it appears quite natural to choose the

regularizing function $\mathcal{R}(\psi_{\boldsymbol{\theta}}(A))$ as follows,

$$\begin{aligned} \mathcal{R}(\psi_{\boldsymbol{\theta}}(A)) \stackrel{\text{def}}{=} & \left[H(N - \text{tr}\psi_{\boldsymbol{\theta}}(A)) + H(\text{tr}\psi_{\boldsymbol{\theta}}(A) - N)e^{-\text{tr}\psi_{\boldsymbol{\theta}}(A)} \right] \times \\ & \times [H(\varepsilon - \det \psi_{\boldsymbol{\theta}}(A)) \ln [1 + (\det \psi_{\boldsymbol{\theta}}(A))^n] + H(\det \psi_{\boldsymbol{\theta}}(A) - \varepsilon)], \end{aligned} \quad (53)$$

where $H(\cdot)$ is the Heaviside step function, N denotes a positive real number with $\mathcal{R}(\psi_{\boldsymbol{\theta}}(A)) \sim O(\text{tr}\psi_{\boldsymbol{\theta}}(A))$ for all $\boldsymbol{\theta}$ such that $\text{tr}\psi_{\boldsymbol{\theta}}(A) \geq N$, and ε is a positive *real* number such that $\mathcal{R}(\psi_{\boldsymbol{\theta}}(A)) \sim o(\det \psi_{\boldsymbol{\theta}}(A))$ for all $\boldsymbol{\theta}$ such that $\det \psi_{\boldsymbol{\theta}}(A) \leq \varepsilon$. It happens that the volume $\mathcal{V}(A)$ in Eq. (51) is well-defined [66].

The geometric entropy (51) is used to single out emergent phenomena occurring in power-law random graphs, in addition to uniform random graphs. The former are 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 $\mathbb{G}_{\gamma,d}$ [67], assigns a uniform probability to all graphs with $n = e^{\alpha}d^{-\gamma}$ where e represents the Euler number. Now, according to the n -asymptotic bifurcation at $\gamma = \gamma_c = 3.47875$ predicted by the Molloy-Reed criterion for the emergence of a giant component in $\mathbb{G}_{\gamma,d}$ [68], the geometric entropy (51) displays the typical phenomenon uncovered in numerical investigations of second-order phase transitions when γ is in the range $1.5 < \gamma < 5.5$ [42].

In the transition from random graphs to real networks, the geometric entropy so far put forward has been considered against some small exponential random graphs which are evaluated by (65) 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 [43]. In Ref. [43], the convex hull of all possible expectation values of the probabilities of the triangles and 3-chains is derived. Furthermore, graphs that correspond to the minimal energy are found to lie on the lower boundary of the mentioned convex hull. The geometric entropy (51) suggests that going up along the lower boundary of this convex hull, the degree of complexity increases [42]. However, for the family of graphs which are intended of maximal energy in Ref. [43], it happens that network complexity is nontrivially influenced by network topology (homology) as was first pointed out in Ref. [63].

An additional step towards analyzing real networks is to consider d -regular graphs [64]. They are networks where each node has the same degree d . A way to generate these special networks is the *configuration model* [69] that is specified in terms of a sequence of degrees $d = (d_1, \dots, d_n)$. 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 [69]. According to the geometric entropy, the larger d , the more complex the network is. Similarly, when we consider networks with one or more hubs, i.e. nodes with degree larger than the others, the

complexity of a network increases with the number of hubs in it [42]. Now, d -regular graphs show homogeneity in the interaction structures entailing topological equivalence of almost all the nodes. On the contrary, most real networks have power law degree distribution $p(d) = \mathcal{A} d^{-\gamma}$, where \mathcal{A} is a positive *real* constant and the exponent γ varies in the range $2 \leq \gamma \leq 3$ [67], and are assessed with high heterogeneity, that is a measure of how far from $p(d_0) = 1$, $p(d) = 1(d_0 \neq d)$ a network is [70]. The degree of heterogeneity h of a network has been quantified in [71] by means of the following definition,

$$h \stackrel{\text{def}}{=} \sum_{i,j \in E} \left(\frac{1}{\sqrt{d_i}} - \frac{1}{\sqrt{d_j}} \right), \quad (54)$$

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. Interestingly, the degree of heterogeneity of some networks as measured by (51) is the same as the ordering produced by the h measure of heterogeneity [42].

Finally, the geometric entropy has been applied to some real networks where its predicted features have been compared with those of three other measures of complexity: one describing modularity structure, one homogeneity, and, finally, one characterizing the heterogeneity of real networks [72]. The findings uncovered in these investigations validate the meaningfulness and the effectiveness of the geometric entropy as a good measure of network complexity [42].

V. STOCHASTIC INTERACTIONS

Other complexity measures can be placed in a geometric framework. When considering the statistical (or, structural) complexity [73] of describing a system made up by an ensemble of interacting units, a reasonable feature of complex systems is the emergence of new structural properties at different (more detailed) levels of description. Starting from this observation, Ay and coworkers exploited the main results presented in Ref. [74, 75] (hierarchy of probability distributions) and [76] (hierarchical graphical models) in order to place well-known complexity measures [43, 77, 78] in a unifying information geometric setting. Within this IG framework, the idea of hierarchical structure of probability distributions is used to characterize the interactions among the units of a complex system at different scales of description by means of an explicit orthogonal decomposition of the stochastic dependency with respect to an interaction hierarchy (or, order) [45]. In Ref. [79], Ay and coworkers introduced a notion of complexity based on the interaction among parts of a system that rigorously incorporates the idea to quantify the amount of pairwise or triplewise interacting components and beyond, by using the exponential families of k -interactions and the notion of

distance to such an exponential family. This leads to quantify complexity in terms of the degree of k -wise interaction that cannot be explained by a $(k-1)$ -wise interaction. For further details on this approach by Ay and coworkers, we refer to Refs. [80, 81]. In particular, for an investigation on the effects of both spatial and temporal interdependencies in complex systems of stochastically interacting units, we refer to Ref. [81].

In what follows, we shall present the essential tools needed to understand the geometric complexity measure introduced by Ay and coworkers.

A. Exponential families of k -interactions

The set of states of a system is given in terms of compositional structure, i.e. for each site index v in a set $V \stackrel{\text{def}}{=} \{1, \dots, N\}$ we have a finite configuration space \mathcal{X}_v . The set of all possible configuration is $\mathcal{X}_V \stackrel{\text{def}}{=} \times_{v \in V} \mathcal{X}_v$ and likewise $\mathcal{X}_A \stackrel{\text{def}}{=} \times_{v \in A} \mathcal{X}_v$ for each subset $A \subset V$. Given the set of real-valued functions on \mathcal{X}_v , $\mathbb{R}^{\mathcal{X}_v} \stackrel{\text{def}}{=} \{f \mid f : \mathcal{X}_v \rightarrow \mathbb{R}\}$, elements of the subset $\mathcal{F}_{\mathcal{X}_V}$,

$$\mathcal{F}_{\mathcal{X}_V} \stackrel{\text{def}}{=} \left\{ p \in \mathbb{R}^{\mathcal{X}_V} : p(\mathbf{x}) \geq 0, \sum_{\mathbf{x} \in \mathcal{X}_V} p(\mathbf{x}) = 1 \right\}, \quad (55)$$

are probability measures on \mathcal{X}_V . Indeed, due to the compositional structure, these probability measures are only joint probabilities of a set of random variables $\{X_v \mid v \in V\}$, where X_v takes values in \mathcal{X}_v . Observe that from a geometrical viewpoint, $\mathcal{F}_{\mathcal{X}_V}$ in Eq. (55) can be regarded as a $(|\mathcal{X}_V| - 1)$ -dimensional probability simplex. Now, the support of a probability distribution p is given by,

$$\text{supp}(p) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathcal{X}_V : p(\mathbf{x}) > 0\}. \quad (56)$$

and probability distributions with full support are denoted as $\mathcal{P}(\mathcal{X}_V)$,

$$\mathcal{P}(\mathcal{X}_V) \stackrel{\text{def}}{=} \{p \in \mathcal{F}_{\mathcal{X}_V} : p(\mathbf{x}) > 0, \forall \mathbf{x} \in \mathcal{X}_V\}. \quad (57)$$

In order to define useful exponential families for quantifying the amount of k -interactions, we consider the linear space of functions depending on only k of their arguments,

$$\mathcal{I}_k \stackrel{\text{def}}{=} \text{span}_{A \subseteq V, |A|=k} (\mathcal{I}_A), \quad (58)$$

where \mathcal{I}_A is a subspace of $\mathbb{R}^{\mathcal{X}_V}$ formed by all the functions that do not depend on the configurations outside A ,

$$\mathcal{I}_A \stackrel{\text{def}}{=} \{f \in \mathbb{R}^{\mathcal{X}_V} : f(\mathbf{x}_A, \mathbf{x}_{V \setminus A}) = f(\mathbf{x}_A, \tilde{\mathbf{x}}_{V \setminus A})\}, \quad (59)$$

for all $\mathbf{x}_A \in \mathcal{X}_A$, $\mathbf{x}_{V \setminus A}$ and $\tilde{\mathbf{x}} \in \mathcal{X}_{V \setminus A}$. After taking the exponential map of sums of such functions, we find the probability distributions with only k interactions $\mathcal{E}_k \stackrel{\text{def}}{=} \exp(\mathcal{I}_k)$, where

$$\exp : \mathbb{R}^{\mathcal{X}_V} \ni f(x) \mapsto \frac{\exp[f(x)]}{\sum_{\mathbf{x} \in \mathcal{X}_V} \exp[f(\mathbf{x})]} \in \mathcal{P}(\mathcal{X}_V). \quad (60)$$

Proceeding along these lines of reasoning, a hierarchy of exponential families can be defined in terms of the following chain of inclusions,

$$\mathcal{E}_1 \subseteq \mathcal{E}_2 \dots \subseteq \mathcal{E}_k. \quad (61)$$

This hierarchy was widely studied in Refs. [74, 82] where various applications in the theory of neural networks were considered. The connection between this formalism and statistical physics relies upon the following correspondence: $f \in \mathcal{I}_k$ corresponds to an interaction energy which has only k interactions, but no higher interactions. This fact gives rise to a probability distribution $p(\mathbf{x})$,

$$p(\mathbf{x}) \stackrel{\text{def}}{=} \frac{\exp[f(\mathbf{x})]}{\mathcal{Z}} \in \mathcal{E}_k, \quad (62)$$

where \mathcal{Z} denotes the partition function of the system and it is a normalization factor. Let us note that, in general, the image of the exponential map is given by distributions with full support. Therefore, probability zero, which corresponds to infinite energy, is achievable only by limits of sequences of probability measures in the exponential family. This is not a serious drawback for considering the Kullback-Leibler divergence for introducing a complexity measure based on the notion of distance to an exponential family \mathcal{E} .

B. Relative entropy and interaction structures

According to Eq. (22) the Kullback-Leibler divergence (or, relative entropy) $\mathcal{D}(p||q)$ describes the distance between two probability distributions p and q in $\mathcal{F}_{\mathcal{X}_V}$ even though it is not a metric; it is not a metric since it is not symmetric and does not satisfy the triangle inequality. As a consequence, one can consider the distance to an exponential family \mathcal{E} defined in terms of the Kullback-Leibler divergence as,

$$\mathcal{D}(p||\mathcal{E}) \stackrel{\text{def}}{=} \inf_{q \in \mathcal{E}} \mathcal{D}(p||q), \quad (63)$$

which satisfies the following relation,

$$\mathcal{D}(p||\mathcal{E}) - \mathcal{D}(p||\mathcal{F}) \geq 0 \quad (64)$$

whenever $\mathcal{E} \subseteq \mathcal{F}$. At this point, we are ready to describe the complexity measure proposed by Ay and coworkers. This measure is a vector-valued quantity $I(p) \stackrel{\text{def}}{=} (I^{(1)}(p), \dots, I^{(N)}(p))$ with,

$$I^{(k)}(p) \stackrel{\text{def}}{=} \mathcal{D}(p||\mathcal{E}_{k-1}) - \mathcal{D}(p||\mathcal{E}_k), \quad (65)$$

for any $1 \leq k \leq N$. The component $I^{(k)}(p)$ is a quantitative measure of the dependencies among k nodes that fail to be captured by interaction among smaller subsets of nodes. An explicit illustrative example of this nice theoretical approach to complexity appears in Ref. [79] where, by considering a one-dimensional binary cellular automata and the symbolic dynamics of a coupled tent map lattice [83, 84], a numerical analysis of such information geometric complexity measures was presented. In particular, for elementary binary cellular automata [85] with rules uniformly distributed, no interaction can be detected. On the contrary, the complexity measure in the Eq. (65) is able to detect long-range correlations for rules that do not have the uniform distribution as their invariant measure. As far as the coupled tent maps are concerned, the fully connected graph and the circle graph are considered. For $k > 6$, the $I^{(k)}$ quantities are very small and depend on the random initial condition, so they are neglected. For the fully connected graphs, the $I^{(k)}$ suggest that complex dynamics takes place on the edge of synchronized regimes. Furthermore, interesting results are uncovered when the simulation of the symbolic dynamics of a coupled tent map lattice is performed by means of a circle graph specified by a next neighbor coupling. Specifically, when two nodes of the lattice are constant and four of them are periodic, all $I^{(k)}$ with $k \neq 2$ are zero. This is a signature of a high degree of regularity. However, when the average activity does not follow one of the individual nodes and a complex dynamical structure driven by the higher-order correlations among the nodes exists, very high values of $I^{(k)}$ with $k \neq 2$ are observed.

Relying on the correspondence with statistical physics described by Eq. (62), the complexity measure (65) is also able to quantify structure in networks [86]. Indeed, whenever the function $f(\mathbf{x})$ is an Hamiltonian of the observables of the system, the exponential family happens to be the collection of maximum entropy distributions for fixed expectation values of these observables. The counts of subgraphs with no more than k links is a sufficient statistic for hierarchically structured exponential families. Hence, within this theoretical framework, the study of the dependencies among several observables such as the degree distribution, cluster, and assortativity coefficients, leads to quantifiable structures in networks.

