Chaos

Characterization and comparison of large directed networks through the spectra of the magnetic Laplacian

Cite as: Chaos **30**, 073141 (2020); https://doi.org/10.1063/5.0006891
Submitted: 07 March 2020 . Accepted: 29 June 2020 . Published Online: 27 July 2020

Bruno Messias F. de Resende, and Luciano da F. Costa

Chaos ARTICLE scitation.org/journal/cha

Characterization and comparison of large directed networks through the spectra of the magnetic Laplacian

Cite as: Chaos 30, 073141 (2020); doi: 10.1063/5.0006891
Submitted: 7 March 2020 · Accepted: 29 June 2020 ·
Published Online: 27 July 2020

View Online Export Citation CrossMark

Bruno Messias F. de Resende[a] and Luciano da F. Costa

AFFILIATIONS

Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, SP 13566-590, Brazil

[a]**Author to whom correspondence should be addressed:** messias.physics@gmail.com

ABSTRACT

In this paper, we investigated the possibility of using the magnetic Laplacian to characterize directed networks. We address the problem of characterization of network models and perform the inference of the parameters used to generate these networks under analysis. Many interesting results are obtained, including the finding that the community structure is related to rotational symmetry in the spectral measurements for a type of stochastic block model. Due the hermiticity property of the magnetic Laplacian we show here how to scale our approach to larger networks containing hundreds of thousands of nodes using the Kernel Polynomial Method (KPM), a method commonly used in condensed matter physics. Using a combination of KPM with the Wasserstein metric, we show how we can measure distances between networks, even when these networks are directed, large, and have different sizes, a hard problem that cannot be tackled by previous methods presented in the literature.

The Laplacian operator of a directed network is not Hermitian. This property hampers the interpretation of the spectral measurements and restricts the use of computational methods developed in network science. In this work, we propose a framework and novel measurements based on the spectrum of the magnetic Laplacian to study directed networks. By using the properties of circulant matrices, we show analytically that novel measurements are able to grasp information about the structure of directed networks. It shows that the number of modular structures in networks is related to the rotational symmetry of the spectrum and, therefore, can contribute to characterize the parameters of the directed networks. To infer the generative parameters of networks, we propose the application of the Wasserstein metric to measure the distance between the spectra of the magnetic Laplacian, allowing networks to be compared. All the proposed methods depend on the diagonalization of the magnetic Laplacian operator, which implies a high computational cost. Therefore, the calculations can become unfeasible. To overcome this limitation, we implemented the Kernel Polynomial Method (KPM) using TensorFlow package. This method approximates the spectrum density with a lower computational cost, allowing the spectral characterization of large networks containing hundreds of thousands of nodes.

I. INTRODUCTION

In the seminal work *Can one hear the shape of a drum?*,[1] Mark Kack discusses the relationship between a membrane and the set of eigenvalues (spectrum) of the Laplacian operator. However, this relationship was identified to be not unique,[2] in the sense that two distinct membranes (non-isometric manifolds) can have the same spectrum. Nevertheless, despite such degeneracies, spectral information can provide valuable insights about the real world. For instance, spectral geometry has been used to study physical phenomena such as quantum gravity[3] and provided the basis for developing algorithms in computer science.[4]

Although the analysis of continuous regions such as those considered by Kack remains an interesting issue, several phenomena in nature and society need to be modeled in terms of discrete structures such as networks. In this case, we can adapt Kack's question as *Can one hear the shape of a network?* The answer to this question is analogous to what has been verified for the original question, i.e., two nonisomorphic networks can share the same spectrum.[5,6] Despite such a limitation, the spectral approach to discrete structures can still be useful in some practical and theoretical problems.[7,8] An example of a spectral approach that has been applied to characterize networks is the von-Neumann entropy.[9–11]

More recently, the concept of entropy of a graph has been used to measure the similarity between two given networks.[12] Examples of this approach include the entropic similarity applied to the inference of parameters of network models.[12,13] However, this measure cannot be immediately extended to directed networks and, as has been shown in Ref. 14, the directed edges have substantial implications in dynamics on graphs.

In addition, the entropic similarity depends on the product of the matrices associated with the given networks. This implies that this similarity measurement is not invariant with respect to permutations of the indices associated with the nodes. Given these dependencies, such measurements are not well defined when the nodes cannot be associated with fixed indices.

Given that directed networks can accurately model several real-world problems, it is essential to develop new methodologies capable of dealing with network directionality. An immediate difficulty implied by this scenario is that the associated Laplacian operator will often have complex values, because the adjacency matrix associated with the directed networks is non-symmetric. A promising approach to address this problem consists of studying directed complex networks while considering their magnetic Laplacian operator.[15–17] As an example, in Ref. 16, the authors showed that the results of community detection algorithms could be improved by considering the magnetic Laplacian associated with the directed network.

In this work, we show that the magnetic Laplacian approach can be used to characterize complex networks, including those with hundreds of thousands of nodes. By characterization, we mean that measurements taken from this operator contribute to identify the network model responsible for generating a given network, as well as performing the inference of parameters responsible for generating a given specific network configuration. Several results were obtained. First, for simpler models (i.e., modular regular networks), the number of modular structures is related to the specific heat rotational symmetry. Subsequently, we showed that these spectral measurements, combined with the Wasserstein distance between spectral densities,[18–20] can provide valuable contributions to infer the original parameters used for getting those networks, with relative errors smaller than 1%.

II. METHODS

A. Magnetic Laplacian

A directed network can be expressed by a tuple $G = (V, E, w)$, where V is the set of vertices, and $|V|$ stands for the number of the vertices; E is the set of edges such that for each $u, v \in V$ the ordered tuple $e = (u, v) \in E$ assigns a directed edge from vertex u to v and $w : E \to \mathbb{R}$. A directed network can be associated with an undirected counterpart $G^{(s)} = (V, E^{(s)}, w^{(s)})$, where $w^{(s)}(u, v) = \frac{w(u,v)+w(v,u)}{2}$. However, the directionality of G is lost in $G^{(s)}$.

