Spectral Dimension Reduction of Complex Dynamical Networks

Edward Laurence,[1,2] Nicolas Doyon,[2,3,4] Louis J. Dubé,[1,2] and Patrick Desrosiers[1,2,4]

[1]*Département de physique, de génie physique, et d'optique, Université Laval, Québec G1V 0A6, Canada*
[2]*Centre interdisciplinaire en modélisation mathématique de l'Université Laval, Québec G1V 0A6, Canada*
[3]*Département de mathématiques et de statistique, Universite Laval, Québec G1V 0A6, Canada*
[4]*Centre de recherche CERVO, Québec G1J 2G3, Canada*

(Received 21 September 2018; revised manuscript received 16 January 2019; published 4 March 2019)

Dynamical networks are powerful tools for modeling a broad range of complex systems, including financial markets, brains, and ecosystems. They encode how the basic elements (nodes) of these systems interact altogether (via links) and evolve (nodes' dynamics). Despite substantial progress, little is known about why some subtle changes in the network structure, at the so-called critical points, can provoke drastic shifts in its dynamics. We tackle this challenging problem by introducing a method that reduces any network to a simplified low-dimensional version. It can then be used to describe the collective dynamics of the original system. This dimension reduction method relies on spectral graph theory and, more specifically, on the dominant eigenvalues and eigenvectors of the network adjacency matrix. Contrary to previous approaches, our method is able to predict the multiple activation of modular networks as well as the critical points of random networks with arbitrary degree distributions. Our results are of both fundamental and practical interest, as they offer a novel framework to relate the structure of networks to their dynamics and to study the resilience of complex systems.

DOI: 10.1103/PhysRevX.9.011042

Subject Areas: Complex Systems, Nonlinear Dynamics, Statistical Physics

I. INTRODUCTION

Critical breakdowns generally arise unexpectedly in complex dynamical systems [1]. Noteworthy examples are financial crises [2,3], epileptic seizures [4], and species extinctions [5]. These breakdowns are typically identified by using global-scale indicators that collapse at the critical point, such as stock market indices, neural synchronization, and species biomass. While much effort has been devoted to forecast breakdowns [6], no simple and universal method has yet been found. This is mostly due to the inherent complexity of the problem: Real systems are composed of multiple units that participate in the global state in highly complicated patterns of interactions.

Network science addresses this problem and offers a unifying framework where a complex system with N fundamental units is described as a network of N components (nodes). The state of each unit is encoded into an activity variable, and the evolution of the states in the whole system is governed by N coupled dynamical equations that depend on both the activity variables and a set of weighted interactions (links). Therefore, the dynamical properties of the system strongly depend on the underlying network structure. Although recent developments have clarified how small targeted perturbations in the network structure can provoke drastic changes in the structure itself [7,8], much less is known about the dynamical effects of these perturbations. One promising approach is to use dimension reduction to transform the original N-dimensional representation into a simplified version with $n \ll N$ effective dimensions.

Recently, Gao *et al.* presented a dimension reduction formalism that collapses any N-dimensional network into a one-dimensional effective version and used it to predict the global activity of the original network [9]. The authors proposed measuring the global activity as the degree-weighted average activity in which the nodes with high degree, i.e., high number of links, contribute more to the average than those with low degree. The rationale behind this choice is that the highly connected nodes have a higher impact on the dynamics. Moreover, they have shown that the degree-weighted connectivity is sufficient to explain the global level of activity. Their formalism can be applied to a wide variety of complex systems, thus suggesting that the degree-weighted averages are in fact *universal* predictors. For instance, they accurately predict the minimum level of interaction between species to prevent biomass extinction,

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.

the so-called critical point, of real ecosystems. Yet, this spectacular outcome is not totally satisfactory since no fundamental reason is provided that would explain why degree is the key property to any network structure, particularly those with degree correlations.

