







**Network science: Ising states of matter**Hanlin Sun <sup>1,2</sup> Rajat Kumar Panda <sup>3,4,5,6</sup> Roberto Verdel <sup>3</sup> Alex Rodriguez <sup>3,7</sup>  
Marcello Dalmonte <sup>3,4</sup> and Ginestra Bianconi <sup>1,8</sup><sup>1</sup>*School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom*<sup>2</sup>*Nordita, KTH Royal Institute of Technology and Stockholm University, Hannes Alfvéns väg 12, SE-106 91 Stockholm, Sweden*<sup>3</sup>*The Abdus Salam International Centre for Theoretical Physics (ICTP), Strada Costiera 11, 34151 Trieste, Italy*<sup>4</sup>*SISSA—International School of Advanced Studies, via Bonomea 265, 34136 Trieste, Italy*<sup>5</sup>*INFN Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italy*<sup>6</sup>*Department of Physics, University of Trieste, 34127 Trieste, Italy*<sup>7</sup>*Dipartimento di Matematica e Geoscienze, Università degli Studi di Trieste, via Alfonso Valerio 12/1, 34127 Trieste, Italy*<sup>8</sup>*The Alan Turing Institute, 96 Euston Road, London NW1 2DB, United Kingdom*

(Received 29 August 2023; accepted 26 February 2024; published 6 May 2024)

Network science provides very powerful tools for extracting information from interacting data. Although recently the unsupervised detection of phases of matter using machine learning has raised significant interest, the full prediction power of network science has not yet been systematically explored in this context. Here we fill this gap by providing an in-depth statistical, combinatorial, geometrical, and topological characterization of 2D Ising snapshot networks (IsingNets) extracted from Monte Carlo simulations of the 2D Ising model at different temperatures, going across the phase transition. Our analysis reveals the complex organization properties of IsingNets in both the ferromagnetic and paramagnetic phases and demonstrates the significant deviations of the IsingNets with respect to randomized null models. In particular percolation properties of the IsingNets reflect the existence of the symmetry between configurations with opposite magnetization below the critical temperature and the very compact nature of the two emerging giant clusters revealed by our persistent homology analysis of the IsingNets. Moreover, the IsingNets display a very broad degree distribution and significant degree-degree correlations and weight-degree correlations demonstrating that they encode relevant information present in the configuration space of the 2D Ising model. The geometrical organization of the critical IsingNets is reflected in their spectral properties deviating from the one of the null model. This work reveals the important insights that network science can bring to the characterization of phases of matter. The set of tools described hereby can be applied as well to numerical and experimental data.

DOI: [10.1103/PhysRevE.109.054305](https://doi.org/10.1103/PhysRevE.109.054305)**I. INTRODUCTION**

Networks [1–5] encode the information present in a large variety of natural and artificial interacting systems by representing them as graphs, i.e., a set of nodes (representing the element of the system) connected by links or edges (representing typically the interactions). In particular, the underlying architecture of complex systems is encoded in networks that strongly deviate from random graphs whose information content can be mined by exploiting their statistical, combinatorial as well as the geometrical and topological nature [6]. Networks are hence a simple yet very powerful framework that has been able in the last 20 years to transform our understanding of complex systems. These complex networks obey relevant organization principles while retaining a stochastic nature.

Recently great attention has been addressed to formulating unsupervised machine learning algorithms to detect

different phases of matter [7–12]. In this research line, important progress in characterizing phase transitions has recently been made by using interpretable machine learning tools such as graphical models and restricted Boltzmann machines [13–17]. The core of such approaches is that learning methods—in particular, unsupervised—shall be capable of revealing hidden structures in data sets, that correspond to physical information (such as, e.g., the existence of order parameters). However, the very advanced, complementary tools of network science to extract information from complex data of interacting systems have not yet been systematically employed for this task.

Here we want to show how network science can enrich and enhance our unsupervised characterization of phases of matter. Indeed, we will show evidence that network science provides a very transparent set of methods to investigate the characteristics of different phases across critical phase transitions and allows the determination of unsupervised indicators of their critical points.

Historically networks have been used in condensed matter for describing physical interactions existing among the elements of a system, as well as structured in configuration and Hilbert space. In principle, they can be used as well to

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

represent abstract data structures coming from numerical simulations or directly from experiments, giving access to a whole new toolbox to define and interpret many-body correlations (of arbitrary order, if expressed in terms of local observables).

Reflecting this twofold possible application of network science, recently several works have explored the use of networks to model or represent interactions in condensed matter systems. For instance, networks can be considered to define Hamiltonians whose interaction terms are determined by a complex network rather than by a lattice. In particular networks can be used to define different critical phenomena including the Ising model [18–20] and the inverse Ising model [21], as well as quantum critical phenomena including the transverse Ising model [22–24], the Bose-Hubbard model [25], the Jaynes-Cummings-Hubbard model [26], in addition to others classical collective phenomena and inverse problems [27–30]. Alternatively, networks can define quantum environments [31,32] or even multilayer couplings between interdependent superconductor networks [33]. Networks also can be used to represent correlations in complex and financial networks [34,35] as well as in quantum systems. In particular recently network structures have been shown to encode the quantum long-range mutual information (and other measures of quantum correlations) existing among the nodes of quantum lattice models in one [36–38] and two dimensions [39], providing in some cases indicators for quantum critical points.

As we will see in this work, weighted networks [40] are amenable to be analysed and treated with topological data analysis (TDA) [41–44] which provides a very efficient way to probe topological, large scale and global network properties. TDA, although until now only applied to point clouds, is raising significant interest to do unsupervised inference of phase transitions [45–51], and has wide applications, including the characterization of universal dynamics in quantum gases [52] and of confinement in lattice field theory [53–55].

Finally and most relevantly for our work, networks have been proposed to capture the underlying structure of quantum spin systems as revealed by wave function snapshots that can be probed experimentally as well as sampled from Monte Carlo simulations [56]. However, despite these very pioneering works [36,56], little attention has been so far addressed to study phases of matter using network science.

Here we launch a large-scale systematic study of the phase of matter based on network science. We leverage a multiplicity of tools developed in network science, and we reveal the combinatorial, statistical, geometrical, and topological network representation of different phases of matter. We provide an in-depth characterization of the networks generated from single snapshots of spin system configurations.

In Ref. [56] it was shown that wave function networks constructed starting from quantum wave function snapshots are strongly deviating from random graphs and for a wide range of values of the threshold distances they give rise to networks with very broad degree distribution. An open question is whether the complex properties of these networks are inherently quantum effects or they can be observed in classical systems as well.

To characterize phase transitions and critical behaviors, the Ising model is arguably the most studied model in statistical

physics. Apart from the Ising model defined on lattices, the model has also been generalized on networked structures such as small-world networks [57,58], random scale-free networks [18–20,59,60], and spatially embedded scale-free networks [61]. The phase diagram of the model is shown to be highly sensible on the value of the branching ratio of the network. These results have been also extended to the transverse Ising models [22,23]. Moreover some of the machine learning tools to study the Ising model such as restricted Boltzmann machine and graphical models [13–17] strongly leverage on their underlying network structure. However, to the best of our knowledge, the tools from network science have not been used to the analysis the simulation snapshots of the Ising model.

In this work we consider 2D Ising snapshot networks (IsingNets) following a construction proposed in Ref. [56] applied to classical 2D Ising model snapshots and we characterize their structure using advanced statistical, combinatorial, geometrical and topological tools of network theory. In order to provide an in-depth analysis of the IsingNet, across different phases we focus our attention on IsingNets obtained starting from Monte Carlo simulations of the 2D Ising model performed across the phase transition. IsingNets are obtained from a sample of state configurations of the 2D Ising model which constitute the set of nodes of the IsingNets. Each pair of nodes of an IsingNet is associated with a distance, here taken to be the Euclidean distance between the configuration snapshots. IsingNets are constructed starting from the fully connected distance matrix between the nodes by connecting only the nodes whose distance is smaller than a threshold value of the distance.

Our in-depth network analysis of the IsingNets will allow us to go well beyond the characterization of these networks based solely on the degree distribution. Possibly in the future, this in-depth analysis can be conducted also on networks built from quantum wave function snapshots in order to assess which are the properties inherently quantum in the latter networks.

Our analysis is conducted following two main directions whose goal is different but complementary. First, we will perform an analysis of the IsingNets that is agnostic about the choice of the distance threshold. In particular, we will study network properties as a function of the distance threshold. These include percolation properties, persistence homology, network embedding and statistical characterization of the distance matrices. Second, we will consider specific choices of the distance threshold and we will characterize the statistical, combinatorial and geometrical and spectral properties of the IsingNets, showing the important roles of degree-degree and weight-degree correlations in these systems.

Anticipating our main results we have found that IsingNets reflect the symmetry of the configuration space of the 2D Ising model in a prominent way. In particular, the percolation properties of the IsingNets strongly deviate from the percolation properties of networks in which the same distances among the nodes are distributed randomly. In fact, below the critical temperature of the 2D Ising model, the IsingNets are characterized by two giant components whereas their randomized counterpart displays only a single giant component. When the threshold distance defining the IsingNets is raised significantly these two giant clusters merge, but interestingly they

keep a rather compact structure as revealed by our persistent homology results highlighting that the Betti number of their clique complex are strongly suppressed with respect to their random counterpart. The Weisserstein distance between the persistence diagrams of the IsingNets and their randomized counterpart is here shown to be an unsupervised indicator of criticality as it displays a maximum in correspondence of the critical temperature. Our statistical and combinatorial analysis of the IsingNets strongly demonstrates the complex organization of these networks which display strong heterogeneity in both their topological (degree, clustering coefficient, degree correlations) and their weighted network properties. In particular nodes of higher-degree are characterized by having neighbors connected by stronger affinity weights (smaller distances). Finally, the IsingNets possess a significant geometrical organization at criticality as it is revealed by their interesting spectral properties.

Among the most important benefits of our approach is that the proposed analysis is fully interpretable. Moreover, we emphasize the applicability of the approach. Indeed while our analysis is here performed on data coming from Monte Carlo simulations the approach can be readily applied also to experimental data. However, a possible limitation of the approach is due to the significant computational cost of our proposed unsupervised TDA analysis.

## II. THE 2D ISING MODEL MONTE CARLO SIMULATIONS

We consider a square 2D lattice of dimension  $L \times L$  where on each site  $n$  is located the spin  $S_n \in \{-1, 1\}$ . The nearest-neighbor spins are interacting through the Hamiltonian

$$H = - \sum_{\langle n,m \rangle} S_n S_m. \quad (1)$$

The 2D Ising model is characterized by  $\mathbb{Z}_2$  spontaneous symmetry breaking and undergoes a second-order phase transition at  $T_c = 2/\ln(1 + \sqrt{2}) \approx 2.269$  [62]. Starting from Markov chain Monte Carlo simulations of this model, for each temperature single snapshots  $\vec{x}_i = \{S_1, S_2, \dots, S_{L^2}\}$  of the spin system are sampled at equilibrium [63]. More specifically, we use the Wolff cluster algorithm [64,65], starting from the configuration with either all up spins or all down spins, chosen at random. Next, 30 000 to 50 000 “cluster flips” are performed for the system to equilibrate. After this, we collect snapshots every 1000 to 1500 cluster flips to ensure that the collected state configurations are as uncorrelated as possible [63]. In total, we gather 10 000 snapshots during a Monte Carlo run.

For each temperature, five independent Monte Carlo simulations are performed as prescribed. By combining the sampled configuration snapshots of these runs, we thus obtain a data set with  $N_r = 50\,000$  independent thermal configurations  $\{\vec{x}_i\}_{i=1}^{N_r}$ . The starting point to construct the IsingNets is a set of  $N$  configuration snapshots  $i \in \{1, 2, \dots, N\}$  randomly selected from the data set described above, and the fully connected distance matrix  $\mathbf{d}$  of elements  $d_{ij}$  between these states. Here the distance  $d_{ij}$  between two generic snapshots  $\vec{x}_i$  and  $\vec{x}_j$  is taken to be their Euclidean distance. As discussed in previous works, such manifolds are typically living in very high dimensional subspaces [12,63], so that simple

dimensional reductions are not applicable, and a full-fledged network analysis is needed.