VI. ENTROPIC MOTION

We now move to temporal complexity arising in entropic inference processes. In recent years, entropic inference methods [87] have been combined with information geometric techniques [5] in order to quantify the complexity of statistical models used to provide probabilistic descriptions of physical systems in the presence of limited information. Within this context, the complexity of statistical models is regarded as the difficulty of inferring macroscopic predictions in the absence of detailed knowledge about the microscopic nature of the system being investigated. This fascinating line of research, originally called *Information Geometric Approach to Chaos* (IGAC), was presented by Cafaro in his Ph. D. doctoral dissertation in Ref. [88]. The essence of the IGAC theoretical construct can be described as follows: given an arbitrary complex system, once the microscopic degrees of freedom of the system are identified and the relevant information constraints about them are selected, entropic methods are used to provide a preliminary static statistical model of the complex system. The statistical model is defined by probability distributions parametrized in terms of statistical macrovariables which, in turn, depend on the particular functional form of the information constraints that have been assumed to be relevant for carrying out the statistical inferences. Then, assuming that the complex system evolves, the evolution of the corresponding statistical model from an initial to a final configuration is described by means of the so-called Entropic Dynamics (ED [89]). The ED is a form of information-constrained dynamics built on curved statistical manifolds where elements of the manifold are probability distributions that are in a one-to-one relation with a suitable set of statistical macrovariables that define the so-called parameter space and provide a convenient parametrization of points on the statistical manifold. Within the ED framework, the evolution of probability distributions is governed by an entropic inference principle: starting from the initial configuration, motion towards the final configuration occurs by maximizing the logarithmic relative entropy (Maximum relative Entropy method- MrE method, [87]) between any two consecutive intermediate configurations. It is worthwhile mentioning that ED only provides the *expected* and not the *actual* trajectories of the system [90]. Inferences in ED rely on the nature of the chosen information constraints employed in the MrE algorithm. The validation of this type of modeling scheme can be checked only *a posteriori*. If discrepancies between the experimental observations and the inferred predictions occur, a new set of information constraints needs to be selected. The evolution of probability distributions is described in terms of the outcome of this maximization procedure, namely a geodesic evolution of the statistical macrovariables [87]. The distance between two different probability distributions can be viewed as a measure of the distinguishability between the two distributions and is quantified, as mentioned in Section II, in terms of the so-called Fisher-Rao information metric [5]. Given the information metric, the evolution arising from

the maximization procedure results in geodesics with respect to the Levi-Civita (or Riemannian) connection $\bar{\nabla}$ of the Fisher-Rao metric. Now, the Riemannian connection $\bar{\nabla}$ gives rise to a self dual structure of the statistical model meaning that $\nabla^* = \nabla$. Within the IGAC approach to complexity, the complete geometric structure supplying the entropic dynamics is given in terms of the Fisher-Rao metric and the Riemannian connection $\bar{\nabla}$ due to minimization requirements on the Lagrangian metric functional [88]. For this reason, within the present section, we refer to statistical models endowed with a Riemannian geometric structure. In this way one can apply standard methods of Riemannian differential geometry to study the geometric structure of the manifold underlying the entropic motion that characterizes the evolution of probability distributions. Generally speaking, conventional Riemannian geometric quantities such as Christoffel connection coefficients, Ricci tensor, Riemannian curvature tensor, sectional curvatures, Ricci scalar curvature, Weyl anisotropy tensor, Killing fields, and Jacobi fields can be computed in the conventional manner. More specifically, the chaoticity (or, in other words, the temporal complexity) of such statistical models can be investigated using suitably chosen measures such as the signs of the scalar and sectional curvatures of the statistical manifold underlying the entropic motion, the asymptotic temporal behavior of Jacobi fields, the IG analogue of Lyapunov exponents, the existence of Killing vectors, and the existence of a non-vanishing Weyl anisotropy tensor. In addition to the above mentioned ordinary measures of complexity imported into the information geometric framework from the ordinary Riemannian geometry platform, complexity within the IGAC approach can also be characterized in terms of the so-called information geometric entropy (IGE), originally presented in [88].

A. Information geometric entropy

In what follows, we describe the concept of the IGE within the IGAC framework. Let us assume that the points $\{p(\mathbf{x}; \boldsymbol{\theta})\}$ of an n -dimensional curved statistical manifold \mathcal{M}_s are parametrized by means of n *real* valued variables $(\theta^1, \dots, \theta^n)$,

$$\mathcal{M}_s \stackrel{\text{def}}{=} \left\{ p(\mathbf{x}; \boldsymbol{\theta}) : \boldsymbol{\theta} = (\theta^1, \dots, \theta^n) \in \mathcal{D}_{\boldsymbol{\theta}}^{(\text{tot})} \right\}. \quad (66)$$

The microvariables \mathbf{x} belong to the microspace \mathcal{X} while the macrovariables $\boldsymbol{\theta}$ belong to the parameter space $\mathcal{D}_{\boldsymbol{\theta}}^{(\text{tot})}$ defined as,

$$\mathcal{D}_{\boldsymbol{\theta}}^{(\text{tot})} \stackrel{\text{def}}{=} \bigotimes_{k=1}^n \mathcal{I}_{\theta^k} = (\mathcal{I}_{\theta^1} \otimes \mathcal{I}_{\theta^2} \dots \otimes \mathcal{I}_{\theta^n}) \subseteq \mathbb{R}^n. \quad (67)$$

The quantity \mathcal{I}_{θ^k} with $1 \leq k \leq n$ in Eq. (67) is a subset of \mathbb{R} and represents the entire range of allowable values for the macrovariables θ^k . The information geometric entropy (IGE, [88, 91–93]) is a proposed measure of temporal

complexity of geodesic paths within the IGAC framework and is defined as,

$$\mathcal{S}_{\mathcal{M}_s}(\tau) \stackrel{\text{def}}{=} \ln \widetilde{\text{vol}}[\mathcal{D}_{\boldsymbol{\theta}}(\tau)], \quad (68)$$

where the average dynamical statistical volume $\widetilde{\text{vol}}[\mathcal{D}_{\boldsymbol{\theta}}(\tau)]$ is given by,

$$\widetilde{\text{vol}}[\mathcal{D}_{\boldsymbol{\theta}}(\tau)] \stackrel{\text{def}}{=} \frac{1}{\tau} \int_0^\tau \text{vol}[\mathcal{D}_{\boldsymbol{\theta}}(\tau')] d\tau'. \quad (69)$$

Observe that the tilde symbol in Eq. (69) denotes the operation of temporal average. The volume $\text{vol}[\mathcal{D}_{\boldsymbol{\theta}}(\tau')]$ in the RHS of Eq. (69) is defined as,

$$\text{vol}[\mathcal{D}_{\boldsymbol{\theta}}(\tau')] \stackrel{\text{def}}{=} \int_{\mathcal{D}_{\boldsymbol{\theta}}(\tau')} \rho(\theta^1, \dots, \theta^n) d^n \boldsymbol{\theta}, \quad (70)$$

where $\rho(\theta^1, \dots, \theta^n)$ is the so-called Fisher density and is equal to the square root of the determinant $g(\boldsymbol{\theta})$ of the Fisher-Rao information metric tensor $g_{\mu\nu}(\boldsymbol{\theta})$,

$$\rho(\theta^1, \dots, \theta^n) \stackrel{\text{def}}{=} \sqrt{g(\boldsymbol{\theta})}. \quad (71)$$

We remark that the expression in Eq. (70) can be made more transparent for statistical manifolds with information metric tensor whose determinant can be factorized as follows,

$$g(\boldsymbol{\theta}) = \prod_{i=1}^n g_i(\theta^i). \quad (72)$$

Indeed, in such cases, the IGE in Eq. (68) becomes

$$\mathcal{S}_{\mathcal{M}_s}(\tau) = \ln \left\{ \frac{1}{\tau} \int_0^\tau \left[\prod_{i=1}^n \left(\int_{\tau_0}^{\tau_0+\tau'} \sqrt{g_i[\theta^i(\alpha)]} \frac{d\theta^i}{d\alpha} d\alpha \right) \right] d\tau' \right\}. \quad (73)$$

Observe that within the IGAC framework, the asymptotic behavior of $\mathcal{S}_{\mathcal{M}_s}(\tau)$ is used to characterize the complexity of the statistical models being analyzed. For this reason, one usually considers the quantity $\mathcal{S}_{\mathcal{M}_s}^{(\text{asymptotic})}(\tau)$,

$$\mathcal{S}_{\mathcal{M}_s}^{(\text{asymptotic})}(\tau) \approx \lim_{\tau \rightarrow \infty} [\mathcal{S}_{\mathcal{M}_s}(\tau)], \quad (74)$$

the leading asymptotic term in the IGE. The integration space $\mathcal{D}_{\boldsymbol{\theta}}(\tau')$ in Eq. (70) is given by,

$$\mathcal{D}_{\boldsymbol{\theta}}(\tau') \stackrel{\text{def}}{=} \{\boldsymbol{\theta} : \theta^k(\tau_0) \leq \theta^k \leq \theta^k(\tau_0 + \tau')\}, \quad (75)$$

where $\theta^k = \theta^k(\alpha)$ with $\tau_0 \leq \alpha \leq \tau_0 + \tau'$ and τ_0 denoting the initial value of the affine parameter α such that [Appendix A],

$$\frac{d^2 \theta^k(\alpha)}{d\alpha^2} + \bar{\Gamma}_{lm}^k \frac{d\theta^l}{d\alpha} \frac{d\theta^m}{d\alpha} = 0, \quad (76)$$

with $\bar{\Gamma}_{lm}^k$ denoting the Christoffel symbols of the Riemannian connection $\bar{\nabla}$. The integration space $\mathcal{D}_{\theta}(\tau')$ in Eq. (75) is an n -dimensional subspace of $\mathcal{D}_{\theta}^{(\text{tot})}$ whose elements are n -dimensional macrovariables $\{\theta\}$ whose components θ^k are bounded by specified limits of integration $\theta^k(\tau_0)$ and $\theta^k(\tau_0 + \tau')$. The integration of the n -dimensional set of coupled nonlinear second order ordinary differential equations in Eq. (76) determines the (temporal) functional form of such limits.

The IGE in Eq. (68) at a given instant is essentially the logarithm of the volume of the effective parameter space explored by the system at that very instant. The temporal average has been introduced in order to average out the possibly very complex fine details of the entropic dynamical description of the system on the curved statistical manifold. Furthermore, the long-term asymptotic temporal behavior is considered in order to characterize in a proper fashion the selected dynamical indicators of chaoticity, including Lyapunov exponents and entropic quantities, by eliminating the effects of transient effects which enter the computation of the expected value of the volume of the effective parameter space. Therefore, the IGE is built to provide an asymptotic coarse-grained inferential description of the complex dynamics of a system in the presence of limited information. The construction and interpretation of the IGE shares similarities with the logical and thermodynamic depths [38, 39]. The logical depth [38] is considered to be a good measure of statistical complexity [24] where the correlated structure of the system's constituents is essential to determine the complex path connecting the initial and final states of the system being considered. Specifically, it is a time measure of complexity and represents the run time needed by a universal Turing machine executing the minimal program to reproduce a given pattern. The run time is obtained by a suitable averaging procedure over the various programs that will accomplish the task by weighting shorter programs more heavily. Therefore, the logical depth of any system is defined if a suitably coarse-grained description of it is encoded into a bit string. On the other hand, the thermodynamic depth of a process is a structural measure of complexity and it represents the difference between the system's coarse- and fine-grained entropy. The depth of a macrostate reached by a particular trajectory is proportional to $-\ln p_j$ where p_j is the probability of j -th trajectory. The set $\{p_j\}$ represents probabilities which are consistent with all the measurements that have been made on the system during its history. The arbitrariness and lack of explanation of how the macrostates of the system leading to the formation of the path-trajectory are selected [94] is an important objection to the thermodynamic depth. Within the IGE construction, the selection of explored macrostates by the system occurs in a manner as objective as possible since it relies on the universal MrE updating method [95]. The MrE method of determining macroscopic paths makes no mention of randomness or other incalculable quantities. It simply chooses the distribution (that is, the macrostate) with the maximum entropy

allowed by the information constraints. Thus, it selects the most uninformative distribution of microstates possible. If we chose a probability distribution with lower entropy, then we would assume information we do not possess. In addition, to choose one with a higher entropy would violate the constraints of the information we do possess.

The effectiveness of the IGAC theoretical approach so far described has been proved also for particular subsets of the exponential family Eq. (12), namely Gaussian statistical manifolds. Using the notation used in Section II, an n -dimensional multivariate Gaussian statistical model is defined in terms of the following relations,

$$\begin{aligned}\mathcal{X} &\stackrel{\text{def}}{=} \mathbb{R}^k, \quad n \stackrel{\text{def}}{=} k + \frac{k(k+1)}{2}, \quad \boldsymbol{\theta} \stackrel{\text{def}}{=} [\mu, \Sigma], \\ \Theta &\stackrel{\text{def}}{=} \{[\mu, \Sigma] \mid \mu \in \mathbb{R}^k, \Sigma \in \mathbb{R}^{k \times k} : \text{positive definite}\}, \\ p(\mathbf{x}; \boldsymbol{\theta}) &\stackrel{\text{def}}{=} \frac{\exp \left[-\frac{1}{2} (\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu) \right]}{[(2\pi)^k \det \Sigma]^{1/2}}.\end{aligned}\tag{77}$$

In Eq. (77), μ is the k -dimensional mean vector while Σ denotes the $k \times k$ (symmetric) covariance matrix. Specifically, in Refs. [91, 96], the IGAC is used to investigate the geometry and the entropic dynamics of an uncorrelated Gaussian model with l degrees of freedom, each one characterized by two pieces of relevant information, its mean and its variance. It is observed that the scalar curvature of the $2l$ dimensional statistical manifold is proportional to the number of degrees of freedom of the system, $\mathcal{R}_{\mathcal{M}_s} = -l$. Similarly, from a dynamical standpoint, it happens for the IGE in (68),

$$\mathcal{S}_{\mathcal{M}_s}(\tau) \stackrel{\tau \rightarrow \infty}{\sim} l\lambda\tau,\tag{78}$$

where λ denotes the maximum positive Lyapunov exponent that characterizes the statistical model. Finally, by integrating the geodesic deviation equations, it is found that, in the asymptotic limit, the Jacobi vector field intensity $J_{\mathcal{M}_s}$ exhibits exponential divergence and is proportional to the number of degrees of freedom l , $J_{\mathcal{M}_s}(\tau) \stackrel{\tau \rightarrow \infty}{\sim} l \exp(\lambda\tau)$. Observing that the exponential divergence of the Jacobi vector field intensity $J_{\mathcal{M}_s}$ is a classical feature of chaos, it was argued that $\mathcal{R}_{\mathcal{M}_s}$, $\mathcal{S}_{\mathcal{M}_s}$ and $J_{\mathcal{M}_s}$ behave as proper indicators of chaoticity and are proportional to the number of Gaussian-distributed microstates of the system. This proportionality was the first example in the literature of a possible substantial link among information geometric indicators of chaoticity for probabilistic descriptions of dynamical systems.