In order to preserve the hermiticity and the information about directionality , define γ, as $\gamma : E \to \mathcal{G}$, where $\mathcal{G}$ is a group, such that $\gamma(u, v)^{-1} = \gamma(v, u)$, choosing $\mathcal{G} = U(1)$ and expressing γ as

$$\gamma_q(u, v) = \exp(2\pi i q f(u, v)), \tag{1}$$

where $q \in [0, 1]$ and $f(u, v) = w(u, v) - w(v, u)$ represents the flow in a given vertex u due to another vertex v.

The symmetric network equipped with γ_q has information about directed edges and, at the same time, the adjacency matrix is Hermitian.

Now, we consider the following operator, associated with $(G^{(s)}, \gamma_q)$, where $\odot$ is the Hadamard product:

$$\mathbf{L}_q = \mathbf{D} - \mathbf{\Gamma}_q \odot \mathbf{W}^{(s)}, \tag{2}$$

where $\mathbf{D}$ is the degree matrix that contains the node degrees along its main diagonal; $[\mathbf{\Gamma}_q]_{u,v} = [\mathbf{\Gamma}_q^\dagger]_{v,u} = \gamma_q(u, v)$ and $[W^{(s)}]_{u,v} = [W^{(s)}]_{v,u} = w^{(s)}(u, v)$.

It is interesting to observe that this operator corresponds to the magnetic Laplacian,[15,22] L_q. The reason for the term magnetic is that the operator can be used to describe the phenomenology of a quantum particle subject to the action of a magnetic field.[23] Due to this physical context, the parameter q is named charge.

By construction, $\mathbf{D}$ and $\mathbf{W}^{(s)}$ are both symmetric and $\mathbf{\Gamma}_q$ is Hermitian. Consequently, $\mathbf{L}_q$ is Hermitian. In addition, it is sometimes convenient to use a normalized version of $\mathbf{L}_q$, which is given by

$$\mathbf{H}_q = \sqrt{\mathbf{D}^{-1}}\mathbf{L}_q\sqrt{\mathbf{D}^{-1}}, \tag{3}$$

where the $\mathbf{H}_q$ is defined only if the network is at least weakly connected.

A given eigenvector of $\mathbf{H}_q$, $|\psi_{l,q}\rangle \in \mathbb{C}^{|V|}$, can be obtained as the solution of

$$\mathbf{H}_q|\psi_{l,q}\rangle = \lambda_{l,q}|\psi_{l,q}\rangle, \tag{4}$$

where $\lambda_{l,q} \in \mathbb{R}$ and $\lambda_{1,q} \le \lambda_{l,q} \le \cdots \le \lambda_{|V|,q}$.

It is possible to enhance the analogy with physical systems by including a temperature parameter $T \in \mathbb{R}_+$. By using this parameter, the network properties can be studied from the statistical mechanics viewpoint.

Here, we adopted the Boltzmann–Gibbs statistical mechanics formulation as a means to associate the partition function

$$Z(T, q) = \sum_{l=1}^{|V|} e^{-\frac{\lambda_{l,q}}{T}} \tag{5}$$

with G.

By using Eq. (5), the expected value at temperature T of an operator O can be expressed in terms of its eigenvalues $\{o_l\}$ as

$$\langle O\rangle = \frac{1}{Z(T, q)} \sum_{l=1}^{|V|} e^{-\frac{\lambda_{l,q}}{T}} o_l. \tag{6}$$

In this work, we use Eq. (6) to define the measure of specific heat, c_λ, associated with a network. This novel measurement is given by

$$c_\lambda(q, T) = \frac{\langle H_q^2\rangle - \langle H_q\rangle^2}{T^2}. \tag{7}$$

Equation (7) has two free parameters, namely, q and T. Because of this free choice of parameters and, owing to the fact that we have a rotation associated with directed edges (γ_q), we plot c_λ in two dimensions, setting $2\pi q$ as the polar coordinate and T as the radial one. Regarding the interpretation and justification of physics-related quantities such as the specific heat, it is directly related to the variance of the eigenvalue spectrum. As a consequence, that quantity provides a signature of the spectrum properties, contributing to the characterization of the network structure.

B. Directed modular networks

In this work, we resort to a type of Stochastic Block Model (SBM) in order to obtain a good control of the network properties such as community size and also because of its potential for facilitating analytical studies. The adopted stochastic block model networks were obtained as follows:

1. Split the set V onto N_f equal-size sets $(f_1, f_2, \ldots, f_{N_f})$.
2. For each $u, v \in f_i$ create a directed edge (u, v) with probability p_c.
3. For each $u \in f_i$ and a $v \in f_{i+1}$ (assuming $f_{N_f+1} = f_1$), create a directed edge (u, v) with probability p_d.

C. Spectral entropy of directed networks

Recent works reported how to use entropic measurements to quantify the similarity between two undirected networks.[12,13] The entropy of a network is derived from the usual Laplacian spectrum (all eigenvalues are real). By contrast, these measurements cannot be used in the case of directed networks because the adjacency matrix is not Hermitian. However, the magnetic Laplacian methodology yields a Hermitian operator H_q, which is here used to define an entropic measurement for directed networks.