## III. NETWORK CHARACTERIZATION ACROSS THE DISTANCE FILTRATION

### A. Weight filtration

As anticipated in the introduction, in this first section, our analysis focuses on the properties of IsingNets observed as a function of the distance filtration. We consider IsingNets which are graphs  $G = (V, E)$  formed by a set of  $N$  nodes  $V$  and a set of links  $E$  with  $(i, j) \in E$  only if the nodes (state configurations)  $i$  and  $j$  have distance  $d_{ij} < r$ . Here  $r$  determines the distance filtration and indicates a tunable parameter that ranges from the minimum of the distances  $r_{\min} = \min_{i,j} d_{ij}$  between the  $N$  nodes to their maximum distance  $r_{\max} = \max_{i,j} d_{ij}$ . We are thus here completely agnostic about the best choice of  $r$  as we study the properties of IsingNets across all possible choices of the distance threshold  $r$ .

In particular as a function of  $r$  we will explore the percolation properties of the IsingNets which define an agglomeration from  $N$  disconnected nodes to a single connected component as  $r$  is raised from  $r_{\min}$  to  $r_{\max}$  and we will compare this process with the corresponding null model obtained from the same process applied to a randomized distance  $\mathbf{d}^{rand}$  matrix. The randomized distance matrix  $\mathbf{d}^{rand}$  is constructed by randomly reshuffling the upper triangular elements of  $d$  and subsequently symmetrizing the matrix. Therefore the null model networks display the same distribution of “distances” as the true IsingNet while being completely randomized. Note, however, that one of the main differences between the distance matrix  $\mathbf{d}$  and the randomized distance matrix  $\mathbf{d}^{rand}$  is that the entries of  $\mathbf{d}^{rand}$  are not proper distances as they do not obey the triangular inequality.

In this section, we will use a combination of tools coming from network science to analyze the IsingNets described above. In particular, anticipating the detailed description of the methods used to perform this analysis in the following paragraphs, we will use persistent homology [42–44] to further characterize topologically the mentioned aggregation process. In this way, we will show that persistent homology is able to detect the position of the critical temperature of the 2D Ising model under study. This analysis will be accompanied by the visualization of the network using minimum spanning trees [35,66], the results of the network embedding conducted using the UMAP (uniform manifold approximation and projection for dimension reduction) algorithm [67], and the statistical characterization of the distance matrix as a function of the temperature conducted using the closeness centrality [68] distribution.

### B. Percolation process

We start our investigation exploring the percolation process [2,3,69–71] monitoring the connected component of the network as a function of the filtration parameter  $r$ . To contrast the behavior of the percolation process of IsingNets with a null-hypothesis percolation process, we consider the process defined on the actual IsingNet distance matrix  $\mathbf{d}$  with the

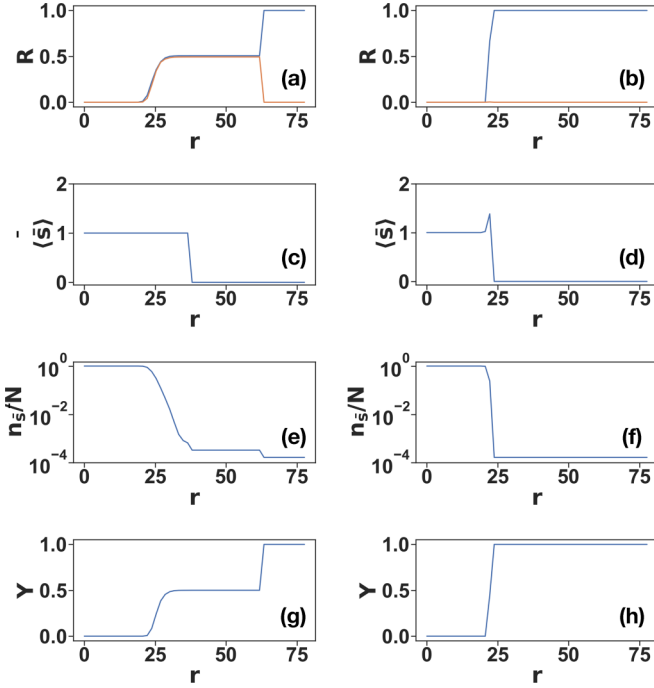


FIG. 1. Percolation properties of the IsingNet (left panels) generated from 2D Ising model Monte Carlo simulations on spin systems of linear size  $L = 40$  at temperature  $T = 2.12 < T_c$  are shown as a function of filtration parameter  $r$ . Nodes are connected if their distance is less than  $r$ . Five quantities are measured: the fraction of nodes in the largest connected component (the first row, blue line) and the fraction of nodes in the second largest connected component (the first row, orange line), the average size of components that are smaller than the second largest component  $\langle s \rangle$  (the second row), the number of components  $n_s$  (the third row), and the inverse participation ratio  $Y$  (the fourth row). The results are compared with these quantities obtained from corresponding percolation properties obtained from a randomly permuted distance matrix (right panels). The number of nodes of the IsingNets is  $N = 6000$ .

same process defined starting from a matrix  $\mathbf{d}^{rand}$  obtained by randomly permuting distances among pair of nodes.

The percolation process of IsingNets reveals a major difference with the percolation process on the randomized distance matrix: mainly the IsingNets obtained for the 2D Ising model below the phase transition, i.e.,  $T < T_c$  display for a very significant range of values of the filtration parameter  $r$ , two giant components while the randomized process only displays one giant component. This phenomenon is evident from the plot in Figs. 1 and 2 showing the relative size of the largest component  $R$  and the second largest component  $R_2$ , which are both giant, i.e., extensive for a wide range of  $r$  values. Indeed, below the critical temperature  $T_c$ , the IsingNets display two transitions as the value of the filtration parameter is raised (see Figs. 1 and 2). The first transition is characterized by the emergence of the two equal size giant components corresponding to the symmetry of the configuration snapshots of the 2D Ising model for  $T < T_c$  and the second one is characterized by the merging of these two giant components for very large values of  $r$ , characterized by the disappearance of a significant second largest connected component [orange line in Fig. 1(a) and Fig. 2(a)]. This phenomenology is dramatically different

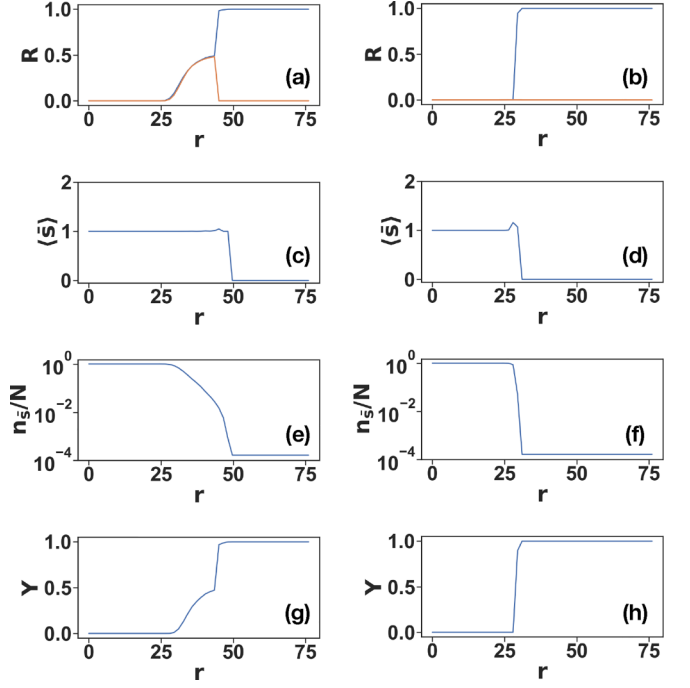


FIG. 2. Same as Fig. 1 but with IsingNets obtained from 2D Ising model simulations at  $T = 2.25$ .

from the percolation obtained in the randomized null model where the giant component is unique for every value of  $r$  (see Figs. 1 and 2). For temperatures above the critical one (see Fig. 3), instead, only one giant component is observed corresponding to the paramagnetic state of the 2D Ising model. In order to further characterize the percolation process

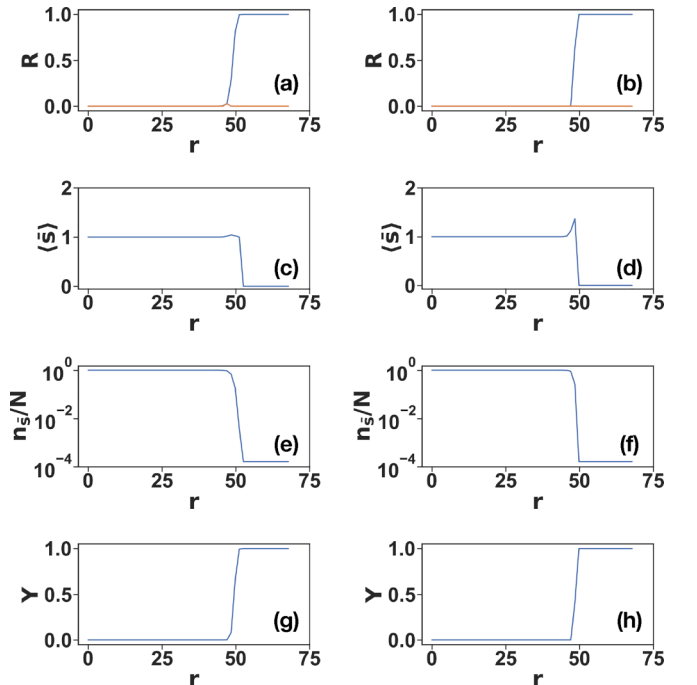


FIG. 3. Same as Fig. 1 but with IsingNets obtained from 2D Ising model simulations at  $T = 2.50$ .

we also monitor as a function of  $r$  the average size of finite components  $\langle \bar{s} \rangle$ , the number of components  $n_{\bar{s}}$  and the inverse participation ratio  $Y$  whose inverse determines the number of typical clusters. Specifically the inverse participation ratio  $Y$  is defined as

$$Y = \sum_p \left( \frac{\bar{s}_p}{\sum_q \bar{s}_q} \right)^2 = \frac{1}{N^2} \sum_p \bar{s}_p^2, \quad (2)$$

where  $\bar{s}_p$  indicates the size of the  $p$ th largest component. The abrupt increase of  $Y$  at large  $r$  further indicates the merging of two giant components [Fig. 1(g) and Fig. 2(g)].

On an Erdős-Rényi random graph the average size of finite component  $\langle \bar{s} \rangle$  plays the role of the percolation susceptibility [69], diverging in correspondence with the emergence of the (single) giant component. In the IsingNets (see Figs. 1–3) the average size of finite component  $\langle \bar{s} \rangle$  has a very suppressed maximum with respect to the same quantity measured on the considered null model, indicating that the agglomeration of the two giant clusters proceeds by subsequent agglomeration of very small components and isolated nodes rather than by the agglomeration of finite clusters of diverging average size as in a random graph. This is also confirmed by the behavior of the number of clusters  $n_{\bar{s}}$  as a function of the filtration parameter  $r$  which for the IsingNets decays less steeply than in the randomized null model. Finally, the inverse participation ratio  $Y$  for low temperatures reveals a significant plateau at  $Y = 1/2$  indicating the existence of two giant components of approximately equal size (see Fig. 1). Above the critical temperature, for  $T > T_c$  as the filtration parameter is raised only one giant component emerges and the difference with respect to the randomized null model is reduced (see Fig. 3).