Further investigation of the IGAC is carried out for correlated Gaussian statistical models in [97] to analyze the information constrained dynamics of a system with two microscopic degrees of freedom. These degrees of freedom are

assumed to be represented by two correlated Gaussian-distributed microvariables characterized by the same variance. The presence of correlations at the microscopic level leads to the emergence of an asymptotic information geometric compression of the statistical macrostates explored by the system at a faster rate than that observed in absence of correlations. In particular, it is found that in the asymptotic limit [97],

$$[\exp(\mathcal{S}_{\mathcal{M}_s}(\tau))]_{\text{correlated}} \stackrel{\tau \rightarrow \infty}{\sim} f(r) [\exp(\mathcal{S}_{\mathcal{M}_s}(\tau))]_{\text{uncorrelated}}, \quad (79)$$

where $f(r)$, with $0 \leq f(r) \leq 1$, is a monotonic decreasing compression factor for any value of the correlation coefficient r in the open interval $(0, 1)$. This finding represents an important and explicit connection between correlations at the *microscopic* level and complexity at the *macroscopic* level in the probabilistic description of dynamical systems within the IGAC theoretical approach. The importance of this result is twofold: first, it provides a clear description of the effect of information encoded in microscopic variables on experimentally observable quantities defined in terms of macroscopic variables; second, it neatly exhibits the behavioral change of the macroscopic complexity of a statistical model caused by the presence of correlations in the underlying microscopic level.

In Ref. [98], bivariate and trivariate Gaussian statistical models with different correlational structures were considered and their effect on the asymptotic behavior of the IGE (68) was observed. Essentially, it was reported that the complexity of entropic inferences not only depends on the amount of available pieces of information but also on the manner in which such pieces are correlated. In particular, for a trivariate statistical model with only two correlated degrees of freedom, the asymptotic temporal behavior of the information geometric complexity ratio between the correlated and uncorrelated cases exhibits a non monotonic behavior in terms of the correlation parameter r assuming a value equal to zero at the extrema of the allowed range of r ,

$$\frac{(\exp[\mathcal{S}_{\mathcal{M}_s}^{(\text{mildly connected})}(\tau)])_{\text{correlated}}}{(\exp[\mathcal{S}_{\mathcal{M}_s}^{(\text{mildly connected})}(\tau)])_{\text{uncorrelated}}} \stackrel{\tau \rightarrow \infty}{\sim} \sqrt{\frac{3(1-2r^2)}{3-4r}}. \quad (80)$$

However, for closed configurations (that is to say, bivariate and trivariate models with all microscopic variables correlated to each other) the complexity ratio exhibits a monotonic behavior in terms of the correlation parameter. For instance, in the fully-connected trivariate Gaussian case, it was found that

$$\frac{(\exp[\mathcal{S}_{\mathcal{M}_s}^{(\text{fully connected})}(\tau)])_{\text{correlated}}}{(\exp[\mathcal{S}_{\mathcal{M}_s}^{(\text{fully connected})}(\tau)])_{\text{uncorrelated}}} \stackrel{\tau \rightarrow \infty}{\sim} \sqrt{1+2r}, \quad (81)$$

with $0.5 < r < 1$. Note that a fully (mildly) connected lattice denotes a network with a higher (lower) connectivity structure. From Eq. (81), it is clear that in the fully connected trivariate case no peak arises and a monotonic behavior in r of the information geometric complexity ratio is observed. In the mildly connected trivariate case of

Eq. (80), instead, a peak in the information geometric complexity ratio is recorded at $r_{\text{peak}} = 0.5 \geq 0$. In analogy to the ideal scenario of minimum energy spin configurations in statistical physics [99, 100], within the IGE based approach, one would desire a configuration of minimum complexity in order to make reliable macroscopic predictions. Our findings in Eqs. (80) and (81) show a dramatically distinct behavior between the mildly connected and the fully connected trivariate Gaussian configurations. This behavior can be ascribed to the fact that when carrying out statistical inferences with positively correlated Gaussian random variables, the system appears frustrated in the fully connected case. This happens because the maximum entropy favorable scenario seems to be incompatible with the ideal scenario of minimum complexity. Just like certain spin configurations are not particularly favorable from an energy standpoint, certain lattice configurations in the presence of correlations are not especially favorable from a statistical inference perspective of minimum complexity. In conclusion, based upon these findings, it was argued in Ref. [98] that the impossibility of reaching the most favorable configuration for certain correlational structures from an entropic inference viewpoint leads to an information geometric analog of the frustration effect that occurs in statistical physics when loops are present [101].

For further details on the IGAC and its applications, we refer to [44].

VII. QUANTUM OUTLOOK

In this section, we consider the use of IG methods to characterize quantum PTs of second-order. We also briefly discuss how to capture the concept of complexity in quantum systems.

A. Phase transitions

As pointed out in Section III, within the classical information geometric setting, one of the main established findings concerns the fact that the Ricci scalar curvature \mathcal{R} is positive definite and diverges as ξ^d at the critical point of a second order PT. Information geometric methods can also be applied to quantum systems. Pioneering works along this line of investigation were carried out by Provost and Vallee in Ref. [102] where they introduced a Riemannian metric tensor from the Hilbert space structure of quantum states. The application of such theoretical research efforts in order to study zero-temperature second-order quantum phase transitions is more recent [103]. We remark that the IG of quantum PTs has different qualitative features with respect to those of the IG of classical PTs. For classical systems, the geometric parameters are thermodynamic variables and their Legendre transforms. Instead, for quantum systems,

geometric parameters are represented by quantities that appear in the Hamiltonian of the system, for instance, the coupling constants of the theory.

1. Curvature analysis

Let us consider a smooth family of Hamiltonians $\mathcal{H}(\lambda)$ in a Hilbert-space \mathbb{H} where $\lambda \in \mathcal{M}$ denotes a parameter that lives in the parameter manifold \mathcal{M} . Given, for the sake of simplicity, a unique ground-state $|\Psi_0(\lambda)\rangle \in \mathbb{H}$ for the Hamiltonian $\mathcal{H}(\lambda)$, we can define the map $\Psi_0 : \mathcal{M} \rightarrow \mathcal{H}$ in such a manner that it associates to each parameter value the ground-state of the corresponding Hamiltonian. Indeed, this map can be regarded as a map between \mathcal{M} and the projective space \mathbb{PH} . This latter space is a metric space equipped with the Fubini-Study distance

$$d_{\text{FS}}(\psi, \phi) \stackrel{\text{def}}{=} \arccos[\mathcal{F}(\psi, \phi)], \quad (82)$$

where $\mathcal{F}(\psi, \phi)$ is the fidelity between the quantum pure states $|\psi\rangle$ and $|\phi\rangle$,

$$\mathcal{F}(\psi, \phi) \stackrel{\text{def}}{=} |\langle\psi|\phi\rangle|. \quad (83)$$

In Ref. [104], it was noticed that $d_{\text{FS}}(\psi, \phi)$ in Eq. (82) is the maximum of the Fisher-Rao information metric distance between the probability distributions that arise from $|\psi\rangle$ and $|\phi\rangle$. Specifically, given a complete set of rank one projectors $\{|i\rangle\langle i|\}_{i=1}^{\dim \mathbb{H}}$, one can consider two probability distributions $p_i \stackrel{\text{def}}{=} |\langle i|\psi\rangle|^2$ and $q_i \stackrel{\text{def}}{=} |\langle i|\phi\rangle|^2$. In particular, for p and q infinitesimally close to each other, the Fisher-Rao information metric distance (that represents the distinguishability metric on the space of probability distributions) between p and q becomes [105],

$$ds_{\text{FR}}^2 \stackrel{\text{def}}{=} \sum_{i=1}^{\dim \mathbb{H}} \frac{dp_i^2}{p_i} = \sum_{i=1}^{\dim \mathbb{H}} p_i (d \ln p_i)^2. \quad (84)$$

The previously mentioned maximum of the Fisher-Rao information metric distance is computed over all possible projective measurements. In addition to being a metric space, the complex projective space \mathbb{PH} is also endowed with a Riemannian structure, that is, a metric tensor. Indeed, when the fidelity \mathcal{F} is very close to one, it can be shown that

$$d_{\text{FS}}^2(\psi, \psi + \delta\psi) \simeq 2(1 - \mathcal{F}). \quad (85)$$

In addition, since

$$\mathcal{F}(\psi, \psi + \delta\psi) \simeq |1 + \langle\psi|\delta\psi\rangle + (1/2)\langle\psi|\delta^2\psi\rangle|^2, \quad (86)$$

using this approximate expression for \mathcal{F} together with the normalization condition for the quantum state $|\psi\rangle$, one obtains

$$ds^2 \stackrel{\text{def}}{=} d_{\text{FS}}^2(\psi, \psi + \delta\psi) = \langle \delta\psi | \delta\psi \rangle - |\langle \psi | \delta\psi \rangle|^2 = \langle \delta\psi | (1 - |\psi\rangle\langle\psi|) \delta\psi \rangle. \quad (87)$$

Eq. (87) defines a complex metric over \mathbb{PH} also known as quantum geometric tensor [102]. The real (imaginary) part of the quantum geometric tensor defines a Riemannian metric tensor (a symplectic form) on \mathbb{PH} . To uncover the expression of the metric in the parameter manifold \mathcal{M} induced via a pull-back by the ground state mapping Ψ_0 introduced above, we define

$$\delta|\Psi_0(\lambda)\rangle \stackrel{\text{def}}{=} \sum_{\mu=1}^{\dim \mathcal{M}} |\partial_\mu \Psi_0\rangle d\lambda^\mu, \quad (88)$$

where $\partial_\mu \stackrel{\text{def}}{=} \partial/\partial\lambda^\mu$ with $1 \leq \mu \leq \dim \mathcal{M}$. Then, using Eq. (87), we readily obtain $ds^2 = g_{\mu\nu} d\lambda^\mu d\lambda^\nu$ with,

$$g_{\mu\nu} \stackrel{\text{def}}{=} \text{Re} \langle \partial_\mu \Psi_0 | \partial_\nu \Psi_0 \rangle - \langle \partial_\mu \Psi_0 | \Psi_0 \rangle \langle \Psi_0 | \partial_\nu \Psi_0 \rangle, \quad (89)$$

where Re denotes the real part of a complex number. Inserting in (89) the spectral resolution $I = \sum_n |\Psi_n(\lambda)\rangle \langle \Psi_n(\lambda)|$ and differentiating the eigenvalue equation $\mathcal{H}(\lambda)|\Psi_0(\lambda)\rangle = E_0(\lambda)|\Psi_0(\lambda)\rangle$, we find

$$g_{\mu\nu} = \text{Re} \sum_{n \neq 0} \frac{\langle \Psi_0(\lambda) | \partial_\mu H | \Psi_n(\lambda) \rangle \langle \Psi_n(\lambda) | \partial_\nu H | \Psi_0(\lambda) \rangle}{[E_n(\lambda) - E_0(\lambda)]^2}, \quad (90)$$

where $|\Psi_n(\lambda)\rangle$ are the eigenvectors of $\mathcal{H}(\lambda)$. The expression of $g_{\mu\nu}$ in Eq. (90) clearly suggests that at the critical points, where one of $[E_n(\lambda_c) - E_0(\lambda_c)] \geq 0$ vanishes in the thermodynamical limit [9], the metric tensor exhibits a singular behavior.

a. The transverse XY spin chain model. In order to explicitly see how the singularities of $g_{\mu\nu}$ arise, we will discuss the paradigmatic case of the XY model following [103]. Let us start from systems of quasi-free fermions that are defined by the quadratic Hamiltonian

$$\mathcal{H} \stackrel{\text{def}}{=} \sum_{i,j=1}^L c_i^\dagger A_{ij} c_j + \frac{1}{2} \sum_{i,j=1}^L \left(c_i^\dagger B_{ij} c_j^\dagger + c_j B_{ji} c_i \right), \quad (91)$$

where c_i (c_i^\dagger) are annihilation (creation) operators of L fermionic modes, while A and B are $L \times L$ real matrices symmetric and anti-symmetric, respectively. The set of ground states for the Hamiltonian system described in Eq. (91) is parametrized by orthogonal $L \times L$ real matrices T that yield the unitary part of the polar decomposition of the matrix $Z \stackrel{\text{def}}{=} A - B$. Then, one can prove that,

$$\mathcal{F}(Z, Z') \stackrel{\text{def}}{=} |\langle \Psi_Z | \Psi_{Z'} \rangle| = \sqrt{|\det[(T + T')/2]|}, \quad (92)$$

where T and T' denote the unitary parts of the polar decompositions of Z and Z' , respectively [106]. By expanding the expression of $\mathcal{F}(Z, Z')$ in Eq. (92) with respect to T' around T , it is possible to find an explicit expression for the infinitesimal metric distance ds^2 ,

$$ds^2 \approx 2(1 - \mathcal{F}) = \frac{1}{8} \text{tr} (dK)^2, \text{ with } K \stackrel{\text{def}}{=} \ln T. \quad (93)$$

From Eq. (93), if $K = K(\lambda)$ with $\lambda \in \mathcal{M}$, one obtains the following expression for the metric tensor induced over \mathcal{M} ,

$$g_{\mu\nu} = \frac{1}{8} \text{tr} (\partial_\mu K \partial_\nu K). \quad (94)$$

For translationally invariant Hamiltonians (91), the matrix K can always be cast in the form $K = i \oplus_k \theta_k \sigma_k^y$ with k being a momentum label while i denotes the complex imaginary unit. Therefore, in this working assumption, we have

$$g_{\mu\nu} = \frac{1}{4} \sum_k \frac{\partial \theta_k}{\partial \lambda^\mu} \frac{\partial \theta_k}{\partial \lambda^\nu}. \quad (95)$$