Recall that a quantum system at finite temperature, T, is defined by its respective density matrix, $\rho(T)$.[24] For a network G and charge q, this operator can be expressed in terms of the eigenvalues and eigenvectors associated to H_q as

$$\boldsymbol{\rho}_q(T)\frac{1}{Z(T,q)}\sum_{l=1}^{|V|} e^{-\frac{\lambda_{q,l}}{T}}|\psi_{l,q}\rangle\langle\psi_{l,q}|. \tag{8}$$

The previously defined density matrix can be used in order to define measurements associated with a directed (or undirected) network. For instance, by using the previous definition, the concepts of spectral entropy of a network can be extended for the directed case by using the following equation:

$$S(G,q,T) = \mathrm{Tr}\left[\boldsymbol{\rho}_q(T)\log\boldsymbol{\rho}_q(T)\right], \tag{9}$$

where log is the matrix logarithm and Tr corresponds to the trace operation.

Given the definition of spectral entropy, we can extend the entropic dissimilarity between two directed networks, $\tilde{G}$ and G, as

$$S_d(\tilde{G},G,q,T) = S(\tilde{G},q,T) - \mathrm{Tr}\left[\tilde{\boldsymbol{\rho}}_q(T)\mathrm{Log}\boldsymbol{\rho}_q(T)\right]. \tag{10}$$

However, as can be noted, the term within the trace depends on the product of distinct matrices. Thus, even if two networks presented in Fig. 2 are isomorphic, the measure of entropic dissimilarity is nonzero when it is desirable to be null. Another issue related with the entropic similarity approach is that measure cannot be used to compare networks with different numbers of nodes. This is interesting also because the largest weakly connected component generated by a model does not necessarily have the same size as the overall number of nodes. At the same time, such a measure has a high computational cost.

In this work, we suggest the application of the kernel polynomial method jointly with the Wasserstein metric in order to quantify

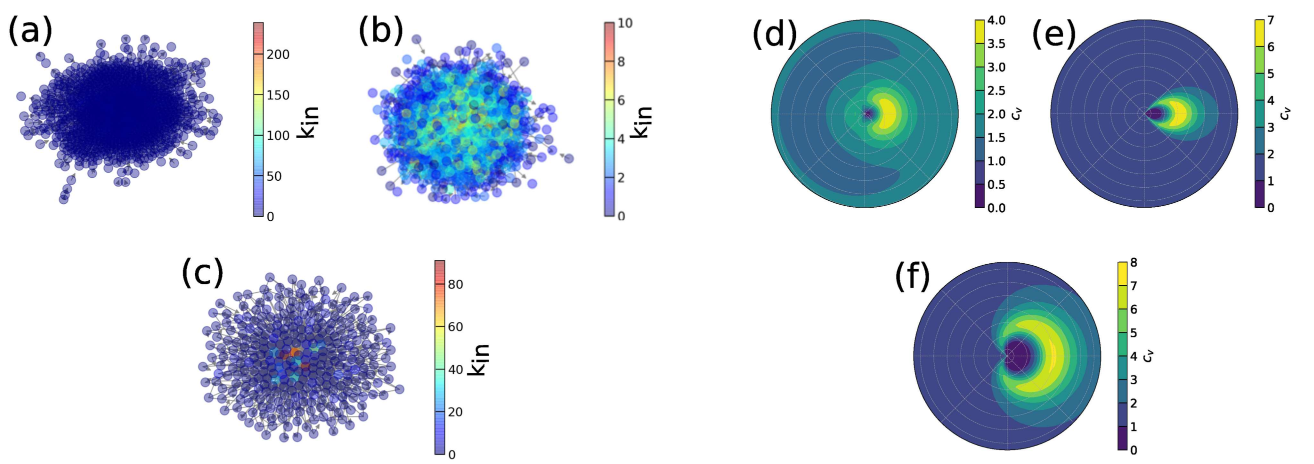

FIG. 1. In (a), (b), and (c), we have SF, ER, and BA networks, respectively. In the color maps, k_{in} is the indegree of a given node. In (d), (e), and (f), the specific heat in terms of the charge $2\pi q$ (polar coordinates) and temperature (radial coordinate) for a scale-free network of Bollobás *et al.*,[21] ER, and BA network, respectively, is shown. The parameters used to generate those networks were $|V| = 1000$; the edge probability for ER was $p = 0.003$; the number of outgoing edges for BA network was $m = 3$. The temperature range and charge are uniformly sampled from intervals [0.01, 0.15] and [0, 1/2] with 30 points each. As can be noted, c_λ shows a specific pattern for each network. This *fingerprint pattern* for each network explains why the SOM (self-organization map) was so successful in the task of organizing networks belonging to the same classes onto the same groups using only the specific heat, without any knowledge about that classes. It follows from Eq. (1) that the eigenvalues and therefore c_λ are symmetric with respect to the addition of integer values to the charge $\gamma_q = \gamma_{q+j}\forall j \in \mathbb{Z}$, reflecting in the bilateral symmetry with respect to the horizontal axis in (d), (e), and (f).

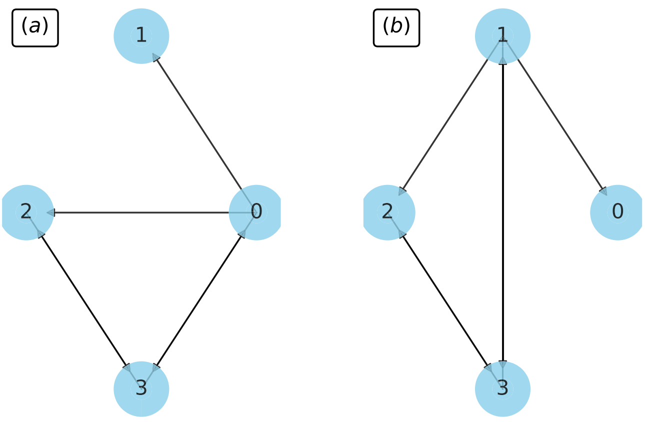

FIG. 2. The networks in (a) and (b) are isomorphic in the sense that they can be mapped one into the other by changing the indexes $0 \to 1, 1 \to 0$.

the dissimilarity between the directed networks. It should be emphasized that this combination of approaches is only possible given that the magnetic Laplacian is a Hermitian operator.