### C. Persistent homology

Topology is the study of shapes and their invariant properties under continuous deformations (see for an introduction [6,41]). Major examples of topological invariants are the Betti numbers. The Betti number  $\beta_0$  indicates the number of connected components, the Betti number  $\beta_1$  indicates the number of 1D holes, the number of  $\beta_2$  indicates the number of 2D holes, etc. For instance, a point has Betti numbers  $\beta_0 = 1$  and  $\beta_n = 0$  for any other value of  $n$ , a circle has nonzero Betti numbers  $\beta_0 = \beta_1 = 1$  and a sphere has nonzero Betti numbers  $\beta_0 = \beta_2 = 1$ . An important result of algebraic topology [6,41–44] is that the  $\bar{n}$ -dimensional Betti number  $\beta_{\bar{n}}$  is the rank of the  $\bar{n}$ -dimensional homology group of the considered topological space.

In the discrete setting, the Betti numbers are defined in general for simplicial complexes. Simplicial complexes are a type of higher-order network formed by a set of simplices such as nodes, links, triangles, tetrahedra, etc. They have the additional property of being closed under the inclusion of faces of each simplex. This last property implies that if a triangle belongs to the simplicial complexes also all its links and nodes belong to the simplicial complex.

Topological data analysis [40–44,72] and in particular persistent homology allows to characterize the topological properties of data as a function of a filtration

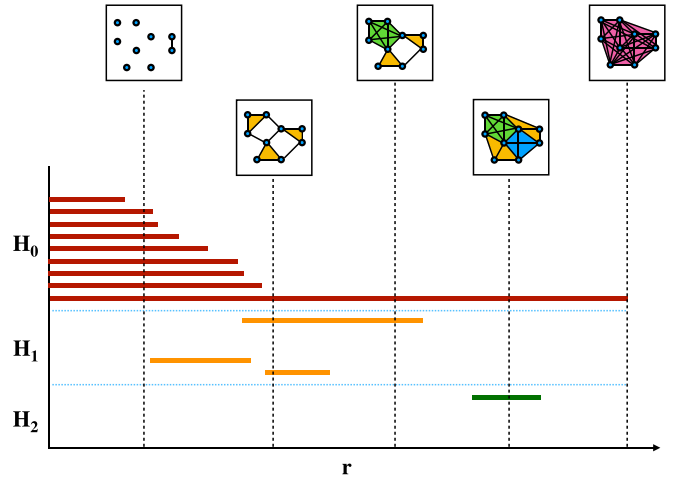


FIG. 4. Schematic illustration of the filtration process. The bar codes are used to show the appearance and disappearance of topological features corresponding to different homology classes as the filtration parameter  $r$  is increased. The filtration process ends at  $r = r_{\max}$  when all  $N$  nodes are fully connected and forms an  $N$  simplex. Simplices of different dimensions are indicated by different colors.

parameter and is becoming increasingly important in network and data science with applications ranging from the study of gene expression to the investigation of brain networks. As mentioned in the introduction TDA is recently becoming a very popular computational tool to study phase transitions as well [45–55]. However, in this context, most of the TDA so far are performed on point clouds rather than on networks. In our setting, when each pair of nodes  $(i, j)$  is assigned a distance  $d_{ij}$ , persistent homology characterizes the topology of data by forming a simplicial complex representation of the data and characterizing its homology, i.e., the connected components ( $H_0$  homology classes), the independent cycles (one-dimensional holes;  $H_1$  homology classes), the two-dimensional holes ( $H_2$  homology classes), etc. The simplicial complexes [6,40,42,43] that we use to perform the topological data analysis are the so-called Vietoris-Rips complexes of the network generated by filling all the simplices having all links at distance  $d_{ij} < r$ . Thus as a function of  $r$  the set of simplicial complexes forms a filtration.

As the filtration parameter  $r$  is raised, first each node belongs to a different connected component, then connected components merge progressively as described also by the previous paragraph (see Fig. 4 for a schematic description of the filtration). However, as  $r$  increases there is the potential also for one-dimensional holes (or network cycles) to emerge with each independent cycle represented by a different  $H_1$  homological class. Eventually, as  $r$  increases these cycles will become filled and thus disappear. Moreover, also higher-dimensional holes represented by higher-dimensional homological classes  $H_{\bar{n}}$  might arise and eventually coalesce as  $r$  is further increased giving rise to bar codes representing the topology of the data.

By monitoring the homology classes as a function of  $r$ , the results of persistent homology are typically summarized by

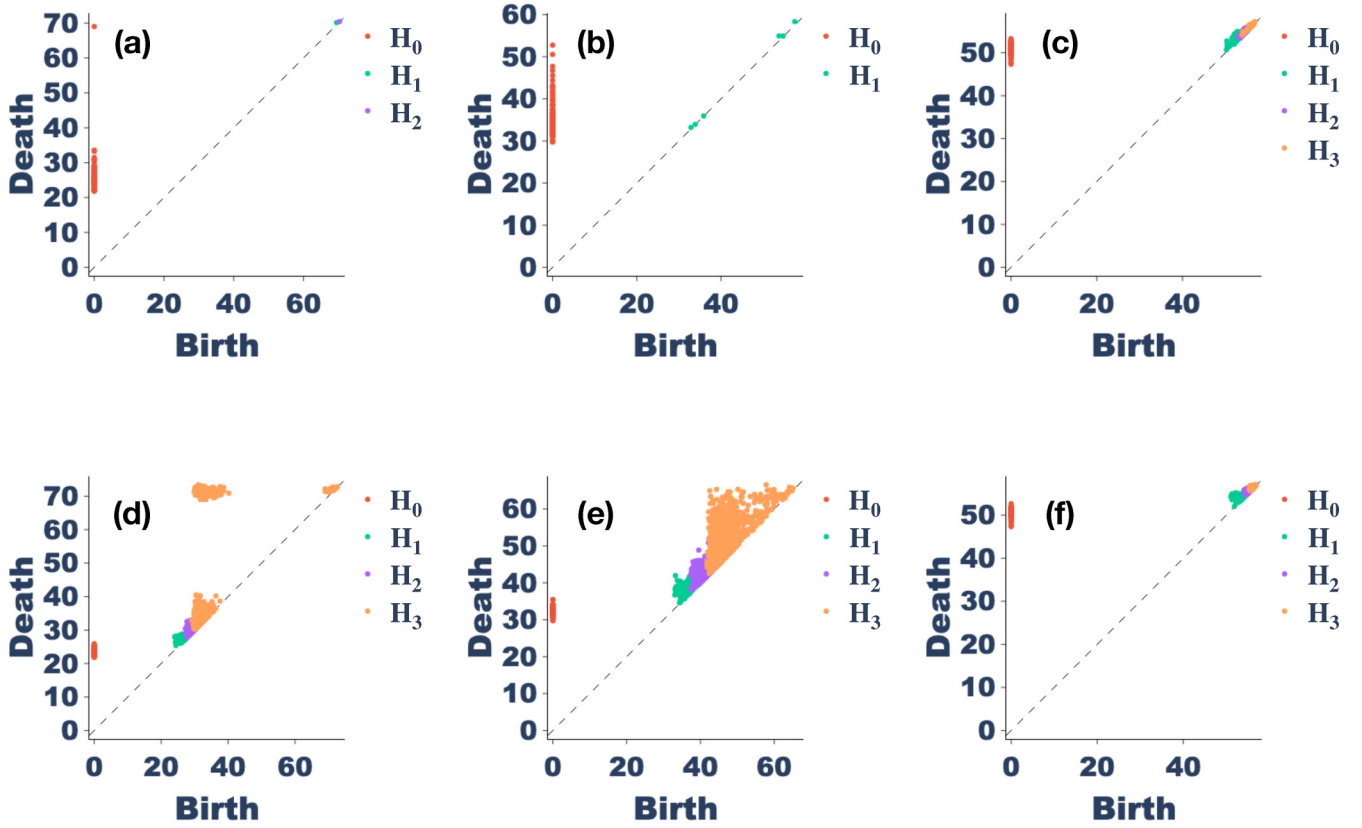


FIG. 5. The persistent diagram corresponding to homology classes in  $H_0$ ,  $H_1$ ,  $H_2$ , and  $H_3$  of the IsingNet clique complexes are plotted as a function of the filtration parameter  $r$ . Panels (a), (b), and (c) show the persistent diagrams of IsingNets obtained from the spin system of linear size  $L = 40$  at  $T = 2.12$  (a),  $T = 2.25$  (b), and  $T = 2.50$  (c). Panels (d), (e), and (f) show the persistent diagram of corresponding randomized null models obtained at  $T = 2.12$  (d),  $T = 2.25$  (e), and  $T = 2.50$  (f). The networks are formed by  $N = 100$  nodes.

bar codes where each homology class is represented as a bar that extends through the corresponding range of values of  $r$  for which the homology class is observed (see Fig. 4). The bar codes are then represented by persistent diagrams where for each homology class the value of  $r$  where the homology class is disappearing (death) is plotted versus the value of  $r$  corresponding to the first appearance of the homology class (birth); see, for instance, Fig. 5. Significant homology classes are the ones represented by points far from the diagonal which last for a large interval of values of the filtration parameter  $r$ .

The investigation of the persistent diagrams for the IsingNets can further characterize these discrete structures by revealing important properties about how the clusters merge as a function of  $r$ . In particular, the persistent diagrams of the IsingNets allow us to show that clusters remain compact with a suppression of the Betti numbers with respect to the corresponding persistent diagram of the randomized null models. This result is evident from Figs. 5 and 6 where we plot the persistent diagram (for homology  $H_{\bar{n}}$  with  $\bar{n} \in \{0, 1, 2, 3\}$ ) and the Betti numbers  $\beta_{\bar{n}}$  with  $\bar{n} \in \{0, 1, 2, 3\}$  as a function of the filtration parameter  $r$  for IsingNets and their randomized null models as a function of the temperature  $T$ . Indeed, the persistent diagrams of IsingNets corresponding to homology  $H_{\bar{n}}$  with  $\bar{n} \in \{0, 1, 2, 3\}$  (see Fig. 5) reveal that homology classes for the IsingNets are less persistent than the homology classes of the null model as they are represented in the diagram by points closer to the diagonal

than in the null model. Moreover also the Betti numbers  $\beta_{\bar{n}}$  with  $\bar{n} \in \{0, 1, 2, 3\}$  are strongly suppressed with respect to their null model counterpart. The persistent diagrams of IsingNets can be compared to the persistent diagrams of their randomized null model as a function of the temperature. This comparison can be performed by considering different measures of distances between persistent diagrams.