In an attempt to determine the critical points of 59 bipartite mutualistic ecosystems [10], Jiang *et al.* proposed a two-dimensional reduction that divides each original ecosystem into two populations for which they obtain the average interaction strength. From numerical explorations, they conclude that the degree may not always be the key predictive property of a network. Their results also suggest that two-dimensional reductions can lead to better predictions than the one-dimensional formalism of Gao *et al.*

It remains unclear whether the dimension reduction procedure of Gao *et al.* can lead to accurate predictions for arbitrary network structures and why some dynamical networks should require two-dimensional reductions. Strong theoretical foundations are clearly lacking to answer those questions. The goal of this paper is to address these issues. We rely on a simple and strong hypothesis: One can predict the evolution of a small number of variables describing the global activities of a network. These variables are constructed as *a priori* unknown weighted averages of the individual node activities. By enforcing this hypothesis, we provide theoretical justifications for the required number of effective dimensions and quantify the contribution of each component to the universal global activities. Beyond the mere improvement in precision over existing approaches, our method allows the detection of dynamical breakdowns that would be missed altogether with previous reductions.

This study provides a reliable tool for researchers who want to study critical breakdowns of complex systems. Using our approach, once a system is framed as a network, one can first determine the number of effective dimensions required to adequately reduce the system at hand and then find the variables of interest that describe the global state. Moreover, our method also identifies the units that, if perturbed, can induce large reactions in the system. Our findings thus lead to a deeper understanding of how critical breakdowns occur and how to prevent them.

The paper is structured as follows. We first present the general framework for complex dynamical networks (Sec. II). We then describe a general method to obtain a one-dimensional reduction for mutualistic networks (Sec. III). We show that the reduction scheme of Gao *et al.* emerges as an approximation of our general approach when specifically considering random networks. In Sec. IV, we develop the cycle reduction, a multidimensional approach that is useful to reduce heterogeneous and bipartite networks. Next, in Sec. V, we complete the method by including subdominant contributions of the structure. We finally assess the goodness of these reductions, as a function of the structure, and the nature of the dynamics (Sec. VI).

II. MODEL DEFINITION

The diverse nature of complex systems requires that we establish a common ground. In Sec. II A, we regroup dynamical complex networks under a general model that encodes the structure and the dynamics. Then, in Sec. II B, we provide examples of contrasting models of dynamics satisfying the formalism used afterward in the paper to illustrate the dimension reduction methods.

A. General formalism

We consider a complex network of N units, called nodes, for which the interactions are encoded in the weighted and directed adjacency matrix $\boldsymbol{W}$. The element $w_{ij} \in \mathbb{R}$ of $\boldsymbol{W}$ is interpreted as the strength of the directed interaction from node j to node i.

Each node has an activity $x_i \in \mathbb{R}$ whose evolution is governed by the general equation

$$\dot{x}_i = F(x_i) + \sum_{j=1}^{N} w_{ij} G(x_i, x_j), \tag{1}$$

where $F(x_i)$ and $G(x_i, x_j)$ are real-valued functions. For technical reasons that will become clear in the next section, both $F(x_i)$ and $G(x_i, x_j)$ are required to have continuous derivatives of second order. The product $w_{ij}G(x_i, x_j)$ specifies the type of interactions.

If $w_{ij}\partial G(x_i, x_j)/\partial x_j \leq 0$, the interaction is competitive and the increase of activity of node j tends to decrease the activity of node i. If $w_{ij}\partial G(x_i, x_j)/\partial x_j \geq 0$, the interaction is mutualistic, and therefore node j activity benefits node i. For mixed dynamics of mutualistic and competitive interactions, it is common to fix $\partial G(x_i, x_j)/\partial x_j \geq 0$ and use negative weights $w_{ij} < 0$ for competitive interactions. Unless specified, we will only consider mutualistic dynamics with $w_{ij}\partial G(x_i, x_j)/\partial x_j \geq 0$. Furthermore, we concentrate our studies on the possible dynamical models (Sec. II B) in parameter ranges for which $\partial G(x_i, x_j)/\partial x_j \geq 0$ such that w_{ij} are all non-negative as well.