D. Comparison of large directed networks: The KPM method and the Wasserstein metric

In order to compute the spectral distance between two networks, it is necessary to compute the spectral density,

$$\rho_q(\lambda) = \frac{1}{|V|}\sum_{l=1}^{|V|}\delta(\lambda - \lambda_{l,q}), \tag{11}$$

which has complexity order $O(|V|^3)$. As such, this approach becomes unfeasible for larger networks ($|V| > 10^5$). Fortunately, the magnetic Laplacian matrix is Hermitian and often sparse, so that the method known as kernel polynomial (KPM) can be considered[25] for estimating the ρ_q.

The KPM objective consists in calculating a simplex $\{\mathbf{p} \in \mathbb{R}^n_+ : \sum_{i=1}^n p_i = 1\}$, which allows to define a discrete measure

$$\alpha_q = \sum_{i=1}^{n} p_{i,q}\delta_{\lambda_{i,q}} \tag{12}$$

that approximates Eq. (11) with enough accuracy.

This method is based on two approximations. The first is that any continuous real function in an interval $[-1, 1]$ can be expanded in terms of Chebyshev polynomials,[26] allowing the spectral density to be approximated by n terms. The second approximation consists in evaluating the traces associated with the terms of that expansion using Hutchinson's approach.[27] In essence, the trace of a sparse matrix function can be approximated by the product of this function by a set of random vectors. The oscillations induced by these approximations can be smoothed by subsequently applying a known kernel, which, in the present work, corresponds to the Jackson kernel.[28] Therefore, the KPM allows the spectrum of magnetic Laplacian to be estimated by using algorithms with near-linear computational cost. In this way, it becomes possible to estimate the spectral density and measurements such as entropy and specific heat even in the case of very large networks containing millions of nodes.

Given that it is possible to effectively estimate the magnetic Laplacian spectral density, we can employ Wasserstein metric in order to define distances between directed networks.

For instance, let the set of admissible couplings of two probability distributions α_q and $\tilde{\alpha}_q$, be given as

$$U(\alpha, \tilde{\alpha}) = \{\mathbf{U} \in \mathbb{R}_+^{|V|\times|\tilde{V}|} : \mathbf{U}\mathbf{1}_{|\tilde{V}|} = \mathbf{p}, \quad \mathbf{U}^T\mathbf{1}_{|V|} = \tilde{\mathbf{p}}\}. \tag{13}$$

For a $d \geq 1$, a d-Wasserstein distance between the two measures is given by

$$W_d(p, \tilde{p}, q) = \left(\min_{P\in U(\alpha,\tilde{\alpha})}\left[\sum_{i,j}|\lambda_{i,q} - \tilde{\lambda}_{j,q}|^d U_{i,j}\right]\right)^{1/d}. \tag{14}$$

This function has several desired characteristics, such as: it is a metric, it can be applied to networks with different numbers of nodes, and it has relaxed implementations that allow the distance value to be obtained with a smaller computational cost.

The task of estimating the value of a parameter used to generate a given network, such as the connecting probability in the Erdős–Rényi (ER) model, can be approached by seeking for a minimum Wasserstein distance between the original network and a set of n_{exp} networks synthesized by considering several parameters. In this work, we chose a set of n_q charges from which the magnetic Laplacians of each candidate network is obtained and then KPM is used to obtain the respective spectra, and the minimal distance between the latter and the original is determined by using the Wasserstein distance

$$\langle W_d\rangle(p) = \frac{1}{n_{exp}}\sum_{p\in P}\left(\frac{1}{n_q}\sum_{q\in Q}W_d(p, \tilde{p}, q)\right). \tag{15}$$

III. RESULTS

A. Community structures in network and spectral symmetries

As a first step to address the problem of characterizing directed complex networks by using the magnetic Laplacian formalism, we derive some analytic and numerical results relating to the network structure and the spectrum of the magnetic Laplacian operator.

First, we aim at studying the influence of community structure in directed networks on the magnetic Laplacian spectrum and, consequently, on the specific heat, c_λ. We assume that the connections within the communities, $\mathbf{W_{in}}$, as well as between the communities, $\mathbf{W_{out}}$, are not differentiated between the structures. Under this hypothesis, the adjacency matrix can be organized as follows, assuming N_f communities (henceforth, we take $N_f > 2$):

$$\mathbf{W} = \begin{bmatrix} \mathbf{W_{in}} & \mathbf{W_{out}} & \mathbf{0}_{N_c} & \cdots & \mathbf{0}_{N_c} \\ \mathbf{0}_{N_c} & \mathbf{W_{in}} & \mathbf{W_{out}} & \cdots & \mathbf{0}_{N_c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{W_{out}} & \mathbf{0}_{N_c} & \mathbf{0}_{N_c} & \cdots & \mathbf{W_{in}} \end{bmatrix}, \tag{16}$$

where $\mathbf{0}_{N_c}$ is a null matrix $N_c \times N_c$. For generality's sake, $\mathbf{W_{in}}$ and $\mathbf{W_{out}}$ can be constructed in an arbitrary form.

The magnetic Laplacian expressed as discussed above has the following organization:

$$\mathbf{H}_q = \begin{bmatrix} \mathbf{H_{in}} & \mathbf{H_{out}} & \mathbf{0}_{N_c} & \cdots & \mathbf{H_{out}}^\dagger \\ \mathbf{H_{out}}^\dagger & \mathbf{H_{in}} & \mathbf{H_{out}} & \cdots & \mathbf{0}_{N_c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{H_{out}} & \mathbf{0}_{N_c} & \mathbf{0}_{N_c} & \cdots & \mathbf{H_{in}} \end{bmatrix}. \tag{17}$$

Note that this matrix is circulant, i.e.,

$$\mathbf{H}_q = \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_{N_f-1} \\ \mathbf{h}_{N_f-1} & \mathbf{h}_0 & \cdots & \mathbf{h}_{N_f-2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_0 \end{bmatrix}. \tag{18}$$

Observe that $\mathbf{H}_q$ is a specific case of a Toepltiz matrix[29] so that the eigenvalues can be obtained considering the property that all the columns in the original matrix can be expressed as cyclic permutations of the first column.