Here we show both the Weisserstein distance and the Betti distance among the persistent diagrams of the IsingNets and their corresponding randomized null model. The Weisserstein distance captures the minimum distance over all perfect matchings between points in two persistent diagrams,

$$d_W = \left[ \inf_{\eta: I \rightarrow N} \sum_{x \in I} \|x - \eta(x)\|_\infty^2 \right]^{1/2}, \quad (3)$$

where  $x \in I$  indicates a point  $x = (b, d)$  with birth  $b$  and dead  $d$  in the persistent diagram of the IsingNet, while  $\eta(x) \in N$  indicates the matched point in the persistent diagram of the corresponding null model. The map  $\eta$  denotes any bijection between  $I$  and  $N$ . The Betti distance computes the  $L_2$  distance between Betti curves of two persistent diagrams,

$$d_\beta = \left[ \sum_r (\beta^{[I]}(r) - \beta^{[N]}(r))^2 \right]^{1/2} \quad (4)$$

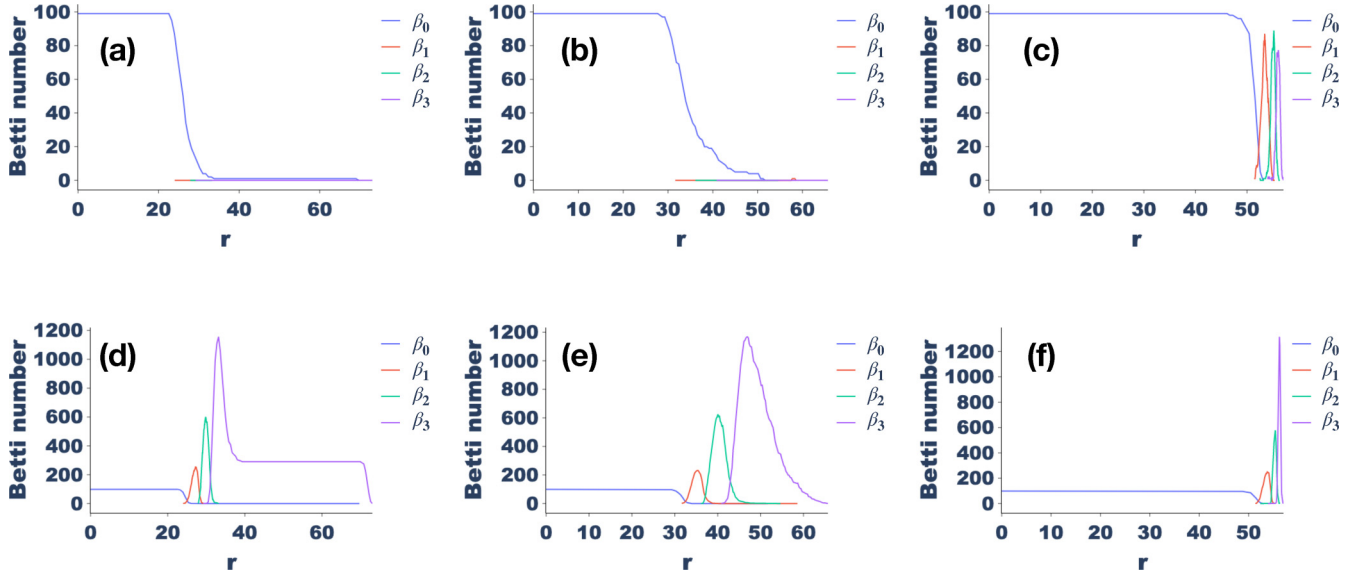


FIG. 6. Betti numbers  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  of the IsingNet clique complexes are plotted as a function of the filtration parameter  $r$ . Panels (a), (b), and (c) show the persistent diagrams of IsingNets obtained from the spin system of linear size  $L = 40$  at  $T = 2.12$  (a),  $T = 2.25$  (b), and  $T = 2.50$  (c). Panels (d), (e), and (f) show the persistent diagram of corresponding randomized null models obtained at  $T = 2.12$  (d),  $T = 2.25$  (e), and  $T = 2.50$  (f). The networks are formed by  $N = 100$  nodes.

where  $\beta^{[l]}(r)$  and  $\beta^{[N]}(r)$  indicate the Betti number of the IsingNet and the corresponding null model with filtration parameter  $r$ .

These distance measures provide good indicators of the critical points as they display a maximum approaching the critical temperature of the 2D Ising model as the size  $L$  of the 2D lattice increases (see Fig. 7). Note that here, due to the computational cost of calculating the persistent diagrams corresponding to higher-order homological classes, we focus only on homology classes  $H_0$  and  $H_1$ .

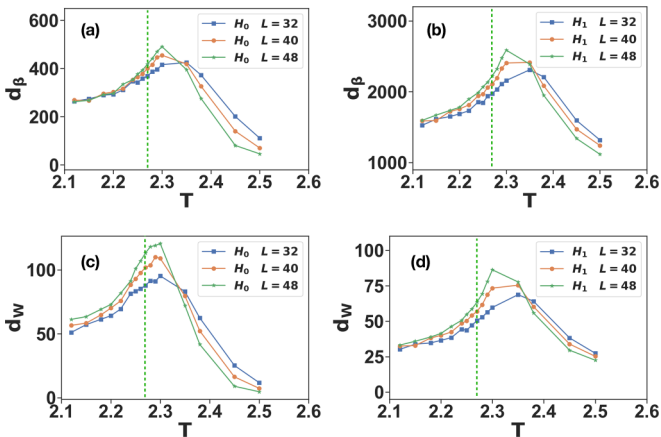


FIG. 7. Betti distance  $d_\beta$  (a, b) and the Wasserstein distance  $d_w$  (c, d) between the persistent diagrams of the IsingNets and its randomized null models are plotted as a function of the temperature  $T$  for homology  $H_0$  and  $H_1$ . The plot shows the finite size scaling of the distances on systems of size  $L = 32$ ,  $L = 40$ , and  $L = 48$ . The dashed line indicates the critical temperature  $T_c$ . The IsingNets on which the persistent diagrams have been calculated have  $N = 300$  nodes.

This is rather clear evidence that the IsingNets have a topology that encodes fundamental properties of the underlying spin system.

#### D. Minimum spanning tree visualization

Our unsupervised analysis of the IsingNets includes their visualization which reveals their highly heterogeneous structure below the critical temperature  $T_c$  and at criticality. A very efficient way to visualize the IsingNets is by plotting their minimum spanning trees (MSTs) [35,66]. The MST is the subtree of the network whose sum of the distances between the connected nodes is minimal, and its topology reveals important properties of the fully connected weighted distance matrix of the IsingNets. In particular, as shown in Fig. 8 for  $T$  deep in the ferromagnetic phase the topology of the MST of the IsingNets is dominated by very relevant hub nodes, and the network displays a clear partition between the two clusters detected by the  $K$ -means algorithm with  $K = 2$  indicated in the figure by two different colors of the nodes. As the temperature is raised to the critical region,  $T \simeq T_c$  the hubs of MST become less dominant. Above the phase transition, the MST becomes clearly more random with a suppression of the hubs in the MST.

#### E. Network embedding

Our analysis of the IsingNets is here enriched by considering the UMAP 2D embedding of the fully connected network endowed with the distance matrix  $\mathbf{d}$ . UMAP is a widely used embedding algorithm exploiting dimension reduction (for more details see, for instance, Ref. [73]). The embedding is here performed as a function of the

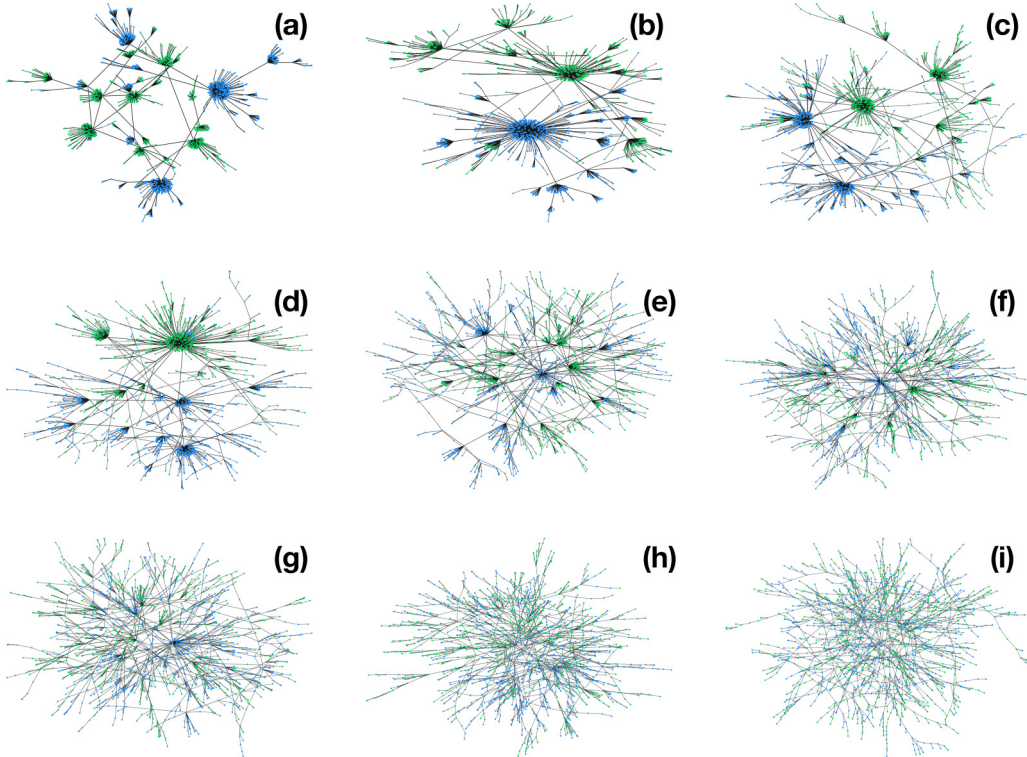


FIG. 8. Minimum spanning trees (MSTs) of the IsingNets obtained from the spin system of linear size  $L = 40$  are plotted for different temperatures  $T$  above and below the phase transition at  $T_c = 2.269\dots$ . The nodes are colored using a  $K$ -means clustering algorithm with  $K = 2$ . The number of nodes of the MSTs is  $N = 2000$ . The MSTs are generated at  $T = 2.12$  (a),  $T = 2.20$  (b),  $T = 2.26$  (c),  $T = 2.28$  (d),  $T = 2.30$  (e),  $T = 2.35$  (f),  $T = 2.38$  (g),  $T = 2.50$  (h),  $T = 3.50$  (i).

temperature  $T$  and the nodes are colored according to their clustering in two groups performed using  $K$  means (see Fig. 9).

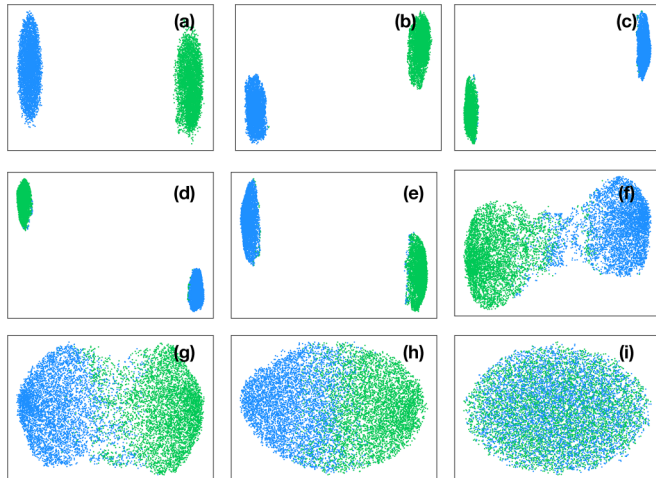


FIG. 9. IsingNets obtained from the spin system of linear size  $L = 40$  at different temperatures  $T$  are embedded into a 2D space using the Uniform Manifold Approximation and Projection (UMAP) embedding algorithm. The nodes are colored using a  $K$ -means clustering algorithm. The number of nodes of the IsingNets is  $N = 10^4$ . The IsingNets are generated at  $T = 2.12$  (a),  $T = 2.20$  (b),  $T = 2.26$  (c),  $T = 2.28$  (d),  $T = 2.30$  (e),  $T = 2.35$  (f),  $T = 2.38$  (g),  $T = 2.50$  (h),  $T = 3.50$  (i).

The UMAP embedding provides a clear visualization of the two main clusters of the nodes of the IsingNets present for  $T < T_c$  and corresponding to the symmetry among configuration snapshots and of their merging as the temperature  $T$  is raised above the critical temperature.

### F. Metric-based centrality measures

To conclude our characterization of the IsingNets without imposing a fixed value of the filtration parameter, we present here the statistical properties of some important geometrical aspects of the IsingNet captured by the closeness centrality of the nodes.