Now, the Hamiltonian for a periodic XY spin chain with an odd number of spins $L = 2M + 1$ in a transverse magnetic field h can be written as [103],

$$\mathcal{H} \stackrel{\text{def}}{=} \sum_{j=-M}^M \left[-\frac{1+\gamma}{4} \sigma_j^x \sigma_{j+1}^x - \frac{1-\gamma}{4} \sigma_j^y \sigma_{j+1}^y + \frac{h}{2} \sigma_j^z \right], \quad (96)$$

where γ is the anisotropy parameter in the x - y plane and h is the magnetic field. The critical points of this model are the lines $h = \pm 1$ and the segment $|h| < 1$, $\gamma = 0$ [9]. The Hamiltonian (96) can be cast in the form (91) by the Jordan-Wigner transformation [9]. Moreover, it is translationally invariant and Eq.(95) can be used with $\lambda^1 = h$, $\lambda^2 = \gamma$ and $\theta_k = \arccos((\cos x_k - h)/\Lambda_k)$. The quantity Λ_k denotes the single particle energy,

$$\Lambda_k \stackrel{\text{def}}{=} \sqrt{[\cos x_k - h]^2 + \gamma^2 \sin^2 x_k}, \quad (97)$$

where $x_k \stackrel{\text{def}}{=} 2\pi k/L$ with $-M \leq k \leq M$. In the thermodynamic limit where L is very large, one replaces the discrete variable x_k with a continuous variable x and substitutes the sum with an integral, that is to say,

$$\sum_{k=1}^M \rightarrow [L/(2\pi)] \int_0^\pi dx. \quad (98)$$

While at critical points this substitution is not generally feasible due to singularities in some terms of the sums in Eq. (98), when away from critical points, the resulting integrals yield analytical formulae. For instance, for $|h| < 1$, it happens that [103],

$$g = \frac{L}{16|\gamma|} \text{diag} \left(\frac{1}{1-h^2}, \frac{1}{(1+|\gamma|)^2} \right). \quad (99)$$

We remark that it is possible to provide a closed analytic expression (although in a less compact form [103]) for g in the thermodynamical limit for $|h| > 1$. We only point out here that for $|h| > 1$, the off-diagonal elements of the metric tensor are non-zero. Having the induced metric tensor, it becomes straightforward to compute the scalar curvature \mathcal{R} . One finds that for $|h| < 1$,

$$\mathcal{R} = -\frac{16}{L} \frac{1 + |\gamma|}{|\gamma|}, \quad (100)$$

while for $|h| > 1$,

$$\mathcal{R} = \frac{16}{L} \frac{|h| + \sqrt{h^2 + \gamma^2 - 1}}{\sqrt{h^2 + \gamma^2 - 1}}. \quad (101)$$

From Eqs. (100) and (101), we notice that \mathcal{R} diverges on the segment $|h| \leq 1$, $\gamma = 0$, while it is only discontinuous on the lines $h = \pm 1$. Thus, while the components of the metric tensor on the parameter manifold generally diverge at a quantum PT (see Eq. (90)), the same is not true for the scalar curvature.

b. The Dicke model. Considering the so-called Dicke model of quantum optics in the thermodynamic limit, for finite values of detuning [107], similar results as those found in Ref. [103] can be obtained. The Dicke model mimics the dipole interaction between a single bosonic mode on a system of N two-state atoms with Hamiltonian given by [108],

$$\mathcal{H}(\omega, \lambda) \stackrel{\text{def}}{=} \omega_0 J_z + \omega a^\dagger a + \frac{\lambda}{\sqrt{N}} (a^\dagger + a) (J_+ + J_-), \quad (102)$$

where

$$2J_z \stackrel{\text{def}}{=} \sum_{i=1}^N \sigma_z^{(i)}, \quad \text{and} \quad 2J_\pm \stackrel{\text{def}}{=} \sum_{i=1}^N (\sigma_x^{(i)} \pm \sigma_y^{(i)}), \quad (103)$$

ω_0 is the difference between the energies of the two states of the atom, ω is the frequency of the bosonic mode with annihilation and creation operators a, a^\dagger , respectively. The quantity λ denotes the atom-field coupling strength and N is the number of two-states atoms. In Ref. [107], assuming to be in the thermodynamic limit (i.e. $N \rightarrow \infty$) and being in the rotating wave approximation (i.e. $(a^\dagger + a)(J_+ + J_-) \approx (a^\dagger J_- + a J_+)$), the scalar curvature of the parameter manifold that corresponds to the Hamiltonian in Eq. (102) exhibits a regular behavior at the phase transitions for any ω value. However, the components of the metric tensor diverge,

$$\lim_{\lambda \rightarrow \lambda_c} \mathcal{R}(\omega, \lambda) = 4, \quad (104)$$

where $\lambda_c \stackrel{\text{def}}{=} \sqrt{\omega_0 \omega}$. The divergence of the components of the metric tensor is not genuine (that is, it is physically meaningless) and can be removed upon a suitable coordinates transformation. Therefore, unlike the classical case

where the scalar curvature diverges at second-order critical points, in the quantum case the geometry of the parameter manifold has to be investigated on a case-by-case basis. The scalar curvature can be regarded as a privileged quantity to study in the analysis of PTs because it is invariant under coordinate transformations of the parameter space, as opposed to the metric components (which, in turn, transform as rank two tensors). Indeed, for two-dimensional parameter manifolds, the Riemann curvature tensor has only one independent component, and hence leads to a unique scalar curvature. Whereas divergences in the components of the metric tensor might be due to specific coordinate choices (and hence can be removed by suitable coordinate transformations) those of the scalar curvature correspond to ones that cannot be removed by such transformations, and, therefore characterize the global (coordinate independent) properties of the manifold.

Overall, however, the above presented results point to the fact that scalar curvature does not seem to provide a universal characterization of second-order quantum PTs. Hence the need to study other geometric quantities, like geodesic curves.

2. Geodesic analysis

In addition to the notions of metric tensor and its corresponding scalar curvature, one could employ a geodesic analysis to investigate PTs. Such a geodesic approach was explored for classical PTs in Ref. [109] and for quantum PTs in Ref. [110]. The main body of work of these investigations was based on extended numerical computations [111] which were necessary due to the difficulties in integrating nonlinear sets of coupled geodesic equations for a variety of classical and quantum models. In summary, the main conclusion achieved was that geodesic trajectories are confined to a single phase. This implies that on a curved parameter manifold, points that lie in different phases are not geodesically connected. Therefore, points on the manifold can be regarded as separated by phase transitions. Furthermore, close to the critical point, geodesics exhibit a turnaround behavior. This is not unexpected since the spinodal curve, being the locus of divergences of the scalar curvature on the parameter manifold, tends to incline the geodesics. Hence, geodesics do not cross spinodal lines. It is well known that geodesics converge, or diverge, at singularities of the manifold. However, the presence of geodesics that tend to converge does not always indicate a singularity. An illustrative example is provided by a two-sphere. In this case, great circles focus at the poles which are, however, regular points on the sphere.

In addition to distinct curvature-based and geodesic-based analyses, phase transitions can also be analyzed by studying the convergence or divergence of geodesic congruences near critical points. This viewpoint relies upon the

understanding of curvature effects on geodesic paths where the evolution of a geodesic congruence is determined by means of three scalar (i.e., coordinate-independent) parameters: expansion, shear, and rotation parameters. In particular, the convergence of geodesics in a two-dimensional manifold can be quantified by a so-called scalar expansion parameter Ξ [112]. To describe and understand this expansion parameter, let us consider coordinates x^a on the parameter manifold. Then, geodesic paths satisfy the equation [Appendix A],

$$(x^a)'' + \Gamma_{bc}^a (x^b)' (x^c)' = 0, \quad (105)$$

with $\Gamma_{bc}^a \stackrel{\text{def}}{=} \frac{1}{2} g^{ad} (g_{db,c} + g_{dc,b} - g_{bc,d})$ denoting the Christoffel connections, and the prime describing the derivative with respect to an affine parameter λ along the geodesic curves. The affine parameter is conventionally taken as the square root of the line element, that is to say, $d\lambda^2 \stackrel{\text{def}}{=} g_{ab} dx^a dx^b$. For such an affinely parametrized geodesic, the geodesic equations can be derived by means of a variational principle from the Lagrangian \mathcal{L} ,

$$\mathcal{L} \stackrel{\text{def}}{=} \frac{1}{2} g_{ab} (x^a)' (x^b)'. \quad (106)$$

Denoting the normalized tangent vectors as $u^a \stackrel{\text{def}}{=} (x^a)'$, curvature effects on geodesics near criticality are measured by the tensor

$$B^a_b \stackrel{\text{def}}{=} \nabla_b u^a, \quad (107)$$

where ∇_a is the covariant derivative defined on a generic vector V^a as,

$$\nabla_a V^b \stackrel{\text{def}}{=} \partial_a V^b + \Gamma_{ac}^b V^c. \quad (108)$$

Finally, the expansion parameter is defined as $\Xi \stackrel{\text{def}}{=} B^a_a$. In the vicinity of critical points of the parameter manifold, Ξ diverges. To compute Ξ , a solution for the vectors u^a is needed and the conditions to find such solutions in an analytical manner are discussed in Ref. [113].

To recap, in a two-dimensional parameter manifold, it is possible to single out three scalar invariants (that is, coordinate independent quantities). These are the scalar curvature \mathcal{R} , the scalar expansion parameter Ξ , and, finally, the line element ds^2 which is identified with an affine parameter that measures infinitesimal distance along geodesics.

a. The transverse XY spin chain model. As pointed out in Ref. [113], there are algebraic relations among these three scalar quantities that can reveal universal behavior in both classical and quantum phase transitions, under the working assumption that the scalar curvature diverges at criticality as a power law. Indeed, for both the (classical) 1D Ising model and the (quantum) transverse XY spin chain with the Hamiltonian defined in Eq. (96), it is possible

to obtain the following relations for the scalar curvature \mathcal{R} and the scalar expansion parameter Ξ [113],

$$\mathcal{R} \sim \lambda^{-2} \text{ and, } \Xi \sim \lambda^{-1}, \quad (109)$$

respectively. The existence of algebraic relations among the above-mentioned three scalar quantities leads one to suspect that there might be a generic way to compute metrics on the parameter manifold, at least near criticality. Indeed, in Ref. [114] a method in which this can be achieved by means of ideas from scaling symmetries. Let K^a be a homothetic vector field on a manifold with metric g_{ab} such that,

$$\mathcal{L}_K g_{ab} \stackrel{\text{def}}{=} D g_{ab}, \quad (110)$$

where $\mathcal{L}_K g_{ab}$ is the Lie derivative of the metric along a curve whose tangent is K^a and, D is a constant that can be identified with the spatial dimension of the system which, in turn, is different from the dimensionality of the parameter manifold. Eq. (110) can be rewritten as,

$$K_{a;b} + K_{b;a} = D g_{ab}, \quad (111)$$

where the semicolon denotes a covariant derivative. Eq. (111) can be further simplified as follows,

$$g_{ac} K_{,b}^c + g_{bc} K_{,a}^c + g_{ab,c} K^c = D g_{ab}, \quad (112)$$

where the comma denotes an ordinary derivative with respect to the coordinate label that follows it and repeated indices imply summation. According to the information geometric analysis presented in Ref. [114] (which, in turn, was inspired by the original work appearing in Ref. [115]), near criticality, one can impose the condition that the so-called beta functions of the theory (for further details on the beta functions, we refer to Ref. [116]) are the components of a tangent vector field which is homothetic. Then, from Eq.(112), we get a set of coupled partial differential equations for the components of the metric. These equations, if solvable, will lead to solutions of the metric without a detailed knowledge of the full solution of the system. For further details, we refer to Ref. [114].

B. Complexity

From a classical physics standpoint, a complex dynamical system is characterized by extreme sensitivity to initial conditions at fixed system parameters. This type of complexity characterization requires use of the concept of trajectory which, in turn, is not present in quantum physics. Furthermore, quantum evolution is governed by a linear evolution equation which leads to a vanishing maximal Lyapunov exponent [117], a suitable quantifier of the

sensitivity to initial conditions of the physical system. The investigation of stability properties of both classical and quantum evolution can be performed by means of the classical and quantum phase-space distributions, respectively. Therefore, given the fact that both classical and quantum mechanics can be formulated in a phase-space setting and since it is also known that the number of Fourier components of the classical distribution function in phase-space grows exponentially for chaotic systems while it grows only linearly for integrable systems [118], it was proposed in Refs. [119, 120] that the complexity of quantum evolution can be quantified by means of the number of harmonics (that is, the Fourier components) of the Wigner function.

Along with this statistical characterization of complexity of quantum systems, the issue of comparing quantum and classical temporal complexity and explaining why the former is weaker than the latter has been carried out within the framework of IGAC. In particular, the tools of Jacobi Levi-Civita (JLC) vector field intensity and Information Geometric Entropy (IGE) are applied to a three-dimensional uncorrelated Gaussian model and a two-dimensional Gaussian statistical model obtained from the higher dimensional model via introduction of an additional information constraint that resembles the canonical minimum uncertainty relation in quantum theory [121].

In view of the description given in Eq. (77), a three dimensional Gaussian statistical model is characterized by the following pdf

$$p(x, y | \sigma_x, \sigma_y, \mu_x) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[-\frac{1}{2\sigma_x^2}(x - \mu_x)^2 - \frac{1}{2\sigma_y^2}y^2 \right], \quad (113)$$

where $\sigma_x, \sigma_y \in \mathbb{R}^+$ and $\mu_x \in \mathbb{R}$. At this point, we can introduce the following macroscopic constraint,

$$\sigma_x\sigma_y = \chi^2, \quad (114)$$

where χ^2 is a positive *real* constant. Therefore, by setting $\sigma \equiv \sigma_x$, the two dimensional Gaussian model arising from the one in Eq. (113) is described by the following pdf,

$$p(x, y | \sigma, \mu_x) = \frac{1}{2\pi\Sigma^2} \exp \left[-\frac{1}{2\sigma^2}(x - \mu_x)^2 - \frac{\sigma^2}{2\chi^4}y^2 \right]. \quad (115)$$

The macroscopic constraint (114) resembles the quantum mechanical canonical minimum uncertainty relation where x denotes the position of a particle and y its conjugate momentum.

By applying the IGE of Eq. (73) to the Gaussian model described above, one obtains the following relation [121],

$$\mathcal{S}^{2\text{Dmodel}}(\tau) \stackrel{\tau \gg 1}{\approx} \frac{1}{\sqrt{2}} \mathcal{S}^{3\text{Dmodel}}(\tau) \quad (116)$$

Eq. (116) quantitatively shows that the IGE is softened when approaching the two dimensional case from the three dimensional one via the introduction of the macroscopic constraint (114) that is reminiscent of Heisenberg's minimum uncertainty relation.