Our objective now is to find the set $\{\lambda_u\}$ such that

$$\mathbf{H}_q|\psi_u\rangle = \lambda_u|\psi_u\rangle. \tag{19}$$

As known from the literature,[29] the eigenvectors of a cyclic matrix can be obtained as

$$|\psi_u\rangle = \begin{bmatrix} |\phi\rangle \\ \rho_u|\phi\rangle \\ \vdots \\ \rho_u^{N_f-1}|\phi\rangle \end{bmatrix}, \tag{20}$$

where $u \in \{0, \ldots, N_f - 1\}$ and $\rho_u = \rho^\star_{N_f-u} = \exp(\frac{2\pi i u}{N_f})$. Substituting this eigenvector equation (20) into Eq. (19) allows the block equation induced by the first row to be solved as

$$\tilde{\mathbf{H}}_u|\psi_u\rangle = \sum_{l=0}^{N_f-1} \mathbf{h}_l \rho_{l\cdot u}|\psi_u\rangle = \lambda_u|\psi_u\rangle. \tag{21}$$

The above equation can be simplified introducing the variable

$$m_f = \begin{cases} \dfrac{N_f+1}{2} & \text{if } N_f \text{ is odd,} \\ \dfrac{N_f}{2} & \text{if } N_f \text{ is even,} \end{cases} \tag{22}$$

and by taking into account that $\mathbf{H_N}$ is Hermitian consequently $\mathbf{h}_j = \mathbf{h}^\dagger_{N_f-j}$.

The simplified version is given as

$$\tilde{\mathbf{H}}_u = \mathbf{h}_0 + \sum_{l=1}^{m_f-1}\left(\mathbf{h}_l\rho_{l\cdot u} + \mathbf{h}_l^\dagger \rho^\star_{l\cdot u}\right) + \mathbf{\Delta}, \tag{23}$$

where

$$\mathbf{\Delta} = \begin{cases} \mathbf{0}_{N_c} & \text{if } N_f \text{ is odd,} \\ (-1)^u \mathbf{h}_{m_f} & \text{if } N_f \text{ is even.} \end{cases} \tag{24}$$

Since in the flow structure $\mathbf{\Delta} = \mathbf{0}_{N_c}$, and only three instances $\mathbf{h}_u$ are non-null, we have

$$\tilde{\mathbf{H}}_u = \mathbf{h}_0 + \mathbf{h}_1\rho_u + \mathbf{h}_1^\dagger \rho_u^\star. \tag{25}$$

Replacing the operators $\mathbf{h}$ by their respective counterparts in Eq. (17), we obtain the following expression for the uth matrix in a network with N_f blocks:

$$\tilde{\mathbf{H}}_u = \mathbf{H_{in}} + e^{\frac{2\pi i u}{N_f}}\mathbf{H_{out}} + e^{-\frac{2\pi i u}{N_f}}\mathbf{H_{out}}^\dagger. \tag{26}$$

In the following sections, we will investigate how distinct $\mathbf{H_{in}}$ influence c_λ.

1. Uniform connections

Uniform connection is characterized by having the degree of each vertex given as $[\mathbf{D}_{ii}] = d = 2N_c - 1$. Consequently, the intrablock of the magnetic Laplacian is

$$\mathbf{H_{in}} = \frac{\mathbf{I}_{N_c}(1+d) - \mathbf{1}_{N_c}}{d}, \tag{27}$$

and the interblock defining the connections between the modular structures is given as

$$\mathbf{H_{out}} = -\frac{\exp(2\pi i q)}{2d}\mathbf{1}_{N_c}. \tag{28}$$

Substituting the two previous equations into Eq. (26), $\tilde{\mathbf{H}}_u$ can be obtained as

$$\tilde{\mathbf{H}}_u = \frac{\mathbf{I}_{N_c}(1+d) - \mathbf{1}_{N_c}}{d} - 2\frac{\cos(2\pi(\frac{u}{N_f} - q))}{2d}\mathbf{1}_{N_c}. \tag{29}$$

Observe that $\tilde{\mathbf{H}}_u$ is a circulant matrix. Due to this, let $v \in \{0, \ldots, N_c - 1\}$, and define

$$m_c = \begin{cases} \dfrac{N_c+1}{2} & \text{if } N_c \text{ is odd,} \\ \dfrac{N_c}{2} & \text{if } N_c \text{ is even.} \end{cases} \tag{30}$$

The eigenvalues of $\tilde{\mathbf{H}}_u$ can be obtained as

$$\lambda_{u,v} = h_0 + \sum_{l=1}^{m_c-1}\left(h_l\rho_{l\cdot v} + h_l^\dagger \rho^\star_{l\cdot v}\right) + \Delta, \tag{31}$$

where

$$\Delta = \begin{cases} 0 & \text{if } N_c \text{ is odd,} \\ (-1)^v h_{m_c} & \text{if } N_c \text{ is even.} \end{cases} \tag{32}$$

Replacing h_l by their counterparts in Eq. (31), the following eigenvalue equation can be obtained:

$$\lambda_{u,v} = 1 - \frac{\cos(2\pi(\frac{u}{N_f} - q))}{d} + \frac{2}{d}\left(1 + \cos\left(2\pi\left(\frac{u}{N_f} - q\right)\right)\right) f(v, N_c, m_c) + \Delta, \tag{33}$$