The closeness centrality  $CL_i$  of a node  $i$  measures how close is the node to the other nodes of the network, and is defined as the inverse of the average distance of node  $i$  to the other nodes of the network:

$$CL_i = \frac{N - 1}{\sum_{j \neq i} d_{ij}}. \tag{5}$$

We investigate the statistical properties (the mean  $\langle CL \rangle$ , the standard deviation  $\sigma(CL)$ , the skewness  $Sk(CL)$  and the kurtosis  $Ku(CL)$ ) of the distribution of the closeness centrality in IsingNets as a function of the temperature  $T$ . In Fig. 10 we show that the average closeness centrality decreases as a function of the temperature, demonstrating that on average the distances between the nodes are higher at higher temperatures. The higher-order statistics of the closeness centrality distribution are even more revealing of the IsingNets organization and the skewness and kurtosis of the closeness distribution



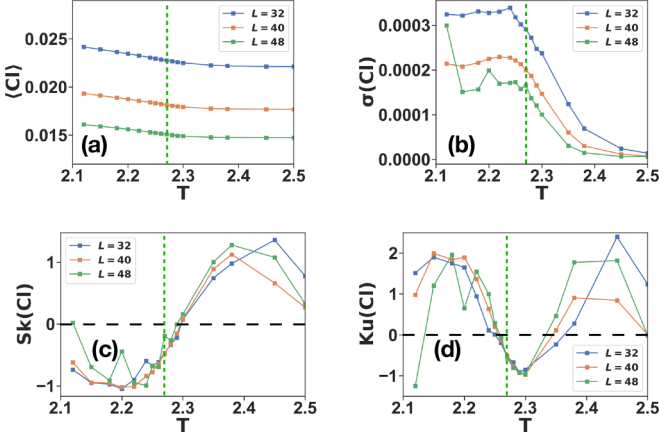


FIG. 10. The mean  $\langle CI \rangle$ , standard deviation  $\sigma(CI)$ , skewness  $Sk(CI)$ , and kurtosis  $Ku(CI)$  of closeness distribution at different temperatures  $T$  and sizes  $L$  are plotted on the IsingNets where all the distance between each pair of nodes is retained. The closeness is calculated via Eq. (5). The corresponding random networks are formed by randomly permuting the distances between node pairs. The average closeness centrality of the IsingNets coincides by construction with the average of the null model (not shown) however the higher-order statistics strongly depart from the null model behavior. The vertical green dashed lines indicate the critical temperature  $T_c$  the horizontal black dashed lines indicate  $Sk(CI) = 0$  in panel (c) and  $Ku(CI) = 0$  in panel (d).

provide a good unsupervised indicator of the critical point (see Fig. 10). Indeed the standard deviation of the closeness centrality displays a maximum for temperatures close to the critical temperature; the skewness of the closeness centrality is negative below the critical temperature and positive above the critical temperature and the kurtosis becomes negative above the critical temperature, while at very high temperature is strongly affected by noise.

#### IV. NETWORK CHARACTERIZATION OF ISINGNETS AT A GIVEN VALUE OF THE THRESHOLD DISTANCE $R$

##### A. IsingNets at given threshold distance $r$

In this section, we study the properties of IsingNets where we fix a given choice of filtration parameter  $r$ . In particular we will consider the IsingNets, whose  $N \times N$  adjacency matrix  $\mathbf{A}$  has elements

$$A_{ij} = \theta(r - d_{ij}), \quad (6)$$

with  $\theta(x) = 1$  for  $x > 0$  and  $\theta(x) = 0$  otherwise. In the following, we will indicate with  $i \sim j$  two neighbor nodes for which  $A_{ij} = 1$ .

We adopt the same choice of the parameter  $r$  proposed in Ref. [56] where it was shown that for a wide range of choices of  $r$  the IsingNets are scale-free presenting often power-law exponents less than two which are known to occur in a variety of context [74–77]. In particular, here we take  $r$  equal to the average distance of the fifth nearest neighbor.

Similarly, the randomized network forming our null model in this section is performed by thresholding the randomized

distance matrix  $\mathbf{d}^{rand}$  with the same threshold used for the corresponding IsingNet.

We consider the statistical and combinatorial properties of these networks where we assign to each link  $(i, j)$  a weight  $w_{ij}$  equal to the inverse of the distance  $d_{ij}$ , provided this distance is smaller than  $r$ :

$$w_{ij} = \frac{1}{d_{ij}} A_{ij}. \quad (7)$$

We provide an in-depth network analysis of these networks investigating their degree and strength distribution, and going beyond these statistical properties providing evidence of the presence of degree correlations, of a nontrivial  $k$ -core structure and of weight degree correlations. Moreover, the investigation of the spectral properties of the IsingNets will demonstrate nontrivial signatures of criticality. This analysis provides clear evidence that not only the degree distribution of IsingNets is strongly different from the one of an Erdős-Rényi random graph, but the network also obeys important higher-order correlations reflecting the correlations existing in the spin system configuration snapshots.

##### B. Degree and strength distribution

One of the most simple yet fundamental property of a network is its degree distribution  $P(k)$  which characterize globally the network starting from the knowledge of the node's degrees where the degree  $k_i$  of the node  $i$  indicates the sum of the links incident to it:

$$k_i = \sum_{j=1}^N A_{ij}. \quad (8)$$

Thus while the degree is a local property of the nodes the degree distribution  $P(k)$  is able to characterize the global properties of the networks. In particular scale-free [78–80] and in general degree distribution with second  $\langle k^2 \rangle$  (and eventually first  $\langle k \rangle$ ) moment diverging with the network size have been shown to have a significant role in determining the response of the network to random damage and the critical behavior of the dynamics defined on top of these networks, such as epidemic spreading and the Ising model [1,79].

For weighted networks it is also possible to define weighted degree also called *strength*  $s_i$  of the generic node  $i$  [81] given by the sum of the weights of its incident links:

$$s_i = \sum_{j=1}^N w_{ij}. \quad (9)$$

From the sequence of the node's strengths, it is possible to extract also the strength distribution  $P(s/s_0)$  where  $s_0$  is the minimal strength of the links. This distribution has been shown to display broad distributions in a number of real weighted networks, such as collaboration networks and airport networks [82].

In Fig. 11 we plot the degree  $P(k)$  and strength  $P(s/s_0)$  distribution for temperatures below and above the phase transition demonstrating that these distributions are broad. We note that below the critical temperature  $T_c$ , the networks are not only broad but also dense, i.e., having an average degree growing with the network size (for models of these networks see [74–77]).

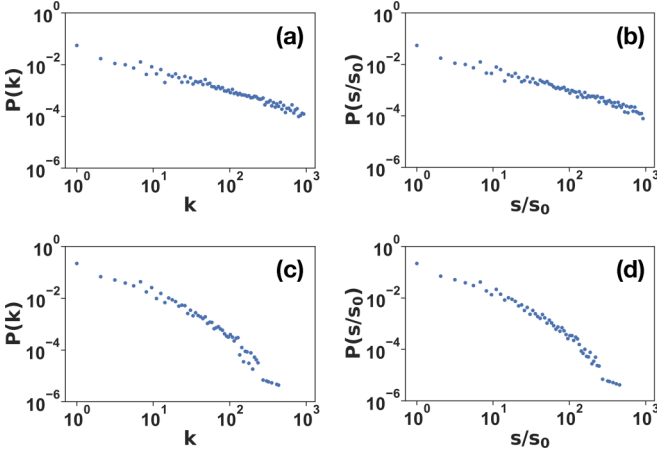


FIG. 11. Degree distribution  $P(k)$  and strength distribution  $P(s/s_0)$  of IsingNets obtained from the spin system of linear size  $L = 40$  formed by  $N = 10^4$  nodes at temperature  $T = 2.12$  and  $T = 2.50$ . Panel (a) shows the degree distribution of IsingNet at  $T = 2.12$ . Panel (b) shows the strength distribution of the same network as panel (a). Panel (c) is the same as panel (a) but obtained at  $T = 2.50$ . Panel (d) is the same as panel (b) but obtained at  $T = 2.50$ .

To quantify the scale-free nature of these distributions we plot the ratio between the second moment, the average degree  $\langle k \rangle$  and the average strength  $\langle s \rangle$  as a function of the temperature, showing that that  $\langle k \rangle$ , and  $\langle k^2 \rangle / \langle k \rangle$  are good indicators of the critical point displaying a maximum for  $T = T_c$  without noticeable finite size effects for the sizes investigated in this work (see Fig. 12). We note that origin of such power laws is not necessarily related to the presence of a critical point, as the data sets in Fig. 11 are representative of regimes where the correlation length is not much larger than  $L$  (and, for  $T = 2.5$ , is smaller). We attribute these features of  $P(k)$  to the strong correlations present in the system (so: not critical behavior, but rather, just correlation length much larger than the lattice spacing).

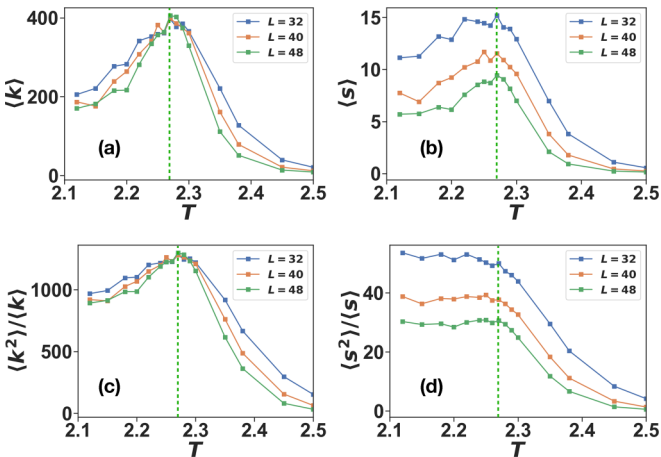


FIG. 12. The average degree  $\langle k \rangle$  (a) and the average strength  $\langle s \rangle$  (b) are plotted together with the ratio  $\langle k^2 \rangle / \langle k \rangle$  (c) and  $\langle s^2 \rangle / \langle s \rangle$  (d) as a function of temperature  $T$ . The dashed line indicates the critical temperature  $T_c$ . The IsingNets are formed by  $N = 10^4$  nodes.

### C. Degree correlations

In order to go beyond single node statistics and to explore how far IsingNets are from random networks with a given degree distribution, in this section we characterize the degree correlations [2,3] of the IsingNets. Degree-degree correlations measure to what extent the degree of two end nodes of the same link are correlated. Degree correlations are typically classified as either assortative or disassortative. Assortative degree correlations imply that highly connected nodes are more likely to be connected to highly connected nodes while low-degree nodes are more likely to be connected to low-degree nodes than in a maximally random network with the same degree distribution. Conversely, disassortative networks are networks in which highly connected nodes are more likely to be connected to low-degree nodes than in the maximally random network with the same degree distribution. Examples of assortative networks are social networks while examples of disassortative networks include the Internet and the protein interaction networks. The degree-degree correlations can be quantified by considering the average degree of the neighbor of a node  $k_{nn}(i)$  defined as

$$k_{nn}(i) = \frac{1}{k_i} \sum_{j \sim i} k_j. \quad (10)$$

When  $k_{nn}(i)$  tends to be higher for nodes of higher degree  $k_i$  the network is classified as assortative. Instead when  $k_{nn}(i)$  is typically lower for nodes of higher degree  $k_i$  then the network is classified as disassortative. The IsingNets are clearly displaying a disassortative behavior for  $T < T_c$  that deviates strongly from the behavior of the null model in which the distance matrix  $\mathbf{d}$  is reshuffled (see Fig. 13). On the contrary, for  $T > T_c$  the trend of  $k_{nn}$  versus  $k$  does not appear to be fully monotonic, while nodes of larger degrees remain more likely to connect to low-degree nodes.

Interestingly the degree correlations are also affecting the average clustering coefficient  $C(k)$  [3,83,84] of nodes of degree  $k$  which display a decay as a function of  $k$  typical of networks with disassortative degree correlations (see Fig. 13).

Degree-degree correlations can also be detected by the Pearson correlation coefficient  $\bar{r}$  [2] between degrees of linked nodes which is plotted as a function of the temperature in Fig. 14(a) showing a clear negative (disassortative) correlations for low temperatures which strongly deviates from the null model. In Fig. 14(b) we also report the average clustering coefficient  $C$  as a function of the temperature showing that IsingNets display a much larger average clustering coefficient than the null model counterpart and that this average clustering coefficient is higher deep in the ferromagnetic phase (lower temperatures).