A further investigation that concerns the attenuation of the asymptotic temporal growth of complexity indicators of entropic motion has been carried out in Ref. [121]. In particular, a JLC vector field intensity analysis was performed. A JLC vector field $J = (J^1, \dots, J^n)$ is a solution of the JLC equation [Appendix A],

$$\frac{\nabla^2 J^k}{d\tau^2} + R_{ijr}^k \frac{d\theta^i}{d\tau} J^j \frac{d\theta^r}{d\tau} = 0, \quad (117)$$

where R_{ijr}^k are the components of the Riemann curvature tensor and $\theta^i(\tau)$ are the components of the geodesic flow upon which the IGAC relies. Eq. (117) quantifies the connection between the geodesic spread and the curvature of the given Riemannian metric. In particular, it is often employed as an indicator of the stability/instability of the dynamics by determining the evolution of perturbations of a given trajectory. Indeed, the JLC vector field accounts for such a perturbation from the geodesic dynamics. By computing JLC vector fields for the previous three and two dimensional Gaussian models in the Eqs. (113) and (115), we arrive at the following asymptotic relation

$$J^{2D}(\tau) \stackrel{\tau \gg 1}{\approx} e^{-(\Lambda_{3D} - \Lambda_{2D})\tau} J^{3D}(\tau), \quad \text{with } \Lambda_{3D} - \Lambda_{2D} > 0. \quad (118)$$

Eq. (118) is quite informative since it shows that the Jacobi vector field intensity is softened when approaching the two-dimensional case from the three-dimensional case via the introduction of the quantum-like macroscopic constraint (114). The approach so far described and implemented in [121] for tackling the issue of complexity of quantum systems relies on the similarity between the information constraint on the variances and the phase-space coarse-graining imposed by the Heisenberg uncertainty relations suggesting that it provides a possible way of explaining the phenomenon of suppression of classical chaos operated on by quantization constraints.

In addition to such investigations, the complexity of quantum systems has recently been studied in terms of less conventional techniques. One approach relies on defining a complexity measure for a pdf and then apply it to the pdf of certain variables of the system, principally the configuration or momentum. Among the measures of statistical complexity applied to quantum systems, the Fisher-Shannon measure plays a key role [122]. Indeed, the statistical complexity of a probability distribution function constructed from the measurement of quantum observables forming a continuous manifold has been proposed as a suitable way to characterize the non-integrable behavior of quantum systems. In the Hilbert space $L_2(\mathbb{R})$, a continuous manifold of observables is defined by the operators,

$$O_\omega \stackrel{\text{def}}{=} Q \cos \omega - P \sin \omega, \quad \omega \in [0, \pi[\quad (119)$$

where (Q, P) is a pair of canonically conjugate observables, namely position and momentum, respectively. By denoting $|o_\omega\rangle$ the eigenvectors of the operator O_ω with eigenvalue o_ω , the probability distribution for the output of the

measurement of O_ω is defined as $\rho(o_\omega) = |\langle o_\omega | \psi \rangle|^2$ for any state $|\psi\rangle$ of the quantum system. Therefore the statistical complexity in the space determined by the observable O_ω is defined in terms of the product of two factors [123],

$$C_{FS} \stackrel{\text{def}}{=} I[\rho(o_\omega)] \times J[\rho(o_\omega)], \quad (120)$$

where

$$I[\rho(o_\omega)] \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \rho(o_\omega) \left[\frac{d}{do_\omega} \ln(\rho(o_\omega)) \right] do_\omega, \quad (121)$$

$$J[\rho(o_\omega)] \stackrel{\text{def}}{=} \frac{1}{2\pi e} \exp(2H_{\text{Shannon}}[\rho(o_\omega)]), \quad (122)$$

are the Fisher information and the entropic power of Shannon entropy, respectively. The Shannon entropy is defined as $H_{\text{Shannon}}[\rho(o_\omega)] \stackrel{\text{def}}{=} -\int_{-\infty}^{\infty} \rho(o_\omega) \ln(\rho(o_\omega)) do_\omega$. The Fisher-Shannon measure is composed of the product of a measure of the spreading of the function and a measure of possible oscillations thereof. By analyzing the complexity (120) it is evident that it takes different values depending on the (phase space) parameter ω .

In [124], a way to overcome this difficulty by considering the entire phase space was proposed. To this end, two complexity measures were defined. First, the global Fisher Shannon (GFS) complexity was considered

$$C_{GFS}[|\psi\rangle] \stackrel{\text{def}}{=} \frac{1}{\pi} \int_0^\pi I[\rho(o_\omega)] \times J[\rho(o_\omega)] d\omega. \quad (123)$$

Second, representing a possibility for generating a base independent measure of complexity, the minimum value of the usual Fisher Shannon complexity in the complete range of ω was taken into consideration,

$$C_{MFS}[|\psi\rangle] \stackrel{\text{def}}{=} \min_{\omega \in [0, \pi]} I[\rho(o_\omega)] \times J[\rho(o_\omega)]. \quad (124)$$

The latter measure relies on Kolmogorov's complexity idea that the complexity of a system must be calculated in its simplest description. The measure of Eq. (124) is then called the *minimum Fisher Shannon* (MFS) measure of complexity.

For both measures, the global one and the minimum one, if on one side they turn out to be independent of the conjugate states, on the other side they will require changing the condition for statistical measures of complexity by a more general condition. Relying on the fact that *both, the GFS and the MFS measures give the minimum value for the states that can be represented as a Gaussian distribution for any o_ω* [124], the new definition of complexity is given in terms of non-Gaussianity, that in turn is compatible with the original Fisher Shannon complexity measure.

The quantification of complexity of motion of physical systems in terms of probability distribution functions constructed from relevant observables of the system, paved the way to apply information geometric methods to quantify complexity of quantum systems. The complexity of quantum energy levels statistics was investigated in Refs.

[125, 126], the complexity of motion in scattering induced quantum entanglement appeared in Refs. [127, 128], while the milder nature of quantum chaoticity with respect to classical chaos was preliminarily addressed from an information geometric standpoint in Refs. [121, 129–131]. The use of differential geometric methods to investigate some aspects of complexity of quantum systems has been very recently presented in Ref. [132]. Here, it was shown that the growth of computational complexity of a quantum system described by a set of strongly coupled Hamiltonians exhibits similar features as those that specify classical geodesics on a compact two-dimensional geometry of uniform negative curvature.

Unfortunately, most of these results are not conclusive. The challenge of uncovering a unifying measure of complexity and the very essence of the concept of complexity itself still remains to be uncovered in full detail. Despite substantial efforts during the last two decades or so [133–141], methods of information geometry extended to the quantum physical settings have not yet reached their full maturity with regard to the quantification of the concept of complexity of quantum evolution. In particular, to the best of our knowledge, there is no information geometric characterization of the concepts of structural complexity of quantum networks [142] nor of quantum statistical complexity [143].

VIII. CONCLUSION

In this manuscript, we reviewed the application of IG methods to describe the notion of complexity in both classical and, when possible, quantum physical settings. We started from classical phase transitions since they are intimately connected to changes in the complexity of physical systems. Then, motivated by the observation of emergent phenomena, we considered a potential measure of complexity that links heterogeneous systems to homogeneous ones. This measure is the so-called Riemann geometric entropy in Eq. (49). We then discussed an additional complexity measure that is very well-grounded on IG methods. This measure appears in Eq. (65) and is based upon the notion of stochastic interaction. Subsequently we reported on chaoticity and complexity properties addressed within the area of entropic motion. Within the entropic framework, the concept of information geometric entropy in Eq. (73) was proposed. Finally, we reviewed IG methods applied to the quantum setting where both features of PTs and complexity were addressed.

A. Phase transitions

We discussed the use of information geometric methods to study second order phase transitions in classical systems by analyzing the divergence and/or scaling behavior of the scalar Ricci curvature in the proximity of a critical point. We recall that the Ricci scalar curvature is a geometric invariant quantity. Specifically, we described the IG characterization of PTs in four models: i) the one-dimensional Ising model [23] reporting the scaling behavior of the scalar curvature in the proximity of the critical region in Eq. (33); ii) the one-dimensional q -state Potts model [52, 53] with the scaling behavior of the Ricci scalar curvature at the critical point given in Eq. (37); iii) the Ising model on planar random graphs [51] with scaling behavior of the scalar curvature reported in Eq. (40); iv) the uniform random graph model reporting the Erdős-Rényi phase transition in terms of a numerical estimation [42] of the Riemannian geometric entropy presented in Eq. (44). For the first two models, the probability distributions that define the parameter manifolds whose curvature properties are investigated are parametrized by two parameters, the inverse temperature β and the external magnetic field intensity h , respectively. Furthermore, both global and local properties of phase transitions described in terms of the scalar curvature of the parameter manifold and the components of the metric tensor, respectively, were considered for such models. For the third model, we only discussed the scaling behavior of the scalar curvature. Finally, in the fourth model, we considered a Riemannian geometric entropy, constructed from the volume of statistical manifold associated to a network (see Eq. (49)), which encodes both structural and statistical complexity aspects. We emphasize that for classical systems, including fluid and spin systems, the curvature always exhibits a divergence at the critical point of a second-order phase transition. Moreover, in the classical scenario, there is a fairly clear connection between the correlation length ξ and the scalar curvature \mathcal{R} as reported in Eq. (24).

B. Complexity

We reviewed three approaches to complexity characterization using IG methods. First, we discussed a Riemannian geometric entropy measure of complexity as reported in Eq. (49) for various complex networks. By associating smooth systems (Riemannian manifolds) to discrete systems (network), this approach allowed to transfer the use of well established geometric tools to the setting of discrete mathematics. In particular, the detection of the Erdős-Rényi phase transition suggested to consider such a Riemannian geometric entropy as a good candidate to measure the degree of organization of networks of complex systems. Indeed, one of the strengths of this Riemannian geometric entropy

is that it allows to quantify both structural and statistical features of complex networks [42]. Second, exploiting the information geometric concept of hierarchical structure of probability distributions, we discussed the complexity of a physical system in terms of the interactions among its interacting units at different scales of description. One of the main advantages of the complexity measure in Eq. (65) is that it takes into account the decomposition of a complex systems in terms of an interaction hierarchy [79]. Third, moving to a dynamical scenario, we described methods of information geometry combined with inductive inference techniques in order to quantify the complexity of entropic motion on curved statistical manifolds underlying a probabilistic description of physical systems in the presence of partial knowledge. This approach uses essentially three information geometric indicators of complexity: curvature, Jacobi vector field intensity, and the so-called information geometric entropy as defined in Eq. (68). The main advantage of this information geometric approach consists in providing a flexible platform for making statistical predictions in the presence of limited information about complex systems of arbitrary nature, in principle. At the same time, one of the essential limitations of this theoretical scheme is the lack of a corresponding fully developed quantum structure for quantifying the complexity of quantum mechanical systems [44]. We point out that the concepts of volumes of curved parameter manifolds play a key role in the first and third information geometric approaches to complexity characterizations. Instead, the Kullback-Leibler divergence $D(p||q)$ and the distance to an exponential family $D(p||\mathcal{E})$ in Eq. (63) are the essential elements needed to introduce a notion of complexity that relies upon the nature of interactions among the interacting units of the physical systems at different scales of description.

C. Quantum framework

Extending our analysis to the quantum setting, we reviewed the use of information geometric techniques to investigate second order phase transitions by analyzing the singular behavior of the metric tensor at the critical points in the thermodynamical limit. In the quantum setting, the metric structure is defined on parameter manifolds whose points are parameterized by means of the coupling constants that appear in the Hamiltonian. More specifically, we considered the information geometric characterization of phase transitions for the periodic XY spin chain with an odd number of spins in a transverse magnetic field. Both a curvature-analysis (see Eqs. (100) and (101) together with Ref. [103]) and a geodesic-analysis (for more details, see Ref. [111]) were considered for this specific model. The geometric parameters used in this case were the anisotropy parameter γ and the magnetic field intensity h . For this system the components of the metric tensor on the parameter manifold diverge at the quantum phase transition, while the Ricci scalar curvature does not (see Eqs. (100) and (101) for the scaling behavior of the curvature). We point out that, in

general, the information geometric analysis of phase transitions in the quantum setting requires more caution since statements are based upon the divergence of metric components which are not coordinate independent quantities and, therefore, singularities may be removable and unphysical. Such a scenario does not appear in the classical setting where the geometric invariance of the Ricci scalar curvature guarantees statements based upon the global properties of the parameter manifold. For example, the information geometric analysis (for more details, see Eq. 104 and Ref. [107]) of the Dicke model of quantum optics in the thermodynamic limit for finite values of detuning confirms this type of behavior. We pointed out that in addition to distinct curvature-based and geodesic-based analyses, phase transitions can be analyzed in terms of the divergence of geodesic congruences in the proximity of critical points. This line of investigation relies upon the fact that curvature affects geodesic paths and, in addition, the evolution of a geodesic congruence is determined by means of three scalar parameters [113]: expansion, shear, and rotation parameters. The scaling behavior of both the scalar curvature and the expansion for the (classical) 1D Ising model and the (quantum) transverse XY spin chain are reported in Eq. (109). For further interesting descriptions of criticality and quantum phase transitions in spin chains by means of the Fisher information, we refer to Refs. [144–146].

Finally, despite the absence of a fully-developed quantum information geometric approach to complexity characterization of any kind, we briefly reported on attempts of such an endeavor. In particular, we mentioned the softening of classical chaos operated on by quantization constraints [121], the information geometric complexity of statistical models arising in the context of quantum energy levels statistics [125] and, finally, a statistical measure of complexity for a continuous manifold of quantum observables [124].

D. Concluding remarks

From the overview presented herein several interesting questions emerge. They include among others, the issue of how to develop an IG characterization of time-dependent networks in which, for instance, the weights of the edges are not kept fixed in time [42]. Furthermore, there is no available IG characterization of the complexity of macroscopic predictions in the presence of limited dynamical information constraints on complex physical systems [90]. Finally, another open problem concerns the IG complexity characterization of a convex combination of two probability distributions and, in particular, the issue of determining whether or not the complexities of the individual constituents are related to the complexity of the convex combination [79]. We are confident that the work reported in this manuscript can offer a valid starting point from which to pursue serious attempts at addressing the above-mentioned open questions. At the same time, we emphasize that the aim of this review is not that of proposing an

information geometric measure of complexity with universal applicability. However, we do have reason to believe that the significance of the selected findings reported here will drive more researchers to consider the possibility of seeking for such a measure within the powerful framework of information geometry. In particular, we point out that, to the best of our knowledge, there does not exist any general theoretical information geometric platform for quantifying and, to a certain extent, understanding the concept of complexity of quantum dynamical motion and its connection to phase transitions.