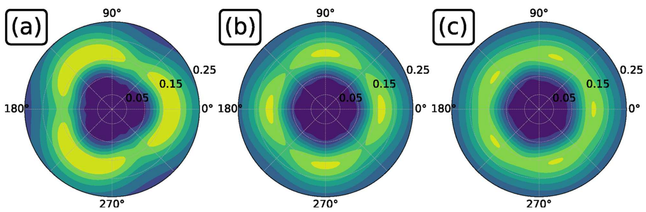

FIG. 3. Specific heat (shown in colors) in terms of the charge $2\pi q$ (polar coordinates) and temperature (radial coordinate) for $N_f = 3$(a), 4(b), and 5(c), assuming $N_c = 45$. This plot was derived from Eq. (33).

where $f(\nu, N_c, m_c) = \sum_{l=1}^{m_c-1} \cos(\frac{2\pi \nu l}{N_c})$ such that

$$f(\nu, N_c, m_c) = \begin{cases} m_c \text{ if } \nu = 0, \\ \dfrac{\sin\left(\frac{\pi \nu m_c}{N_c}\right)}{\sin\left(\frac{\pi \nu}{N_c}\right)} \cos\left(\dfrac{\pi \nu}{N_c}(m_c - 1)\right) \text{ otherwise.} \end{cases} \tag{34}$$

Equation (33) indicates a rotation symmetry related to the charge parameter in the modular directed network. These symmetries also reflect the behavior of the specific heat petal structure shown in Fig. 3.

2. Asymmetries in the specific heat petal structures

The results obtained in the previous section helps us to understand the relationship between the modular structures and the magnetic Laplacian spectrum, as well as the specific heat symmetry. However, these results assume that the inner structures $\mathbf{H_{in}}$ are undirected. The effect of directionality can be inferred by generating random directions inside the intrablocks, i.e., by imposing that $[\mathbf{W_{in}}]_{u,v}$ has probability $p_c < 100\%$ to take value 1. Adopting $p_c = 30\%$, we calculate the specific heat by using numeric diagonalization, yielding the structures in Fig. 4. We can observe that the obtained petals are not symmetric, unlike what had been observed for uniform connections.

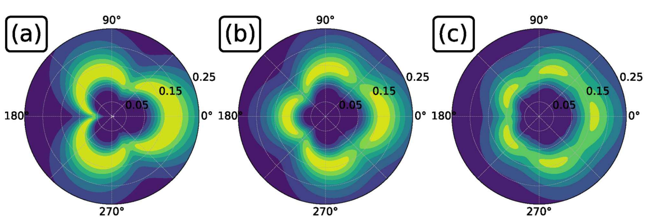

FIG. 4. Specific heat (colors) in terms of the charge $2\pi q$ (angle) and temperature (radius), for $N_f = 3$(a), 4(b), and 5(c), assuming $N_c = 45$. The networks were generated randomly, imposing the probability of having a directed edge as $p_c = 30\%$. Observe the obtained asymmetric petals contrasting with the results obtained previously for the uniform connections.

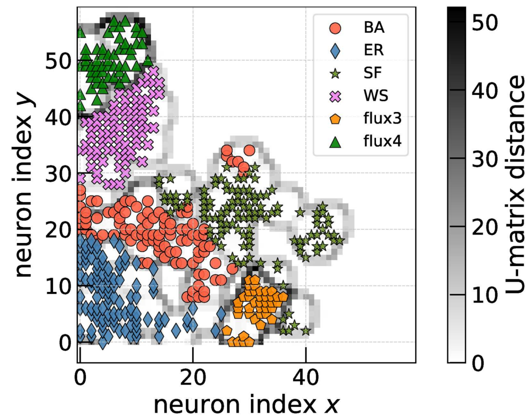

FIG. 5. SOM mapping of six types of complex networks represented by the specific heat approach. The *neuron index x* and *neuron index y* correspond to neurons in the SOM cortical space. The distances between neighboring neurons (U-matrix) are indicated in gray. A good separation between the types of networks can be observed.

B. Model characterization of directed graphs

In this section, we address the task of characterization of distinct networks models through the spectra of magnetic Laplacian. In particular, given a set of measurements obtained from a graph, can we infer which model created that graph? In this work, we opted to use the specific heat, c_λ, as a feature of measurement of graphs, in order to address the question above. As shown in Fig. 1, the c_λ measures yielded specific behavior for different models, therefore providing valuable information that can be used to identify and discriminate between different complex networks models.

In order to evaluate the efficiency of using c_λ as a fingerprint of a directed network, we built a dataset with 2000 network samples with types Erdős–Rényi (ER), Barabási (BA), the scale-free model of Bollobás *et al.*[30] (SF), Watts–Strogatz (WS), and Stochastic Block Model (SBM) with three and four blocks.

Then, self-organizing maps (SOMs), namely, a method for non-supervised clustering,[31] were trained with the obtained c_λ values and the obtained regions were subsequently labeled. This was done by feeding each training data into the SOM and choosing the neuron that exhibited highest activation. As indicated by the results shown in Fig. 5, networks belonging to the same class have been mapped into nearby neurons, defining respective clusters. So, the SOM was able, without previous knowledge to find the patterns of c_λ associated to the considered types of networks.

From what we have seen, we can conclude that the suggested magnetic Laplacian approach is able, at least for the considered cases, to properly characterize the model of given networks. For this reason, in a similar manner to that which has been applied in condensed matter physics, "SOM" proved to be a powerful technique for characterizing complex networks when we see these networks through the lens of statistical mechanics and magnetic Laplacians.