### D. K-core structure

Networks can be decomposed in nested  $K$  cores characterizing their core-periphery structure [85–87]. A  $K$  core is a subgraph of the network formed by a set of  $M(K)$  nodes each having at least  $K$  connections with the other nodes of the set. Power-law networks with exponent  $\gamma \in (2, 3]$  display a significant  $K$ -core structure with the maximum  $K$  diverging with the network size and a power-law decay of  $M(K)$  as a function of  $K$ . On the other hand Erdős-Rényi networks

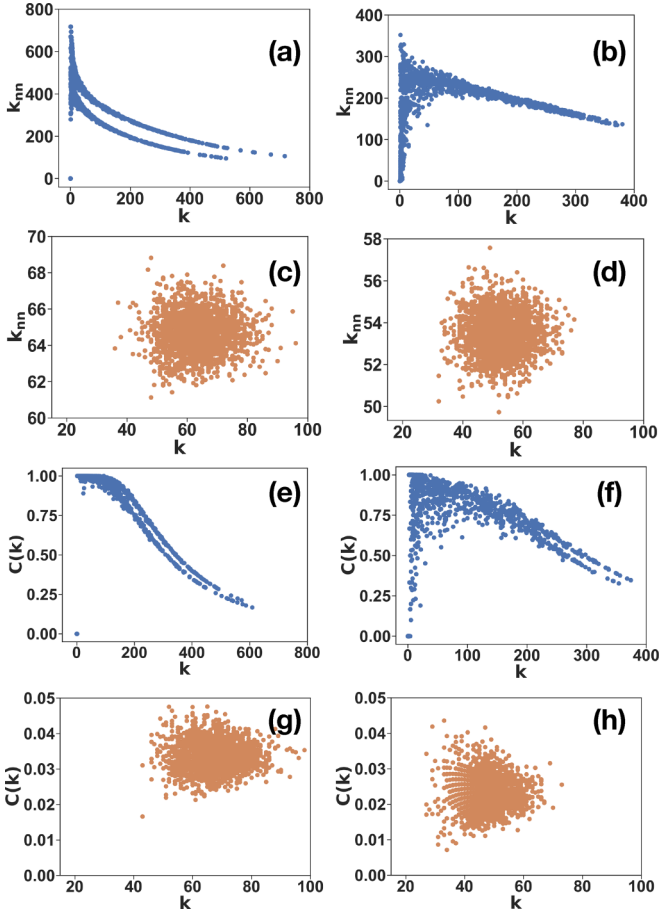


FIG. 13. Average nearest-neighbor degree  $k_{nn}$  and clustering coefficient  $C(k)$  on IsingNets obtained from Monte Carlo simulations of the spin system of linear size  $L = 40$  and corresponding null models with  $N = 2000$  nodes. The threshold distance of connecting two nodes is the fifth nearest-neighbor average distance. The random network is obtained by randomly permuting distances between node pairs and nodes are connected with the same threshold. The average neighbor degree  $k_{nn}$  is shown as a function of degree  $k$  on IsingNets and corresponding null models. Panel (a) and (b) show  $k_{nn}$  obtained at  $T = 2.12$  (a) and  $T = 2.35$  (b). Panel (c) and (d) show  $k_{nn}$  of the corresponding null model at  $T = 2.12$  (c) and  $T = 2.35$  (d). Panel (e) and (f) show  $C(k)$  obtained at  $T = 2.12$  (e) and  $T = 2.35$  (f). Panel (g) and (h) show  $C(k)$  of the corresponding null model at  $T = 2.12$  (g) and  $T = 2.35$  (h).

have a finite number of  $K$  cores also when the average degree diverges. Here we show that IsingNets have a very rich  $K$ -core structure having statistical properties that change below and above the critical temperature (see Fig. 15). Indeed above the critical temperature, we observe a behavior similar to the expected behavior for sparse scale-free networks with power-law exponent between two and three presenting a broad (seemingly power-law straight line on a log-log plot) decay of  $M(K)$  versus  $K$ . However, below the critical temperature where the average degree diverges, the  $K$  cores include more nodes while the decay of  $M(K)$  versus  $K$  is better approximated by an exponential (straight line in a log-linear plot) rather than by a power law.

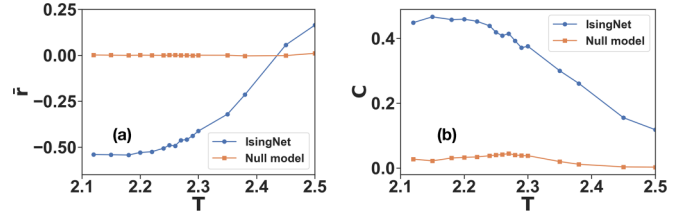


FIG. 14. Pearson correlation coefficient  $\bar{r}$  (a) and average clustering coefficient  $C$  (b) on IsingNets and corresponding null models formed at different temperatures. The Pearson correlation coefficient  $r$  is calculated on networks with  $10^4$  nodes and the average clustering coefficient  $C$  is calculated for IsingNet with  $N = 2000$  nodes coming from Monte Carlo simulation with linear size  $L = 40$ . The threshold distance of connecting two nodes is the fifth nearest-neighbor average distance. The random network is obtained by randomly permuting distances between node pairs and nodes are connected with the same threshold.

### E. Weight-topology correlations

Interestingly in weighted networks not only the network topology can reveal relevant degree correlations showing that the networks deviate from maximally random networks, but also the weights can be distributed in a nonrandom way. In particular, there are two main network analyses that are able to detect weight-topology correlations. The first analysis [81] involves studying the normalized strength  $s/k$  versus the degree  $k$  for each node of the network. If the weights are distributed randomly and independently on the degree of the two end nodes there should not be any significant dependence of  $s/k$  with  $k$ . Conversely if  $s/k$  increases with  $k$  it implies that nodes with higher degrees are incident in average to links with larger weights. The second analysis [88] investigates the weight-topology correlations aiming at revealing the weight heterogeneity among links that connect to the same node. This heterogeneity if any can be quantified by calculating the inverse participation ratio  $Y$  for the weights of the links ending to node  $i$ , defined as

$$Y_i = \sum_{j \sim i} \left( \frac{w_{ij}}{s_i} \right)^2. \quad (11)$$

If the weights  $w_{ij}$  of the links  $(i, j)$  incident to node  $i$  are homogeneous,  $Y_i \sim 1/k_i$ . If the weights are highly heterogeneous,  $Y_i^{-1}$  indicates the effective number of links with

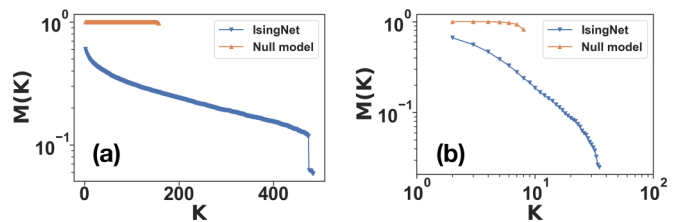


FIG. 15. Fraction of nodes in the  $K$ -core  $M(K)$  as a function of the core size  $K$  on IsingNet and null model with  $N = 10^4$  nodes at  $T = 2.12$  (a) and  $T = 2.50$  (b). The IsingNets are generated from 2D Ising model Monte Carlo simulations of the spin system of linear size  $L = 40$ . Panel (a) is shown with a linear-log scale, and panel (b) is shown with a log-log scale.

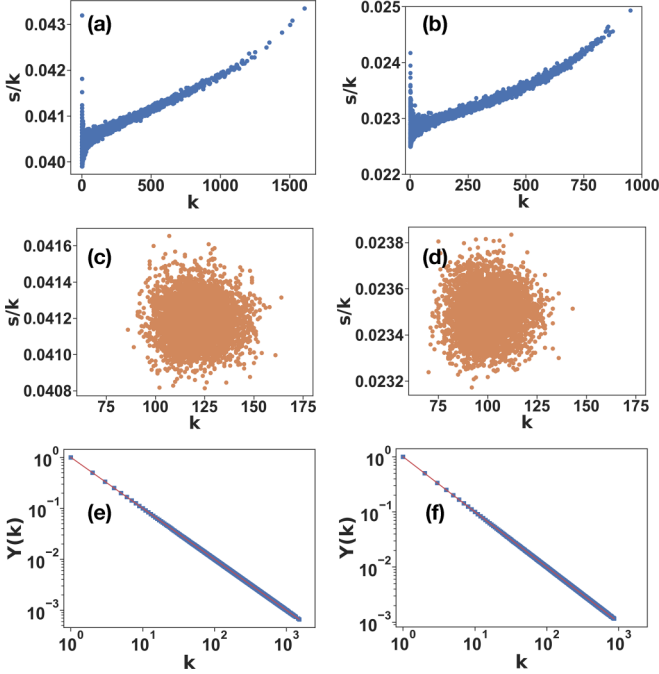


FIG. 16. The ratio of strength and degree and the inverse participation ratio on IsingNets obtained from Monte Carlo simulations of the spin system of linear size  $L = 40$  are shown versus degree on networks formed at different temperatures. The networks are formed by  $N = 5000$  samples. The threshold distance of connecting two nodes is the fifth nearest-neighbor average distance. The random networks are obtained by randomly permuting distances between node pairs and nodes are connected with the same threshold. The first row shows the strength degree ratio  $s/k$  versus degree  $k$  on IsingNets, and the second row shows which on corresponding random networks. The third row shows the inverse participation ratio  $Y(k)$  versus degree  $k$ . The networks are formed by simulation obtained at temperature  $T = 2.12$  (left column) and  $T = 2.35$  (right column).

significant weight. Interestingly when we measure  $Y_i$  for the IsingNets we observe that this second type of weight heterogeneity is missing in the data and that  $Y_i \sim 1/k_i$  indicating that for each node  $i$  the weights of the links incident to it have all weights of comparable order or magnitude (see Fig. 16).

### F. Spectral properties of Ising networks

The IsingNets do not only have very interesting combinatorial and statistical properties encoded in their highly correlated structure but display also relevant geometrical properties reflected in their spectrum. In particular, the critical IsingNets display nontrivial spectral properties characterized by a power-law scaling close to criticality and a highly degenerate spectral gap. The spectral properties of networks are usually probed by considering the spectrum of the graph Laplacian describing diffusion processes. The graph Laplacian  $\Delta$  is defined as  $\Delta = \mathbf{D} - \mathbf{A}$  where  $\mathbf{D}$  is the diagonal matrix whose diagonal elements are the degrees of the nodes (i.e.,  $D_{ii} = k_i$ ), and  $\mathbf{A}$  is the adjacency matrix of the network. The graph Laplacian  $\Delta$  is semidefinite positive and the spectrum always includes a zero eigenvalue with degeneracy given by the number of connected components of the network, i.e., given by the Betti number  $\beta_0$ . The smallest nonzero

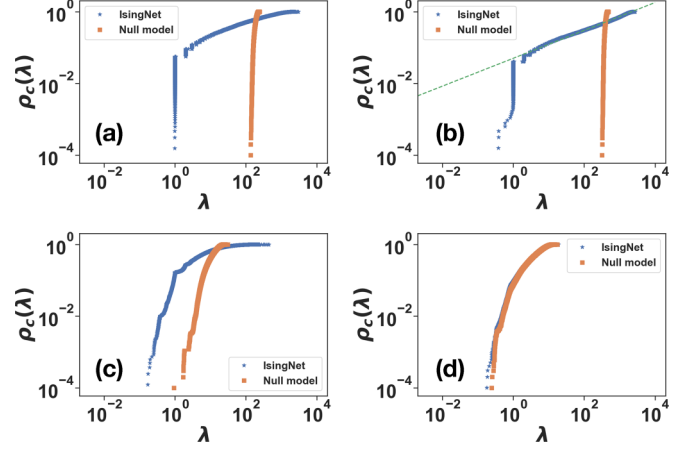


FIG. 17. Cumulative distribution  $\rho_c(\lambda)$  of the eigenvalues  $\lambda$  of graph Laplacian  $\Delta$  of IsingNets and corresponding null models at different temperatures  $T$ . The distribution is shown at  $T = 2.12$  (a),  $T = 2.27$  (b),  $T = 2.50$  (c), and  $T = 3.50$  (d). Data are shown for IsingNets generated from Monte Carlo simulations of spin system of linear size  $L = 40$ . In panel (b), the dashed lines indicate a power-law growth shown in Eq. (13) with exponent  $\hat{d} = 0.78 \pm 0.04$ .

eigenvalue of the graph Laplacian of a network is also called the Fiedler eigenvalue and is typically indicated as  $\lambda_2$  (as it is the second smallest eigenvalue in a connected network). In the literature, often one distinguishes between network models displaying a finite Fiedler eigenvalue  $\lambda_2 \rightarrow \lambda_2^* > 0$  in the limit  $N \rightarrow \infty$  and network models in which  $\lambda_2 \rightarrow 0$  as  $N \rightarrow \infty$ . In the first case, we say that the networks display a spectral gap whereas in the latter case, we say that the “spectral gap closes.” Examples of networks with finite spectral dimension are random graphs above the percolation threshold and examples of networks in which the spectral gap closes are finite dimensional lattices.