Acknowledgements

C. C. acknowledges the hospitality of the United States Air Force Research Laboratory in Rome (New York) where part of his contribution to this work was completed. Finally, the authors thank the referees for constructive criticism leading to an improved version of this manuscript.

Appendix A: Elements of differential geometry

In what follows, we present the main differential geometric concepts of relevance to our work. For a more detailed description we refer to Refs. [147, 148].

An n -dimensional C^∞ manifold \mathcal{M} is a Hausdorff space with a countable basis which is endowed with a differentiable structure,

$$\mathcal{U} = (U_i, \varphi_i)_{i \in I}, \quad (\text{A1})$$

where $U_i \subset \mathcal{M}$ is an open set and $\varphi_i : U_i \rightarrow \varphi(U_i) (\subset \mathbb{R}^n)$ is a homeomorphism, satisfying the following conditions:

- i) $\bigcup_i U_i = \mathcal{M}$;
- ii) For any $i, j \in I$, the mapping $\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_j) \rightarrow \varphi_i(U_i)$ is a $C^\infty(\mathbb{R}^n)$ -function wherever it is well-defined;
- ii) If $V \subset \mathcal{M}$ is an open set, $\varphi : V \rightarrow \varphi(V)$ is a homeomorphism, $\varphi \circ \varphi_i^{-1}$ and $\varphi_i \circ \varphi^{-1}$ are C^∞ wherever they are well-defined, then $(V, \varphi) \in \mathcal{U}$.

The condition ii) is expressed as φ_i and φ_j being compatible. In very simple cases, \mathcal{M} is itself homeomorphic to an open set of \mathbb{R}^n and the differentiable structure is simply given by (\mathcal{M}, φ) and all sets (U_i, φ_i) such that U_i is an open set and $\varphi_i \circ \varphi^{-1}$ is a diffeomorphism. The sets U_i are called *coordinate neighborhoods* and φ_i coordinates. The pair (U_i, φ_i) is called a *local coordinate system*. The differentiable structure allows a natural way for defining a *differentiable function*. Let $f : \mathcal{M} \rightarrow \mathbb{R}$ be a function, we say that it is in $C^\infty(\mathcal{M})$ if the mapping $f \circ \varphi_i^{-1} : \varphi(U_i) \rightarrow \mathbb{R}$ is a C^∞ -function for all $i \in I$. The same applies to any subset of \mathcal{M} that inherits in a natural way the differentiable structure of \mathcal{M} . For $p \in \mathcal{M}$, we denote with $C^\infty(p)$ the set of functions whose restriction to some open neighborhood U of p is in $C^\infty(U)$. Therefore, we identify f and $g \in C^\infty(p)$ if their restrictions to some open neighborhood of p are identical. We define the *tangent space* $T_p\mathcal{M}$ to \mathcal{M} at p as the set of all maps $X_p : C^\infty(p) \rightarrow \mathbb{R}$ satisfying the following conditions:

- i) $X_p(\alpha f + \beta g) = \alpha X_p(f) + \beta X_p(g)$, with $\alpha, \beta \in \mathbb{R}$;
- ii) $X_p(fg) = X_p(f)g + fX_p(g)$ with $f, g \in C^\infty(p)$;

The quantity X_p is called a *tangent vector* and $T_p\mathcal{M}$ is in an obvious way a vector space with dimension $\dim [T_p\mathcal{M}] = n$. For each particular choice of a coordinate system, there is a corresponding canonical basis for $T_p\mathcal{M}$, with basis

vectors being,

$$(\partial_i)_p(f) = \frac{\partial}{\partial x^i} f(\varphi^{-1}(x))|_{x=\varphi(p)}. \quad (\text{A2})$$

A *vector field* is a differentiable function $X : \mathcal{M} \rightarrow \bigcup_p T_p \mathcal{M}$ such that its composition with respect to the projection $\pi : \bigcup_p T_p \mathcal{M} \rightarrow \mathcal{M}$ gives the identity map,

$$(\pi \circ X)(p) = \pi(X_p) = p. \quad (\text{A3})$$

The vector fields on \mathcal{M} are denoted as $\mathcal{T}(\mathcal{M})$. A *Riemannian metric* g is a positive symmetric tensor of rank two,

$$g(X, X) \geq 0 \text{ and } g(X, Y) = g(Y, X). \quad (\text{A4})$$

Since tensors are point-wise, we can think of a metric g_p on each of the tangent space $T_p \mathcal{M}$. A C^∞ manifold with a positive symmetric tensor is called a *Riemannian manifold* (\mathcal{M}, g) . A curve γ in \mathcal{M} is a C^∞ map $\gamma : [\alpha, \beta] \subset \mathbb{R} \rightarrow \mathcal{M}$. It is worth noticing that a curve is more than the sets of points in it. It involves the parametrization and is, thus, not a purely geometric object. On the contrary, the length $|\gamma|$ of the curve γ is a geometric object. It is defined through the vector field $\dot{\gamma}$,

$$\dot{\gamma}_{\gamma(t)}(f) = \frac{\partial}{\partial t} f(\gamma(t)), \forall t \in [\alpha, \beta], \quad (\text{A5})$$

with $f \in C^\infty(\mathcal{M})$ as,

$$|\gamma| \stackrel{\text{def}}{=} \int_\alpha^\beta \sqrt{g(\dot{\gamma}, \dot{\gamma})_{\gamma(t)}} dt. \quad (\text{A6})$$

In order to compare tangent vectors X_p and X_q such that $p, q \in U$, we introduce the notion of an *affine connection* on the manifold \mathcal{M} . It is defined as an operator ∇ ,

$$\nabla : \mathcal{T}(\mathcal{M}) \times \mathcal{T}(\mathcal{M}) \rightarrow \mathcal{T}(\mathcal{M}), \quad (\text{A7})$$

satisfying the following relations:

- i) $\nabla_X Y(\alpha Y + \beta Z) = \alpha \nabla_X Y + \beta \nabla_X Z$, with $\alpha, \beta \in \mathbb{R}$;
- ii) $\nabla_X(fY) = X(f)Y + f \nabla_X Y$, $\forall f \in C^\infty(\mathcal{M})$;
- iii) $\nabla_{fX+gY} Z = f \nabla_X Z + g \nabla_Y Z$, $\forall f, g \in C^\infty(\mathcal{M})$.

An affine connection can be thought of as a directional derivative of vector fields. In particular, $\nabla_X Y$ is the change of the vector field Y in the direction of the vector field X . However, an affine connection can be defined in many

ways since that change is ill defined without giving a rule for comparing vectors in two distinct tangent spaces $T_p\mathcal{M}$ and $T_q\mathcal{M}$. Such a rule is established by the notion of *parallel transport*. First, we say that a vector field X is parallel along the curve $\gamma : [\alpha, \beta] \rightarrow \mathcal{M}$ if,

$$\nabla_{\dot{\gamma}} X = 0, \quad (\text{A8})$$

where $\dot{\gamma}$ denotes any vector field that represents $\frac{\partial}{\partial t}$. Now, for any vector $X_{\gamma(\alpha)} \in T_{\gamma(\alpha)}\mathcal{M}$, there is a unique curve of vectors,

$$X_{\gamma(t)}, \text{ with } t \in [\alpha, \beta] \text{ and } X_{\gamma(t)} \in T_{\gamma(t)}\mathcal{M}, \quad (\text{A9})$$

such that $\nabla_{\dot{\gamma}} X = 0$. We then write,

$$X_{\gamma(\beta)} = \Pi_{\gamma} (X_{\gamma(\alpha)}), \quad (\text{A10})$$

and say that Π_{γ} defines parallel transport along γ . An affine connection can be specified by choosing a local basis $\{\partial_i\}$ with $1 \leq i \leq n$ for the vector-fields and by defining the connections symbols Γ_{ij}^k with $1 \leq i, j, k \leq n$ as

$$\nabla_{\partial_i} \partial_j \stackrel{\text{def}}{=} \Gamma_{ij}^k \partial_k, \quad (\text{A11})$$

where we adopted Einstein's summation convention according to which whenever an index appears in an expression as upper and lower index, we sum over that index. A *geodesic* is a curve with parallel tangent vector field,

$$\nabla_{\dot{\gamma}} \dot{\gamma} = 0 \text{ on } \gamma. \quad (\text{A12})$$

Associated with the notion of a geodesic is the *exponential map* induced by the connection. For all $p \in \mathcal{M}$, $X_p \in T_p\mathcal{M}$ there is a unique geodesic γ_{X_p} such that,

$$\gamma_{X_p}(0) = p, \text{ and } \dot{\gamma}_{X_p}(0) = X_p. \quad (\text{A13})$$

The geodesic is determined by the following differential equations,

$$\ddot{\gamma}^k(t) + \Gamma_{ij}^k \dot{\gamma}^i(t) \dot{\gamma}^j(t) = 0 \quad (\text{A14})$$

together with the initial conditions (A13); here $\gamma_{X_p} = (\gamma^1(t), \dots, \gamma^n(t))$ in local coordinates. Hence, by defining for $X_p \in T_p\mathcal{M}$,

$$\exp(X_p) \stackrel{\text{def}}{=} \gamma_{X_p}(1), \quad (\text{A15})$$

we have $\exp(tX_p) = \gamma_{X_p}(t)$. The exponential map is in general well-defined at least in a neighborhood of zero in $T_p\mathcal{M}$ and, moreover, can be globally defined in special cases. In general, geodesics do not have properties of *minimizing* curve length. However, on any Riemannian manifold there is a unique affine connection ∇ satisfying the following relations,

- i) $\nabla_X Y - \nabla_Y X - [X, Y] = 0$;
- ii) $Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$;

The quantity $[X, Y]$ denotes the Lie bracket and is defined as follows,

$$[X, Y](f) = X(Yf) - Y(Xf), \quad \forall f \in C^\infty(\mathcal{M}). \quad (\text{A16})$$

This connection is called the Riemannian connection or the Levi-Civita connection. Property i) is called *torsion* *freeness* and property ii) means that the parallel transport is isometric. If the connection ∇ is Riemannian, its geodesics will locally minimize curve length. When the manifold is equipped with a Riemannian metric, it is often convenient to specify the symbols $\Gamma_{ij,k}$,

$$\Gamma_{ij,k} = g(\nabla_{\partial_i} \partial_j, \partial_k). \quad (\text{A17})$$

Defining the matrix of the metric tensor and its inverse as,

$$g_{ij} = g(\partial_i, \partial_j) \text{ and } g^{ij} = (g_{ij})^{-1}, \quad (\text{A18})$$

respectively, the symbols Γ_{ij}^k are related to those previously defined as $\Gamma_{ij,l}$ in the following manner,

$$\Gamma_{ij}^k = g^{kl} \Gamma_{ij,l}. \quad (\text{A19})$$

The Riemannian connection $\Gamma_{ij,l}$ is explicitly defined as,

$$\Gamma_{ij,k} \stackrel{\text{def}}{=} \frac{1}{2} (\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}). \quad (\text{A20})$$

An important tensor field associated with a space with an affine connection is the *curvature field*,

$$\mathcal{R} : \mathcal{T}(\mathcal{M}) \times \mathcal{T}(\mathcal{M}) \times \mathcal{T}(\mathcal{M}) \rightarrow \mathcal{T}(\mathcal{M}), \quad (\text{A21})$$

such that,

$$\mathcal{R}(X, Y)Z \stackrel{\text{def}}{=} \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z. \quad (\text{A22})$$

On a Riemannian manifold \mathcal{M} , we also define the *curvature tensor* $\tilde{\mathcal{R}}$ as follows,

$$\tilde{\mathcal{R}}(X, Y, Z, W) \stackrel{\text{def}}{=} g(\mathcal{R}(X, Y)Z, W). \quad (\text{A23})$$

The Riemann curvature tensor $\tilde{\mathcal{R}}$ satisfies the following properties,

- i) $\tilde{\mathcal{R}}(X, Y, Z, W) = -\tilde{\mathcal{R}}(Y, X, Z, W);$
- ii) $\tilde{\mathcal{R}}(X, Y, Z, W) + \tilde{\mathcal{R}}(Y, Z, X, W) + \tilde{\mathcal{R}}(Z, X, Y, W) = 0;$
- iii) $\tilde{\mathcal{R}}(X, Y, Z, W) = -\tilde{\mathcal{R}}(X, Y, W, Z);$
- iv) $\tilde{\mathcal{R}}(X, Y, Z, W) = \tilde{\mathcal{R}}(Z, W, X, Y).$

If $\{\partial_i\}$ is a local basis for $T_p\mathcal{M}$, the curvature tensor $\tilde{\mathcal{R}}(X, Y, Z, W)$ can be calculated as follows

$$\begin{aligned} \tilde{\mathcal{R}}_{ijklm} &= \tilde{\mathcal{R}}(\partial_i, \partial_j, \partial_k, \partial_m) = \\ &= (\partial_i \Gamma_{jk}^s - \partial_j \Gamma_{ik}^s) g_{sm} + (\Gamma_{irm} \Gamma_{jk}^r - \Gamma_{jrm} \Gamma_{ik}^r). \end{aligned} \quad (\text{A24})$$

Furthermore, the sectional curvature $\mathcal{K}(\sigma_{X,Y})$ is given by,

$$\mathcal{K}(\sigma_{X,Y}) \stackrel{\text{def}}{=} \frac{g(\mathcal{R}(X, Y)Y, X)}{g(X, X)g(Y, Y) - g(X, Y)^2}, \quad (\text{A25})$$

where $\sigma_{X,Y}$ is the 2-plane section spanned by X and Y . In a Riemannian manifold, $\mathcal{K}(\sigma_{X,Y})$ completely determines the curvature tensor. The converse is also true whenever \mathcal{R} satisfies properties i)-iv). Two other contractions of the curvature tensor are of interest in our manuscript, the Ricci curvature and the scalar curvature. In local coordinates, the *Ricci curvature* can be written as,

$$\mathcal{Ric} \stackrel{\text{def}}{=} \mathcal{R}_{ij} dx^i \otimes dx^j, \text{ with } \mathcal{R}_{ij} = \tilde{\mathcal{R}}_{kij}^k. \quad (\text{A26})$$

The *scalar curvature* is the function $S = S(p)$ defined as the trace of the Ricci curvature \mathcal{Ric} ,

$$S(p) \stackrel{\text{def}}{=} \text{tr}_g(\mathcal{Ric}) = \mathcal{R}_i^i = \sum_{i,j \text{ with } i \neq j} \mathcal{K}(u_i, u_j), \quad (\text{A27})$$

where $\{u_k\}$ with $1 \leq k \leq n$ is an orthonormal system in $T_p\mathcal{M}$.