Given that many real-world networks contain a large number of nodes, a question arises regarding the feasibility using spectral quantities for their characterization. As described in Sec. II,

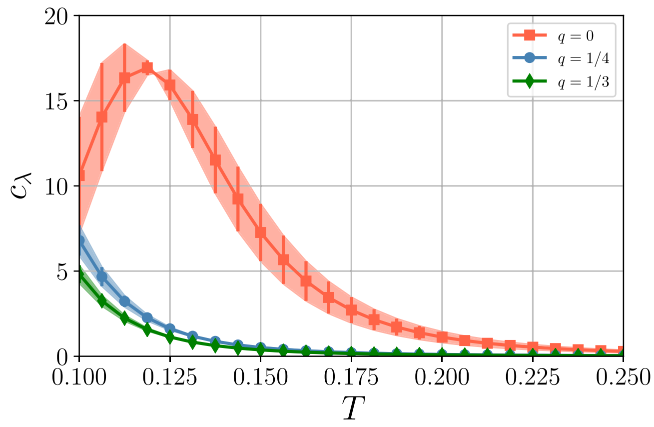

FIG. 6. Approximated specific heat for a network with $|V| = 3000$, $N_f = 3$, $p_c = 0.25$, and $p_d = 0.5$. In the application of the KPM method, the expansion was truncated at 40 first terms and the stochastic trace approximation used 25 random vectors. The error bars represent the deviation between the exact value (obtained numerically) and the approximated value calculated by the KPM method and using numerical integration.

thanks to the magnetic Laplacian formalism, KPM can be used as a means to estimate spectral density measurements. For instance, given a modular directed network, we obtained the exact and KPM-approximated values of the specific heat for different temperatures and charge values. The approximated specific heat is shown in Fig. 6. The error bars indicate a small dispersion, corroborating the potential of the KPM approach for studying the spectral properties of complex networks.

C. Directed network parameter inference

The results shown in Fig. 5 indicates that, given a network $\tilde{G}$, we can infer which model was responsible for generating it. In addition, to complete the task of characterizing a network, it is necessary to find the network which most closely resembles $\tilde{G}$ among several networks created with distinct parameters while fixing the model.

In this section, we explore the problem of inferring the parameters of models using the spectra of the magnetic Laplacian.

In order to argue that the Wasserstein metric can be used combined with the KPM approach as a means to estimate the network model parameters with sufficient precision, we study the problem of inferring the connecting probabilities $\tilde{p}$ of ER networks and the out-degree $\tilde{m}$ of BA networks, both with approximately 10^5 nodes.

In Fig. 7, the continuous vertical lines show the correct value of the parameter and the vertical dashed lines identify the position of the minimal of Eq.(15), which is the inferred value of the parameter. By using KPM with the 100 first terms of the Chebyshev polynomial and approximating the trace by using 20 random vectors, we observe the parameters can be inferred with good accuracy.

IV. CONCLUSIONS

Directed networks can be used to represent several real-world structures and problems. As a consequence, several approaches have been proposed, aimed at characterizing and comparing directed networks. Among these approaches, spectral methods present some particularly interesting properties, such as bearing a direct relationship with the structural and dynamical aspects of given networks. However, when applied to directed networks, the usual Laplacian operator yields complex eigenvalues, which are difficult to treat and interpret. Nevertheless, the hermiticity property of the magnetic Laplacian allows a set of real eigenvalues to be associated with a weighted directed network. We showed here that real eigenvalues and the associated charge parameter convey information about the network, more specifically regarding its mesoscale structures and the spectral and specific heat symmetry.

In order to extend the proposed methodology to larger networks containing hundreds of thousands of nodes, we showed that the KPM method can be combined with the magnetic Laplacian approach. This combination allowed us to estimate the spectral density of the magnetic operator with remarkable efficiency and accuracy. Given that we could estimate the spectral density of the magnetic Laplacian, we showed that the study of spectral geometry under the Wasserstein metric can be used as a tool to infer parameters of networks with low relative errors.

The reported contributions pave the way to a number of future developments and applications involving directed complex networks. For instance, these methods can be applied to study several other theoretical and real-world structures, including fake news dissemination, metabolic networks, neuronal systems, to name but a few possibilities. It would also be interesting to perform studies using random matrix theory in order to infer relationships between topology and spectra for more general complex networks. Since we deal only with spectral information, the results presented in this paper could also be immediately applied to multiplex networks.

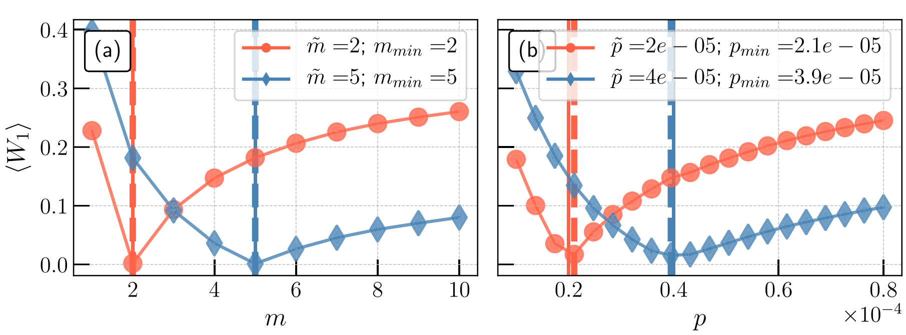

FIG. 7. The curves in (a) and (b) represent the mean of 1-Wasserstein distance equation (15), respectively, to BA and ER, in terms of the parameters adopted for network generation, considering $N_{exp} = 5$, $|V| = 10^5$, and $Q = \{0, 1/3\}$. For spectral estimation using the KPM, 100 terms of expansion and 20 random vectors were used.