In several networks in which the spectral gap closes, it is possible to observe the spectral dimension  $d_S$  [6,89]. The spectral dimension  $d_S$  is the dimension perceived by diffusion processes on the networks encoded in the graph Laplacian. On a lattice, the spectral dimension coincides with the Euclidean dimension  $d$  of the lattice while on general network topology, the spectral dimension can be distinct from the Hausdorff dimension of the network. Interestingly also small world networks with infinite Hausdorff dimension can have a finite spectral dimension  $d_S \geq 2$  [6,90]. The spectral dimension determines the scaling of the cumulative density of the eigenvalues  $\rho_c(\lambda)$  of the graph Laplacian for  $\lambda \ll 1$  in networks where the spectral gap closes. In particular we have that networks with a spectral dimension  $d_S$  have a cumulative distribution  $\rho_c(\lambda)$  that obeys for  $\lambda \ll 1$ ,

$$\rho_c(\lambda) \simeq C\lambda^{d_S/2}, \quad (12)$$

where  $C$  is a constant. In Fig. 17 we show that the IsingNets display nontrivial spectral properties that have very peculiar characteristics strongly deviating from their corresponding null model. Particularly noticeable are the spectral properties of IsingNets close to the critical point where one observes the coexistence of a highly degenerate finite Fiedler eigenvalue  $\lambda_2$  with a power-law scaling of the cumulative

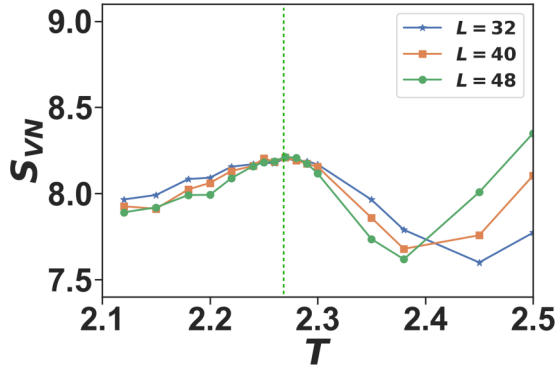


FIG. 18. The von Neumann entropy  $S_{VN}$  of different system size  $L$  is shown versus the temperature  $T$ . The IsingNets are formed by  $N = 10^4$  nodes. The dashed line indicates the critical temperature  $T_c$ .

distribution

$$\rho_c(\lambda) \simeq C\lambda^{\hat{d}/2}, \quad (13)$$

for  $\lambda > \lambda_2$  with a exponent given by  $\hat{d} \simeq 0.78 \pm 0.04$  for  $L = 40$ . Above the critical temperature, the degeneracy of the Fiedler eigenvalue is reduced and one observes a nontrivial spectrum reminiscent of the scale-dependent spectral dimension discussed in Ref. [91] within the critical region that at higher temperatures converges with the spectrum of the null model. Below the critical dimension, the spectral gap remains highly degenerate while the rest of the spectrum remains broadly distributed.

The spectrum of the graph Laplacian is also key to characterizing the network von Neumann entropy  $S_{VN}$  [92–94] defined as

$$S_{VN} = - \sum_{\lambda} \frac{\lambda}{\langle k \rangle N} \ln \left( \frac{\lambda}{\langle k \rangle N} \right). \quad (14)$$

The von Neumann entropy strongly departs from the von Neumann entropy of the null model for low temperatures displaying a local maximum for  $T = T_c$  (see Fig. 18).

An interesting open question that will be addressed in the following works is the relation between these spectral properties of the IsingNets graph Laplacians and the intrinsic dimension and the entropy measures that have been recently proposed starting from the unsupervised PCA analysis of spin systems [12,63].

## V. CONCLUSIONS

In this work we have launched a systematic network analysis of unsupervised learning of different states of matter. The analyzed IsingNets obtained from Monte Carlo simulations of the 2D Ising model are shown to reveal the statistical, combinatorial, geometrical, and topological organization of these networks. Through the paper, we have shown that true IsingNets are highly nonrandom by comparing their structural properties with the structural properties of the randomized counterparts. Importantly, we have also identified several indicators of the phase transition.

We have addressed the characterization of IsingNets following two different approaches. In the first approach, we

have studied the structure of IsingNets as the filtration parameter  $r$  is increased, which enforces an effective percolation process in which nodes are subsequently aggregated by considering connected pairs of nodes at increasing distances. This percolation process reveals the presence of two giant components in IsingNets in the ferromagnetic state, each one corresponding to configurations with different magnetization. The same filtration scheme has also been used here to study the topology of the data by constructing the clique complexes of the IsingNets and calculating their persistent diagram. Interestingly, the persistent diagrams reveal that real IsingNets are formed by compact clusters as the Betti numbers of their clique complex are strongly suppressed with respect to the clique complex of the corresponding randomized null models. This network analysis conducted across the filtration is also enriched by effective visualization of the network embedding conducted using MST and the UMAP embedding and by the statistical characterization of the distribution of closeness centralities.

Second, our investigation of IsingNets has been conducted by considering only links at a distance less than a threshold value  $r$  taken to be the average distance of the fifth nearest-neighbor nodes according to the (fully connected) distance matrix. These networks display a broad degree and strength distribution but their complexity extends well beyond the degree and strength distribution because IsingNets have strong degree-degree correlations and weights-degree correlations, and a rich core structure that changes significantly across the phase transition reflecting the highly nontrivial structure of the spin system that they describe.

This work opens new perspectives for the unsupervised characterization of the study of phases of matter using the tools of network science. This work can be extended in different directions. The analysis performed here for the 2D Ising model can be extended to the study of other classical critical phenomena with the goal of characterizing the possible presence or the lack of universalities among the networks constructed from spin system configuration snapshots. Similarly, the same toolbox can be utilized to attack out-of-equilibrium critical behavior. Widening the class of models where our analysis can lead to physical insight is fundamental to establish ground for possible analytical treatments. In particular, before this is done, a key point to be understood is the relation between correlation length and sampling in models in different universality classes. The reason is that a naive connection (number of samples is proportional to the correlation length resolved, modulo dimensional factors) is likely not correct, based on earlier manifold learning characterization of path integrals [95] (which did not observe such relation, at least in simple quantities such as the intrinsic dimension).

Another natural extension is to consider path integrals of quantum systems [95]. While the data structures of such objects might feature anisotropies due to the different roles of space and imaginary time correlations, they shall be equally amenable to the analysis discussed here. Moreover, single-sliced path integrals can also be represented as networks, as discussed in Ref. [56]: this last route provides a very promising venue for future investigation is the application of the proposed network science tools to study directly experimental data of many-body wave function snapshots.

## ACKNOWLEDGMENTS

We acknowledge useful discussions with M. Heyl, M. Schmitt, and T. Mendes-Santos. The work of M.D. was

partly supported by the MIUR Programme FARE (MEPH), by QUANTERA DYNAMITE PCI2022-132919, and by the PNRR MUR Project No. PE0000023-NQSTI. Nordita is supported in part by NordForsk.

- 
- [1] A.-L. Barabasi, *Network Science* (Cambridge University Press, Cambridge, 2016).
- [2] M. Newman, *Networks* (Oxford University Press, New York, 2018).
- [3] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).
- [4] E. Estrada, *The Structure of Complex Networks: Theory and Applications* (Oxford University Press, New York, 2012).
- [5] S. N. Dorogovtsev and J. F. Mendes, *The Nature of Complex Networks* (Oxford University Press, Oxford, 2022).
- [6] G. Bianconi, *Higher-Order Networks* (Cambridge University Press, Cambridge, 2021).
- [7] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to machine learning for physicists, *Phys. Rep.* **810**, 1 (2019).
- [8] L. Wang, Discovering phase transitions with unsupervised learning, *Phys. Rev. B* **94**, 195105 (2016).
- [9] W. Hu, R. R. P. Singh, and R. T. Scalettar, Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination, *Phys. Rev. E* **95**, 062122 (2017).
- [10] S. J. Wetzel, Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders, *Phys. Rev. E* **96**, 022140 (2017).
- [11] J. F. Rodriguez-Nieva and M. S. Scheurer, Identifying topological order through unsupervised machine learning, *Nat. Phys.* **15**, 790 (2019).
- [12] T. Mendes-Santos, X. Turkeshi, M. Dalmonte, and A. Rodriguez, Unsupervised learning universal critical behavior via the intrinsic dimension, *Phys. Rev. X* **11**, 011040 (2021).
- [13] A. Morningstar and R. G. Melko, Deep learning the Ising model near criticality, *J. Mach. Learn. Res.* **18**, 1 (2018).
- [14] K.-I. Aoki and T. Kobayashi, Restricted Boltzmann machines for the long range Ising models, *Mod. Phys. Lett. B* **30**, 1650401 (2016).
- [15] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, *Nat. Mach. Intell.* **1**, 155 (2019).
- [16] M. Benedetti, J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz, Quantum-assisted learning of hardware-embedded probabilistic graphical models, *Phys. Rev. X* **7**, 041052 (2017).
- [17] R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac, Restricted Boltzmann machines in quantum physics, *Nat. Phys.* **15**, 887 (2019).
- [18] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Ising model on networks with an arbitrary distribution of connections, *Phys. Rev. E* **66**, 016104 (2002).
- [19] G. Bianconi, Mean field solution of the Ising model on a Barabási–Albert network, *Phys. Lett. A* **303**, 166 (2002).
- [20] M. Leone, A. Vázquez, A. Vespignani, and R. Zecchina, Ferromagnetic ordering in graphs with arbitrary degree distribution, *Eur. Phys. J. B* **28**, 191 (2002).
- [21] H. C. Nguyen, R. Zecchina, and J. Berg, Inverse statistical problems: From the inverse Ising problem to data science, *Adv. Phys.* **66**, 197 (2017).
- [22] G. Bianconi, Superconductor-insulator transition on annealed complex networks, *Phys. Rev. E* **85**, 061113 (2012).
- [23] G. Bianconi, Superconductor-insulator transition in a network of 2D percolation clusters, *Europhys. Lett.* **101**, 26003 (2013).
- [24] R. T. C. Chepuri and I. A. Kovács, Complex quantum network models from spin clusters, *Commun. Phys.* **6**, 271 (2023).
- [25] A. Halu, L. Ferretti, A. Vezzani, and G. Bianconi, Phase diagram of the Bose-Hubbard model on complex networks, *Europhys. Lett.* **99**, 18001 (2012).
- [26] A. Halu, S. Garnerone, A. Vezzani, and G. Bianconi, Phase transition of light on complex quantum networks, *Phys. Rev. E* **87**, 022104 (2013).
- [27] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel type of phase transition in a system of self-driven particles, *Phys. Rev. Lett.* **75**, 1226 (1995).
- [28] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, *Proc. Natl. Acad. Sci. USA* **105**, 1232 (2008).
- [29] F. Morcos, A. Pagnani, B. Lunt, A. Bertolino, D. S. Marks, C. Sander, R. Zecchina, J. N. Onuchic, T. Hwa, and M. Weigt, Direct-coupling analysis of residue coevolution captures native contacts across many protein families, *Proc. Natl. Acad. Sci. USA* **108**, E1293 (2011).
- [30] T. Mora and W. Bialek, Are biological systems poised at criticality? *J. Stat. Phys.* **144**, 268 (2011).
- [31] J. Nokkala, F. Galve, R. Zambrini, S. Maniscalco, and J. Piilo, Complex quantum networks as structured environments: Engineering and probing, *Sci. Rep.* **6**, 26861 (2016).
- [32] J. Nokkala, F. Arzani, F. Galve, R. Zambrini, S. Maniscalco, J. Piilo, N. Treps, and V. Parigi, Reconfigurable optical implementation of quantum complex networks, *New J. Phys.* **20**, 053024 (2018).
- [33] I. Bonamassa, B. Gross, M. Laav, I. Volotsenko, A. Frydman, and S. Havlin, Interdependent superconducting networks, *Nat. Phys.* **19**, 1163 (2023).
- [34] M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna, A tool for filtering information in complex systems, *Proc. Natl. Acad. Sci. USA* **102**, 10421 (2005).
- [35] G. Bonanno, G. Caldarelli, F. Lillo, and R. N. Mantegna, Topology of correlation-based minimal spanning trees in real and model markets, *Phys. Rev. E* **68**, 046130 (2003).
- [36] M. A. Valdez, D. Jaschke, D. L. Vargas, and L. D. Carr, Quantifying complexity in quantum phase transitions via mutual