An important geometric aspect of the Riemannian curvature tensor in Eq. (A23) is its connection to the behavior of the geodesics obtained by integrating the nonlinear coupled ordinary differential equations in Eq. (A14). In particular, the Riemannian curvature tensor characterizes the geodesic deviation equation that describes the deviation of two

geodesics which, although assumed initially parallel to each other, may depart from each other in the presence of a gravitational field. The way for measuring the geodesic deviation relies on the definition of the Jacobi vector (or geodesic separation) field and the description of how it changes along a geodesic [149]. Let us proceed first to define the geodesic separation field and then derive its evolution equation along a given geodesic.

Consider a geodesic $\gamma : [0, 1] \rightarrow \mathcal{M}$ with respect to a connection ∇ . A Jacobi variation of γ is a smooth map $\Sigma : (-\varepsilon, \varepsilon) \times [0, 1] \rightarrow \mathcal{M}$ such that $\Sigma(0, t) \equiv \gamma(t)$ and for any $s \in (-\varepsilon, \varepsilon)$ the curve

$$\Sigma_s(t) = \Sigma(s, t), \quad t \in [0, 1] \quad (\text{A28})$$

is a geodesic. In addition, for every fixed $t \in [0, 1]$, the Jacobi variation defines a smooth curve

$$\Sigma^t(s) = \Sigma(s, t), \quad s \in (-\varepsilon, \varepsilon) \quad (\text{A29})$$

and a vector field

$$V(s) = \frac{\partial \Sigma^t}{\partial s}(s), \quad s \in (-\varepsilon, \varepsilon). \quad (\text{A30})$$

The Jacobi vector field on the geodesic $\gamma(t)$ is then defined as

$$J(t) \stackrel{\text{def}}{=} \frac{\partial \Sigma}{\partial s}(0, t), \quad t \in [0, 1]. \quad (\text{A31})$$

On the contrary, given two vectors $X, Y \in T_p M$, there exists on the geodesic $\gamma(t)$ a Jacobi field $J(t)$ such that

$$J(0) = X, \quad \frac{\nabla J}{dt}(0) = Y, \quad (\text{A32})$$

where $\frac{\nabla J}{dt}$ is the covariant derivative of J that reads in local coordinates as follows,

$$\frac{\nabla J^k}{dt} = \frac{dJ^k}{dt} + \Gamma_{ij}^k \dot{\gamma}^i J^j. \quad (\text{A33})$$

This vector $J(t)$ can be constructed by considering a curve $\alpha(s)$ such that

$$\alpha(0) = \gamma(0), \quad \dot{\alpha}(0) = X, \quad (\text{A34})$$

and by constructing on $\alpha(s)$ a vector field $A(s)$ such

$$A(0) = \dot{\gamma}(0), \quad \frac{\nabla A}{dt}(0) = Y. \quad (\text{A35})$$

At this point we can characterize the Jacobi field $J(t)$ as the solution of the Jacobi Levi-Civita equation [149],

$$\frac{\nabla^2 J}{dt^2}(t) + \mathcal{R}(J(t), \dot{\gamma}(t)) \dot{\gamma}(t) = 0, \quad (\text{A36})$$

that in local coordinates reads

$$\frac{\nabla^2 J^k}{dt^2} + \mathcal{R}_{ijr}^k \frac{d\gamma^i}{dt} J^j \frac{d\gamma^r}{dt} = 0. \quad (\text{A37})$$

Let us now briefly comment on the stability of geodesics. Consider again a geodesic $\gamma(t)$ with respect to the connection ∇ . Consider also a perturbed geodesic as

$$\tilde{\gamma}^i(t) = \gamma^i(t) + J^i(t) \quad (\text{A38})$$

where $J(t)$ is a Jacobi field. Therefore from Eq. (A36) it is clear that the stability of geodesics is completely encoded in the curvature tensor \mathcal{R} . In the particular case of constant sectional curvature \mathcal{K} of the Eq. (A25) we can clearly see how the geodesic deviation works. Firstly, Eq. (A36) becomes

$$\frac{\nabla^2 J^k}{dt^2} + \mathcal{K} J^k = 0, \quad (\text{A39})$$

where \mathcal{K} is the constant sectional curvature of the manifold. Choosing a geodesic frame, covariant derivatives become ordinary derivatives, so that the solution of (A39), with initial conditions $J(0) = 0$, $dJ(0)/dt = w(0)$, is

$$J(t) = \begin{cases} \frac{w(0)}{\sqrt{\mathcal{K}}} \sin(\sqrt{\mathcal{K}} t) & (\mathcal{K} > 0); \\ t w(0) & (\mathcal{K} = 0); \\ \frac{w(0)}{\sqrt{-\mathcal{K}}} \sinh(\sqrt{-\mathcal{K}} t) & (\mathcal{K} < 0). \end{cases} \quad (\text{A40})$$

It is evident that, in the particular case of an isotropic manifold (sectional curvature \mathcal{K} constant), geodesics are unstable if $\mathcal{K} < 0$.

-
- [1] C. Adami, *What is complexity?*, BioEssays **24**, 1085 (2002).
 - [2] A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers (1995).
 - [3] S. Amari, *Differential-Geometrical Methods in Statistics*, Springer-Verlag (1985).
 - [4] S. Amari, *Information Geometry and Its Applications*, Springer Japan (2016).
 - [5] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Oxford University Press (2000).
 - [6] N. Ay, J. Jost, H. V. Lê, L. Schwachhöfer, *Information Geometry*, Springer International Publishing (2017).
 - [7] E. P. Gyftopoulos and G. P. Beretta, *Thermodynamics*, DOVER Publications, Inc. (2005).
 - [8] K. Huang, *Statistical Mechanics*, John Wiley & Sons, Inc. (1987).
 - [9] S. Sachdev, *Quantum Phase Transitions*, Cambridge University Press (1999).
 - [10] A. Hamma, *Berry phases and quantum phase transitions*, arXiv:quant-ph/060209 (2006).
 - [11] L. Caiani, L. Casetti, C. Clementi, and M. Pettini, *Geometry of dynamics, Lyapunov exponents, and phase transitions*, Phys. Rev. Lett. **79**, 4361 (1997).
 - [12] R. Franzosi, L. Casetti, L. Spinelli, and M. Pettini, *Topological aspects of geometrical signatures of phase transitions*, Phys. Rev. E **60**, R5009 (1999).
 - [13] M.-C. Firpo, *Analytic estimation of the Lyapunov exponent in a mean-field model undergoing a phase transition*, Phys. Rev. E **57**, 6599 (1998).
 - [14] L. Casetti, E. G. D. Cohen, and M. Pettini, *Topological origin of the phase transition in a mean-field model*, Phys. Rev. Lett. **82**, 4160 (1999).
 - [15] F. Weinhold, *Metric geometry of equilibrium thermodynamics*, J. Chem. Phys. **63**, 2479 (1975).
 - [16] F. Weinhold, *Metric geometry of equilibrium thermodynamics, II. Scaling, homogeneity, and generalized Gibbs-Duhem relations*, J. Chem. Phys. **63**, 2484 (1975).
 - [17] F. Weinhold, *Metric geometry of equilibrium thermodynamics, III. Elementary formal structure of a vector-algebraic representation of equilibrium thermodynamics*, J. Chem. Phys. **63**, 2488 (1975).
 - [18] F. Weinhold, *Metric geometry of equilibrium thermodynamics, IV. Vector-algebraic evaluation of thermodynamic derivatives*, J. Chem. Phys. **63**, 2496 (1975).
 - [19] F. Weinhold, *Metric geometry of equilibrium thermodynamics, V. Aspects of heterogeneous equilibrium*, J. Chem. Phys. **65**, 559 (1976).
 - [20] A. Einstein, *On the general molecular theory of heat*, Annalen der Physik **14**, 354 (1904).
 - [21] A. Einstein, *The theory of the opalescence of homogeneous fluids and liquid mixtures near the critical state*, Annalen der Physik **33**, 1275 (1910).

- [22] G. Ruppeiner, *Thermodynamics: A Riemannian geometric model*, Phys. Rev. A **20**, 1608 (1979).
- [23] H. Janyszek and R. Mrugała, *Riemannian geometry and the thermodynamics of model magnetic systems*, Phys. Rev. A **39**, 6515 (1989).
- [24] R. Landauer, *A simple measure of complexity*, Nature **336**, 306 (1988).
- [25] M. Gell-Mann, *What is Complexity?*, Complexity **1**, 16 (1995).
- [26] R. López-Ruiz, H.L. Mancini, X. Calbetta, *A statistical measure of complexity*, Phys. Lett. A **209**, 321-326 (1995).
- [27] D. P. Feldman and J. P. Crutchfield, *Measures of statistical complexity: Why?*, Phys. Lett. A **238**, 244 (1998).
- [28] S. Wolfram, *Universality and complexity in cellular automata*, Physica D: Nonlinear Phenomena **10**, 1 (1984).
- [29] S. Wolfram, *Origin of randomness in physical systems*, Phys. Rev. Lett. **55**, 449 (1985).
- [30] M. Rasetti, *Uncertainty, predictability and decidability in chaotic dynamical systems*, Chaos, Solitons & Fractals **5**, 133 (1995).
- [31] A. N. Kolmogorov, *Three approaches to the quantitative definition of information*, International Journal of Computer Mathematics **2**, 157 (1968).
- [32] A. N. Kolmogorov, *Logical basis for information theory and probability theory*, IEEE Trans. Inf. Theory, **IT14**, 662 (1968).
- [33] Y. Pesin, Russian Mathematics Survey **32**, 55 (1977).
- [34] A. A. Brudno, *Entropy and the complexity of trajectories of a dynamical system*, Tr. Mosk. Mat. Obs. **44**, MSU, M., 124–149 (1982).
- [35] F. Blume, *On the relation between entropy and the average complexity of trajectories in dynamical systems*, Comp. Complex. **9**, 146 (2000).
- [36] P. Szepfalusy, *Characterization of chaos and complexity by properties of dynamical entropies*, Physica Scripta **T25**, 226 (1989).
- [37] A. Wehrl, *General properties of entropy*, Rev. Mod. Phys. **50**, 221 (1978).
- [38] C. H. Bennett, *How to define complexity in physics, and Why*, in “Complexity, Entropy and the Physics of Information”, SFI Studies in the Sciences of Complexity, vol.VIII, Ed. W. H. Zurek, Addison-Wesley (1990).
- [39] S. Lloyd and H. Pagels, *Complexity as thermodynamic depth*, Annals of Physics **188**, 186 (1988).
- [40] C. M. Papadimitriou, *Computational Complexity*, Addison-Wesley, Massachusetts (1994).
- [41] J. Rissanen, *Stochastic complexity and modeling*, Ann. Stat. **14**, 1080 (1986).
- [42] R. Franzosi, D. Felice, S. Mancini, and M. Pettini, *Riemannian-geometric entropy for measuring network complexity*, Phys. Rev. E **93**, 062317 (2016).
- [43] E. Olbrich, N. Bertschinger, N. Ay, and J. Jost, *How should complexity scale with system size?*, Eur. Phys. J. **B63**, 407 (2008).
- [44] S. A. Ali and C. Cafaro, *Theoretical investigations of an information geometric approach to complexity*, Rev. Math. Phys.

29, 1730002 (2017).

- [45] N. Ay, E. Olbrich, N. Bertschinger, and J. Jost, *A unifying framework for complexity measures of finite systems*, Working Paper 06-08-028, Santa Fe Institute, Proceedings of ECCS'06 (2006).
- [46] C. Cafaro, S. A. Ali, and A. Giffin, *An application of reversible entropic dynamics on curved statistical manifolds*, in Bayesian Inference and Maximum Entropy Methods in Science and Engineering, AIP Conf. Proc. **872**, 243 (2006).
- [47] S. Amari, *Differential geometry of curved exponential families-curvatures and information loss*, The Annals of Statistics **10**, 357 (1982).
- [48] K. Matsumoto, *Any statistical manifold has a contrast function on the c^3 -functions taking the minimum at the diagonal of the product manifold*, Hiroshima Math. J. **23**, 327 (1993).
- [49] G. Ruppeiner, *Riemannian geometry in thermodynamic fluctuation theory*, Rev. Mod. Phys. **67**, 605 (1995).
- [50] G. Ruppeiner, *Thermodynamic curvature measures interactions*, Am. J. Phys. **78**, 1170 (2010).
- [51] W. Janke, D. A. Johnston, and Ranasinghe P. K. C. Malmimi, *Information geometry of the Ising model on planar random graphs*, Phys. Rev. E **66**, 056119 (2002).
- [52] B. P. Dolan, D.A. Johnston, and R. Kenna, *The information geometry of the one-dimensional Potts model*, Journal of Physics **A35**, 9025 (2002).
- [53] B. P. Dolan, *Geometry and thermodynamic fluctuations of the Ising model on a Bethe lattice*, Proceedings of the Royal Society of London A **454**, 2655 (1998).
- [54] D. A. Johnston, W. Janke, and R. Kenna, *Information geometry, one, two, three (and four)*, Acta Physica Polonica B **34**, 4923 (2003).
- [55] W. Janke, D.A. Johnston, and R. Kenna, *Information geometry and phase transitions*, Physica A **336**, 181 (2004).
- [56] V.A. Kazakov, *Ising model on a dynamical planar random lattice: Exact solution*, Phys. Lett. A **119**, 140 (1986).
- [57] D.V. Boulatov and V.A. Kazakov, *The Ising model on a random planar lattice: The structure of the phase transition and the exact critical exponents*, Phys. Lett. B **186**, 379 (1987).
- [58] T. H. Berlin and M. Kac, *The spherical model of a ferromagnet*, Phys. Rev. **86**, 821 (1952).
- [59] W. Janke, D. A. Johnston, and R. Kenna, *The information geometry of the spherical model*, Phys. Rev. E **67**, 046106 (2003).
- [60] D. C. Brody and A. Ritz, *Information geometry of finite Ising models*, Journal of Geometry and Physics **47**, 207 (2003).
- [61] P. Erdős, and A. Renyi, *On the evolution of random graphs*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences **5**, 17 (1960).
- [62] R. Franzosi, D. Felice, S. Mancini, and M. Pettini, *A geometric entropy detecting the Erdős-Renyi phase transition*, Eur. Phys. Lett **111**, 20001 (2015).
- [63] D. Felice, S. Mancini, and M. Pettini, *Quantifying networks complexity from information geometry viewpoint*, J. Math.