ACKNOWLEDGMENTS

The authors thank Thomas Peron, Henrique F. de Arruda, Paulo E. P. Burke, and Filipi N. Silva for all suggestions and useful discussions. Bruno Messias F. de Resende acknowledges the CAPES for financial support. Luciano da F. Costa acknowledges the CNPq (Grant No. 307085/2018-0) and the NAP-PRP-USP for sponsorship. This work has been supported also by the FAPESP (Grant No. 15/22308-2). Research carried out using the computational resources of the Center for Mathematical Sciences Applied to Industry (CeMEAI) was funded by the FAPESP (Grant No. 2013/07375-0).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study. However, our implementation of the KPM method is available at github.com/stdogpkg/emate.

REFERENCES

[1] M. Kac, "Can one hear the shape of a drum?," Am. Math. Mon. **73**, 1–23 (1966).
[2] O. Giraud and K. Thas, "Hearing shapes of drums: Mathematical and physical aspects of isospectrality," Rev. Mod. Phys. **82**, 2213–2255 (2010).
[3] D. Aasen, T. Bhamre, and A. Kempf, "Shape from sound: Toward new tools for quantum gravity," Phys. Rev. Lett. **110**, 121301 (2013).
[4] L. Cosmo, M. Panine, A. Rampini, M. Ovsjanikov, M. M. Bronstein, and E. Rodolà, "Isospectralization, or how to hear shape, style, and correspondence," arXiv:1811.11465 [cs.CG] (2018).
[5] D. M. Cvetković, "Graphs and their spectra," in *Publikacije Elektrotehničkog fakulteta*, Serija Matematika i fizika (University of Belgrade, Serbia, 1971), Vol. 354/356, pp. 1–50.
[6] E. R. van Dam and W. H. Haemers, "Which graphs are determined by their spectrum?," Linear Algebra Appl. **373**, 241–272 (2003).
[7] C. Sarkar and S. Jalan, "Spectral properties of complex networks," Chaos **28**, 102101 (2018).
[8] J. Wang, R. C. Wilson, and E. R. Hancock, "Detecting Alzheimer's disease using directed graphs," in *Graph-Based Representations in Pattern Recognition* (Springer International Publishing, 2017), pp. 94–104.
[9] K. Anand and G. Bianconi, "Entropy measures for networks: Toward an information theory of complex topologies," Phys. Rev. E **80**, 045102 (2009).
[10] M. Dehmer and A. Mowshowitz, "A history of graph entropy measures," Inform. Sci. **181**, 57–78 (2011).
[11] C. Ye, C. H. Comin, T. K. D. Peron, F. N. Silva, F. A. Rodrigues, L. D. F. Costa, A. Torsello, and E. R. Hancock, "Thermodynamic characterization of networks using graph polynomials," Phys. Rev. E **92**, 032810 (2015).
[12] M. De Domenico and J. Biamonte, "Spectral entropies as information-theoretic tools for complex network comparison," Phys. Rev. X **6**, 041062 (2016).
[13] C. Nicolini, V. Vlasov, and A. Bifone, "Thermodynamics of network model fitting with spectral entropies," Phys. Rev. E **98**, 022322 (2018).
[14] J. D. Hart, J. P. Pade, T. Pereira, T. E. Murphy, and R. Roy, "Adding connections can hinder network synchronization of time-delayed oscillators," Phys. Rev. E **92**, 022804 (2015).
[15] G. Berkolaiko, "Nodal count of graph eigenfunctions via magnetic perturbation," Anal. PDE **6**, 1213–1233 (2013).
[16] M. Fanuel, C. M. Alaíz, and J. A. K. Suykens, "Magnetic eigenmaps for community detection in directed networks," Phys. Rev. E **95**, 022302 (2017).
[17] S. Furutani, T. Shibahara, M. Akiyama, K. Hato, and M. Aida, "Graph signal processing for directed graphs based on the Hermitian Laplacian," in *Machine Learning and Knowledge Discovery in Databases*, edited by U. Brefeld, E. Fromont, A. Hotho, A. Knobbe, M. Maathuis, and C. Robardet (Springer International Publishing, Cham, 2020), pp. 447–463.
[18] L. Kantorovitch, "On the translocation of masses," C. R. Acad. Sci. URSS **37**, 199–201 (1942).
[19] V. I. Bogachev and A. V. Kolesnikov, "The Monge-Kantorovich problem: Achievements, connections, and perspectives," Russian Math. Surveys **67**, 785–890 (2012).
[20] G. Peyré and M. Cuturi, "Computational optimal transport," Found. Trends Mach. Learn. **11**, 355–206 (2019).
[21] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, "Directed scale-free graphs," in *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms* (Society for Industrial and Applied Mathematics, 2003), pp. 132–139.
[22] Y. C. de Verdière, "Magnetic interpretation of the nodal defect on graphs," Anal. PDE **6**, 1235–1242 (2013).
[23] E. H. Lieb and M. Loss, "Fluxes, Laplacians, and Kasteleyn's theorem," in *Statistical Mechanics* (Springer, Berlin, 1993), pp. 457–483.
[24] K. Blum, *Density Matrix Theory and Applications* (Springer Science & Business Media, 2012).
[25] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, "The kernel polynomial method," Rev. Mod. Phys. **78**, 275–306 (2006).
[26] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd rev. ed. (Dover Publications, 2001).
[27] M. Hutchinson, "A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines," Commun. Stat. Simul. Comput. **19**, 433–450 (1990).
[28] D. Jackson, "On approximation by trigonometric sums and polynomials," Trans. Am. Math. Soc. **13**, 491–491 (1912).
[29] R. M. Gray, "Toeplitz and circulant matrices: A review," Found. Trends Commun. Inform. Theory **2**, 155–239 (2005).
[30] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, "Directed scale-free graphs," in *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA'03* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2003), pp. 132–139.
[31] A. A. Shirinyan, V. K. Kozin, J. Hellsvik, M. Pereiro, O. Eriksson, and D. Yudin, "Self-organizing maps as a method for detecting phase transitions and phase identification," Phys. Rev. B **99**, 041108 (2019).