- information complex networks, *Phys. Rev. Lett.* **119**, 225301 (2017).
- [37] B. Sundar, M. A. Valdez, L. D. Carr, and K. R. A. Hazzard, Complex-network description of thermal quantum states in the Ising spin chain, *Phys. Rev. A* **97**, 052320 (2018).
- [38] B. Sokolov, M. A. C. Rossi, G. García-Pérez, and S. Maniscalco, Emergent entanglement structures and self-similarity in quantum spin chains, *Philos. Trans. R. Soc. A* **380**, 20200421 (2022).
- [39] A. A. Bagrov, M. Danilov, S. Brener, M. Harland, A. I. Lichtenstein, and M. I. Katsnelson, Detecting quantum critical points in the  $t - t'$  Fermi-Hubbard model via complex network theory, *Sci. Rep.* **10**, 20470 (2020).
- [40] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino, Topological strata of weighted complex networks, *PLoS ONE* **8**, e66506 (2013).
- [41] H. Edelsbrunner and J. L. Harer, *Computational Topology: An Introduction* (American Mathematical Society, Rhode Island, 2022).
- [42] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington, A roadmap for the computation of persistent homology, *EPJ Data Sci.* **6**, 17 (2017).
- [43] F. Vaccarino, U. Fugacci, and S. Scaramuccia, Persistent homology: A topological tool for higher-interaction systems, in *Higher-Order Systems*, edited by F. Battiston (Springer, Cham, 2022), pp. 97–139.
- [44] R. Ghrist, Barcodes: The persistent topology of data, *Bull. Am. Math. Soc.* **45**, 61 (2008).
- [45] I. Donato, M. Gori, M. Pettini, G. Petri, S. De Nigris, R. Franzosi, and F. Vaccarino, Persistent homology analysis of phase transitions, *Phys. Rev. E* **93**, 052138 (2016).
- [46] B. Olsthoorn, J. Hellsvik, and A. V. Balatsky, Finding hidden order in spin models with persistent homology, *Phys. Rev. Res.* **2**, 043308 (2020).
- [47] A. Cole, G. J. Loges, and G. Shiu, Quantitative and interpretable order parameters for phase transitions from persistent homology, *Phys. Rev. B* **104**, 104426 (2021).
- [48] Q. H. Tran, M. Chen, and Y. Hasegawa, Topological persistence machine of phase transitions, *Phys. Rev. E* **103**, 052127 (2021).
- [49] A. Tirelli and N. C. Costa, Learning quantum phase transitions through topological data analysis, *Phys. Rev. B* **104**, 235146 (2021).
- [50] N. Sale, J. Giansiracusa, and B. Lucini, Quantitative analysis of phase transitions in two-dimensional  $XY$  models using persistent homology, *Phys. Rev. E* **105**, 024121 (2022).
- [51] A. Tirelli, D. O. Carvalho, L. A. Oliveira, J. P. de Lima, N. C. Costa, and R. R. dos Santos, Unsupervised machine learning approaches to the  $q$ -state Potts model, *Eur. Phys. J. B* **95**, 189 (2022).
- [52] D. Spitz, J. Berges, M. Oberthaler, and A. Wienhard, Finding self-similar behavior in quantum many-body dynamics via persistent homology, *SciPost Phys.* **11**, 060 (2021).
- [53] D. Sehayek and R. G. Melko, Persistent homology of  $\mathbb{Z}_2$  gauge theories, *Phys. Rev. B* **106**, 085111 (2022).
- [54] N. Sale, B. Lucini, and J. Giansiracusa, Probing center vortices and deconfinement in  $SU(2)$  lattice gauge theory with persistent homology, *Phys. Rev. D* **107**, 034501 (2023).
- [55] D. Spitz, J. M. Urban, and J. M. Pawłowski, Confinement in non-Abelian lattice gauge theory via persistent homology, *Phys. Rev. D* **107**, 034506 (2023).
- [56] T. Mendes-Santos, M. Schmitt, A. Angelone, A. Rodriguez, P. Scholl, H. Williams, D. Barredo, T. Lahaye, A. Browaeys, M. Heyl, and M. Dalmonte, Wave function network description and Kolmogorov complexity of quantum many-body systems, [arXiv:2301.13216](https://arxiv.org/abs/2301.13216).
- [57] C. P. Herrero, Ising model in small-world networks, *Phys. Rev. E* **65**, 066110 (2002).
- [58] A. Pekalski, Ising model on a small world network, *Phys. Rev. E* **64**, 057104 (2001).
- [59] C. P. Herrero, Ising model in scale-free networks: A Monte Carlo simulation, *Phys. Rev. E* **69**, 067109 (2004).
- [60] C. P. Herrero, Ising model in clustered scale-free networks, *Phys. Rev. E* **91**, 052812 (2015).
- [61] S. Bradde, F. Caccioli, L. Dall'Asta, and G. Bianconi, Critical fluctuations in spatial complex networks, *Phys. Rev. Lett.* **104**, 218701 (2010).
- [62] L. Onsager, Crystal statistics. I. A two-dimensional model with an order-disorder transition, *Phys. Rev.* **65**, 117 (1944).
- [63] R. K. Panda, R. Verdel, A. Rodriguez, H. Sun, G. Bianconi, and M. Dalmonte, Non-parametric learning critical behavior in Ising partition functions: PCA entropy and intrinsic dimension, *SciPost Phys. Core* **6**, 086 (2023).
- [64] U. Wolff, Collective Monte Carlo updating for spin systems, *Phys. Rev. Lett.* **62**, 361 (1989).
- [65] U. Wolff, Comparison between cluster Monte Carlo algorithms in the Ising model, *Phys. Lett. B* **228**, 379 (1989).
- [66] R. L. Graham and P. Hell, On the history of the minimum spanning tree problem, *Ann. History Comput.* **7**, 43 (1985).
- [67] L. McInnes, J. Healy, and J. Melville, UMAP: Uniform manifold approximation and projection for dimension reduction, [arXiv:1802.03426](https://arxiv.org/abs/1802.03426).
- [68] A. Bavelas, Communication patterns in task-oriented groups, *J. Acoust. Soc. Am.* **22**, 725 (1950).
- [69] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Critical phenomena in complex networks, *Rev. Mod. Phys.* **80**, 1275 (2008).
- [70] M. Li, R.-R. Liu, L. Lü, M.-B. Hu, S. Xu, and Y.-C. Zhang, Percolation on complex networks: Theory and application, *Phys. Rep.* **907**, 1 (2021).
- [71] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Resilience of the Internet to random breakdowns, *Phys. Rev. Lett.* **85**, 4626 (2000).
- [72] O. Bobrowski and P. Skraba, A universal null-distribution for topological data analysis, *Sci. Rep.* **13**, 12274 (2023).
- [73] A. Baptista, R. J. Sánchez-García, A. Baudot, and G. Bianconi, Zoo guide to network embedding, *J. Phys. Complex.* **4**, 042001 (2023).
- [74] H. Seyed-allaei, G. Bianconi, and M. Marsili, Scale-free networks with an exponent less than two, *Phys. Rev. E* **73**, 046113 (2006).
- [75] O. T. Courtney and G. Bianconi, Dense power-law networks and simplicial complexes, *Phys. Rev. E* **97**, 052303 (2018).
- [76] F. Caron and E. B. Fox, Sparse graphs using exchangeable random measures, *J. R. Stat. Soc. B: Stat. Methodol.* **79**, 1295 (2017).

- [77] G. Timar, S. N. Dorogovtsev, and J. F. F. Mendes, Scale-free networks with exponent one, *Phys. Rev. E* **94**, 022302 (2016).
- [78] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286**, 509 (1999).
- [79] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47 (2002).
- [80] I. Voitalov, P. van der Hoorn, R. van der Hofstad, and D. Krioukov, Scale-free networks well done, *Phys. Rev. Res.* **1**, 033034 (2019).
- [81] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Characterization and modeling of weighted networks, *Physica A* **346**, 34 (2005).
- [82] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, The architecture of complex weighted networks, *Proc. Natl. Acad. Sci. USA* **101**, 3747 (2004).
- [83] D. J. Watts and S. H. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature (London)* **393**, 440 (1998).
- [84] G. Bianconi and M. Marsili, Effect of degree correlations on the loop structure of scale-free networks, *Phys. Rev. E* **73**, 066127 (2006).
- [85] J. Alvarez-Hamelin, L. Dall’Asta, A. Barrat, and A. Vespignani, Large scale networks fingerprinting and visualization using the  $k$ -core decomposition, *Adv. Neural Inf. Process. Syst.* **18** (2005).
- [86] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, A model of internet topology using  $k$ -shell decomposition, *Proc. Natl. Acad. Sci. USA* **104**, 11150 (2007).
- [87] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes,  $k$ -core organization of complex networks, *Phys. Rev. Lett.* **96**, 040601 (2006).
- [88] E. Almaas, B. Kovacs, T. Vicsek, Z. Oltvai, and A.-L. Barabási, Global organization of metabolic fluxes in the bacterium *Escherichia coli*, *Nature (London)* **427**, 839 (2004).
- [89] R. Burioni and D. Cassi, Random walks on graphs: Ideas, techniques and results, *J. Phys. A: Math. Gen.* **38**, R45 (2005).
- [90] J. D. Correia and J. F. Wheeler, The spectral dimension of non-generic branched polymer ensembles, *Phys. Lett. B* **422**, 76 (1998).
- [91] J. Ambjørn, J. Jurkiewicz, and R. Loll, The spectral dimension of the universe is scale dependent, *Phys. Rev. Lett.* **95**, 171301 (2005).
- [92] K. Anand and G. Bianconi, Entropy measures for networks: Toward an information theory of complex topologies, *Phys. Rev. E* **80**, 045102(R) (2009).
- [93] K. Anand, G. Bianconi, and S. Severini, Shannon and von Neumann entropy of random networks with heterogeneous expected degree, *Phys. Rev. E* **83**, 036109 (2011).
- [94] M. De Domenico, V. Nicosia, A. Arenas, and V. Latora, Structural reducibility of multilayer networks, *Nat. Commun.* **6**, 6864 (2015).
- [95] T. Mendes-Santos, A. Angelone, A. Rodriguez, R. Fazio, and M. Dalmonte, Intrinsic dimension of path integrals: Data-mining quantum criticality and emergent simplicity, *PRX Quantum* **2**, 030332 (2021).