- Phys. **55**, 043505 (2014).
- [64] S. Janson, T. Luczak, and A. Rucinski, *Random Graphs*, John Wiley & Sons, New York (2000).
 - [65] G. Leibbrandt, *Introduction to the technique of dimensional regularization*, Rev. Mod. Phys. **47**, 849 (1975).
 - [66] D. Felice, M. Hà Quang, and S. Mancini, *The volume of Gaussian states by information geometry*, J. Math. Phys. **58**, 012201 (2017).
 - [67] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Complex networks: Structure and dynamics*, Phys. Rep. **424**, 175 (2006).
 - [68] W. Aiello, L. Chung, and L. Lu, *A random graph model for power law graphs*, Experimental Mathematics **10**, 53 (2001).
 - [69] E. A. Bender and E. R. Canfield, *The asymptotic number of labeled graphs with given degree sequences*, Journal of Combinatorial Theory A **24**, 296 (1978).
 - [70] J. Wu, Y.-J. Tan, H.-Z. Deng, and D.-Z. Zhu, *A new measure of heterogeneity of complex networks based on degree sequence*, pages 66-73, Springer (2008).
 - [71] E. Estrada, *Quantifying network heterogeneity*, Phys. Rev. E **82**, 066102 (2010).
 - [72] J. Kim and T. Wilhelm, *What is a complex graph?*, Physica A **387**, 2637 (2008).
 - [73] J. P. Crutchfield and K. Young, *Inferring statistical complexity*, Phys. Rev. Lett. **63**, 105 (1989).
 - [74] S.-I. Amari, *Information geometry on hierarchy of probability distributions*, IEEE Trans. Inform. Theory **47**, 1701 (2001).
 - [75] N. Ay, *An information-geometric approach to a theory of pragmatic structuring*, Ann. Probab. **30**, 416 (2002).
 - [76] J. N. Darroch, S. L. Lauritzen, and T. P. Speed, *Markov fields and log-linear interaction models for contingency tables*, The Annals of Statistics **8**, 522 (1980).
 - [77] G. Tononi, O. Sporns, and G. M. Edelman, *A measure for brain complexity: Relating functional segregation and intergation in the nervous systems*, PNAS **91**, 5033 (1994).
 - [78] P. Grassberger, *Toward a quantitative theory of self-generated complexity*, Int. J. Theor. Phys. **25**, 907 (1986).
 - [79] T. Kahle, E. Olbrich, J. Jost, and N. Ay, *Complexity measures from interaction structures*, Phys. Rev. E **79**, 026201 (2009).
 - [80] N. Ay, E. Olbrich, N. Bertschinger, and J. Jost, *A geometric approach to complexity*, CHAOS **21**, 037103 (2011).
 - [81] N. Ay, *Information geometry on complexity and stochastic interaction*, Entropy **17**, 2432 (2015).
 - [82] N. Ay and A. Knauf, *Maximizing multi-information*, Kybernetika **42**, 517 (2006).
 - [83] S. D. Pethel, N. J. Corron, and E. M. Bollt, *Symbolic dynamics of coupled map lattices*, Phys. Rev. Lett. **96**, 034105 (2006).
 - [84] C. Cafaro, W. M. Lord, J. Sun, and E. M. Bollt, *Causation entropy from symbolic representations of dynamical systems*, CHAOS **25**, 043106 (2015).
 - [85] S. Wolfram, *Cellular Automata and Complexity*, Addison-Wesley, Reading, MA (1994).

- [86] E. Olbrich, T. Kahle, N. Bertschinger, N. Ay, and J. Jost, *Quantifying structure in networks*, Eur. Phys. B **77**, 239 (2010).
- [87] A. Caticha, *Entropic Inference and the Foundations of Physics*; USP Press: São Paulo, Brazil (2012); Available online: <http://www.albany.edu/physics/ACaticha-EIFP-book.pdf>.
- [88] C. Cafaro, *The Information Geometry of Chaos*, Ph.D. Thesis in Physics, State University of New York, Albany, NY, USA (2008).
- [89] A. Caticha, *Entropic dynamics*, AIP Conf. Proc. **617**, 302 (2002).
- [90] C. Cafaro and S. A. Ali, *Maximum caliber inference and the stochastic Ising model*, Phys. Rev. E **94**, 052145 (2016).
- [91] C. Cafaro and S.A. Ali, *Jacobi fields on statistical manifolds of negative curvature*, Physica D **234**, 70 (2007).
- [92] C. Cafaro, *Works on an information geometrodynamical approach to chaos*, Chaos, Solitons & Fractals **41**, 886 (2009).
- [93] C. Cafaro, A. Giffin, S. A. Ali, and D.-H. Kim, *Reexamination of an information geometric construction of entropic indicators of complexity*, Appl. Math. Comput. **217**, 2944 (2010).
- [94] J. P. Crutchfield and C. R. Shalizi, *Thermodynamic depth of causal states: objective complexity via minimal representations*, Phys. Rev. E **59**, 275 (1999).
- [95] A. Caticha and A. Giffin, *Updating probabilities*, AIP Conf. Proc. **872**, 31 (2006).
- [96] C. Cafaro, *Information-geometric indicators of chaos in Gaussian models on statistical manifolds of negative Ricci curvature*, Int. J. Theor. Phys. **47**, 2924 (2008).
- [97] S.A. Ali, C. Cafaro, D.-H. Kim, and S. Mancini, *The effect of microscopic correlations on the information geometric complexity of Gaussian statistical models*, Physica A **389**, 3117 (2010).
- [98] D. Felice, C. Cafaro, and S. Mancini, *Information geometric complexity of a trivariate Gaussian statistical model*, Entropy **16**, 2944 (2014).
- [99] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press (2003).
- [100] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press (2005).
- [101] J. F. Sadoc and R. Mosseri, *Geometrical Frustration*, Cambridge University Press (2006).
- [102] J. P. Provost and G. Vallee, *Riemannian structure on manifolds of quantum states*, Comm. Math. Phys. **76**, 289 (1980).
- [103] P. Zanardi, P. Giorda, and M. Cozzini, *Information-theoretic differential geometry of quantum phase transitions*, Phys. Rev. Lett. **99**, 100603 (2007).
- [104] W. K. Wootters, *Statistical distance and Hilbert space*, Phys. Rev. D **23**, 357 (1981).
- [105] S. L. Braunstein and C. M. Caves, *Statistical distance and the geometry of quantum states*, Phys. Rev. Lett. **72**, 3439 (1994).
- [106] P. Zanardi, M. Cozzini, and P. Giorda, *Ground state fidelity and quantum phase transitions in free Fermi systems*, J. Stat. Mech. Vol. **2007**, L02002 (2007).

- [107] A. Dey, S. Mahapatra, P. Roy, and T. Sarkar, *Information geometry and quantum phase transitions in the Dicke model*, Phys. Rev. E **86**, 031137 (2012).
- [108] C. C. Gerry and P. L. Knight, *Introductory Quantum Optics*, Cambridge University Press (2005).
- [109] L. Diosi, B. Lukacs, and A. Racz, *Mapping the van der Waals state space*, The Journal of Chemical Physics **91**, 3061 (1989).
- [110] A. T. Rezakhani, W.-J. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi, *Quantum adiabatic brachistochrone*, Phys. Rev. Lett. **103**, 080502 (2009).
- [111] P. Kumar, S. Mahapatra, P. Phukon, and T. Sarkar, *Geodesics in information geometry: Classical and quantum phase transitions*, Phys. Rev. E **86**, 051117 (2012).
- [112] S. Kar and S. Sengupta, *The Raychaudhuri equations: A brief review*, Pramana **69**, 49 (2007).
- [113] P. Kumar and T. Sarkar, *Geometric critical exponents in classical and quantum phase transitions*, Phys. Rev. E **90**, 042145 (2014).
- [114] R. Maity, S. Mahapatra, and T. Sarkar, *Information geometry and the renormalization group*, Phys. Rev. E **92**, 052101 (2015).
- [115] L. Diosi, G. Forgacs, B. Lukacs, and H. L. Frish, *Metricization of thermodynamic-state space and the renormalization group*, Phys. Rev. A **29**, 3343 (1984).
- [116] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press (1995).
- [117] A. Crisanti, M. Falcioni, G. Paladin, M. Serva, and A. Vulpiani, *Complexity in quantum systems*, Phys. Rev. E **50**, 138 (1994).
- [118] A. K. Pattanayak and P. Brumer, *Chaos and Lyapunov exponents in classical and quantal distribution dynamics*, Phys. Rev. E **56**, 5174 (1997).
- [119] G. Benenti and G. Casati, *How complex is quantum motion?*, Phys. Rev. E **79**, 025201(R) (2009).
- [120] V. V. Sokolov, O. V. Zhirov, G. Benenti, and G. Casati, *Complexity of quantum states and reversibility of quantum motion*, Phys. Rev. E **78**, 046212 (2008).
- [121] C. Cafaro, A. Giffin, C. Lupo and S. Mancini, *Softening the complexity of entropic motion on curved statistical manifolds*, Open Syst. & Inf. Dyn. **19**, 1250001 (2012).
- [122] J. Sañudo, R. López-Ruiz, *Some features of the statistical complexity, Fisher Shannon information and Bohr like orbits in the quantum isotropic harmonic oscillator*, J. Phys. A: Math. Gen. **41**, 265303 (2008).
- [123] J. S. Dehesa, S. López-Rosa, D. Manzano, *Configuration complexities of hydrogenic atoms*, The European Physical Journal D **55**, 539 (2009).
- [124] D. Manzano, *Statistical measures of complexity for quantum systems with continuous variables*, Physica A **391**, 6238 (2012).

- [125] C. Cafaro, *Information geometry, inference methods and chaotic energy levels statistics*, Mod. Phys. Lett. B **22**, 1879 (2008).
- [126] C. Cafaro and S. A. Ali, *Can chaotic quantum energy levels statistics be characterized using information geometry and inference methods?*, Physica A **387**, 6876 (2008).
- [127] D.-H. Kim, S. A. Ali, C. Cafaro and S. Mancini, *Information geometric modeling of scattering induced quantum entanglement*, Phys. Lett. A **375**, 2868 (2011).
- [128] D.-H. Kim, S. A. Ali, C. Cafaro and S. Mancini, *Information geometry of quantum entangled Gaussian wave-packets*, Physica A **391**, 4517 (2012).
- [129] C. Cafaro, A. Giffin, C. Lupo, and S. Mancini, *Insights into the softening of chaotic statistical models by quantum considerations*, in Bayesian Inference and Maximum Entropy Methods in Science and Engineering, AIP Conf. Proc. **1443**, 366 (2012).
- [130] A. Giffin, S. A. Ali, and C. Cafaro, *Local softening of chaotic statistical models with quantum consideration*, in Bayesian Inference and Maximum Entropy Methods in Science and Engineering, AIP Conf. Proc. **1553**, 238 (2013).
- [131] A. Giffin, S. A. Ali, and C. Cafaro, *Local softening of information geometric indicators of chaos in statistical modeling in the presence of quantum-like considerations*, Entropy **15**, 4622 (2013).
- [132] A. R. Brown, L. Susskind, and Y. Zhao, *Quantum complexity and negative curvature*, Phys. Rev. D **95**, 045010 (2017).
- [133] D. Petz and Cs. Sudar, *Geometries of quantum states*, J. Math. Phys. **37**, 2662 (1996).
- [134] D. Petz and Cs. Sudar, *Extending the Fisher metric to density matrices*; published in Geometry in Present Days Science, eds. O. E. Barndorff-Nielsen and E. B. Vendel Jensen, 21-34, World Scientific (1999).
- [135] D. Petz, *Covariance and Fisher information in quantum mechanics*, J. Phys. A: Math. Gen. **35**, 929 (2002).
- [136] P. Gibilisco and T. Isola, *Wigner-Yanase information on quantum state space: the geometric approach*, J. Math. Phys. **44**, 3752 (2003).
- [137] P. Gibilisco and T. Isola, *On the monotonicity of scalar curvature in classical and quantum information geometry*, J. Math. Phys. **46**, 023501 (2005).
- [138] P. Gibilisco and T. Isola, *Uncertainty principle and quantum Fisher information*, Ann. Inst. Stat. Math. **59**, 147 (2006).
- [139] P. Gibilisco, D. Imparato, and T. Isola, *Uncertainty principle and quantum Fisher information. II*, J. Math. Phys. **48**, 072109 (2007).
- [140] P. Facchi, R. Kulkarni, V. I. Man'ko, G. Marmo, E. C. G. Sudarshan, and F. Ventriglia, *Classical and quantum Fisher information in the geometrical formulation of quantum mechanics*, Phys. Lett. A **374**, 4801 (2010).
- [141] F. M. Ciaglia, F. Di Cosmo, D. Felice, S. Mancini, G. Marmo, and J. M. Perez-Pardo, *Hamilton-Jacobi approach to potential functions in information geometry*, J. Math. Phys. **58**, 063506 (2017).
- [142] M. Siomau, *Structural complexity of quantum networks*, AIP Conf. Proc. **1742**, 030017 (2016).

- [143] R. Tan, D. R. Terno, J. Thompson, V. Vedral, and M. Gu, *Towards quantifying complexity with quantum mechanics*, Eur. Phys. J. Plus **129**, 191 (2014).
- [144] U. Marzolino and D. Braun, *Precision measurements of temperature and chemical potential of quantum gases*, Phys. Rev. **A88**, 063609 (2013).
- [145] U. Marzolino and T. Prosen, *Quantum metrology with nonequilibrium steady states of quantum spin chains*, Phys. Rev. **A 90**, 062130 (2014).
- [146] U. Marzolino and T. Prosen, *Fisher information approach to nonequilibrium phase transitions in a quantum XXZ spin chain with boundary noise*, Phys. Rev. B **96**, 104402 (2017).
- [147] J. M. Lee, *Riemannian Manifolds*, Springer-Verlag New York, Inc. (1997).
- [148] T. Frankel, *The Geometry of Physics*, Cambridge University Press (1997).
- [149] M. Pettini, *Geometry and Topology in Hamiltonian Dynamics and Statistical Mechanics*, Springer-New York (2007).