To describe the evolution of the whole system at the macroscopic and the mesoscopic scales, it is convenient to focus on observables. We define an observable as a *smooth* function mapping the activities $x_1, \ldots, x_N$ to a real number. Among all observables, the linear observables, functions of the form $L(\boldsymbol{x}) = \sum_i a_i x_i \in \mathbb{R}$, are of particular interest for different reasons.

The first reason to use a linear observable is the numerical evidence given by Ref. [9] that suggests that a linear combination of the activity could be a good indicator of the global state of the network. We have followed this direction, which has led us to a more general linear dimension reduction formalism. The second reason is that a linear combination is far more intuitive than a general nonlinear observable of the form $\Omega(\boldsymbol{x}) = \sum_{n,j} a_{nj} x_j^n$. It is

not excluded that such nonlinear observables could provide adequate dimension reductions, but their interpretability would most surely be limited. Finally, there is a more formal and practical reason to select linear observables. The set of all observables, which are smooth functions from $\mathbb{R}^N$ to $\mathbb{R}$, forms a real vector space $\mathcal{V}$ of infinite dimension. Dimension reduction can thus be seen as a search for $n < N$ observables, among infinitely many possible choices, whose evolution can be well approximated by n coupled differential equations. At first sight, there is no obvious way to determine the best observables to choose. However, $\mathcal{V}$ contains a subspace $\mathcal{L}$ formed by all linear observables. The dimension of $\mathcal{L}$ is precisely N, and one can show that any basis of $\mathcal{L}$ has a dynamics that is exactly described by N differential equations very similar to the original ones. This simplifies the search immensely: Rather than looking in the whole space $\mathcal{V}$, the spectral properties of the adjacency matrix to find n elements of interest in $\mathcal{L}$ offer themselves as a natural choice.

B. Examples of possible dynamics

A number of dynamical systems satisfy the form of Eq. (1). For instance, in computational neuroscience, the Cowan-Wilson model [11] describes the firing-rate activity of a population of neurons as

$$\dot{x}_i = -x_i + \sum_{j=1}^{N} w_{ij} \frac{1}{1 + \exp[-\tau(x_j - \mu)]}, \tag{2a}$$

where τ and μ are parameters controlling the steepness of the activation function and the firing-rate threshold, respectively [12].

In biology, the generalized Lotka-Volterra dynamics describes the evolution of the population of species in an ecosystem as

$$\dot{x}_i = \omega x_i + x_i \sum_{j=1}^{N} w_{ij} x_j, \tag{2b}$$

where ω is the intrinsic growth rate [13], and x_i is the population of individuals of species i. To prevent unbounded growth and account for species migration and the Allee effect, a more complex model of ecological networks has been proposed [9,14]:

$$\dot{x}_i = B_i + x_i\left(1 - \frac{x_i}{K_i}\right)\left(\frac{x_i}{C_i} - 1\right) + \sum_{j=1}^{N} w_{ij} \frac{x_i x_j}{D_i + E_i x_i + H_i x_j}, \tag{2c}$$

where all parameters are real numbers, B_i accounts for the migration rate, $K_i > 0$ for the ecosystem capacity, and $C_i > 0$ for the minimum abundance for species growth. The parameters D_i, E_i, H_i control the strengths of the interactions between the species.

The Michaelis-Menten equation is yet another example [15]. It applies to the gene regulatory networks and governs the concentration of substrates as

$$\dot{x}_i = -c x_i^a + \sum_{j=1}^{N} w_{ij} \frac{x_j^b}{x_j^b + 1}, \tag{2d}$$

where a, b, c are positive parameters.

In social networks, the spreading of a virus or rumors can be described using the susceptible-infected-susceptible model (SIS) [16]. In this context, the activity $x_i \in [0, 1]$ is interpreted as the probability of being infected, and it evolves according to

$$\dot{x}_i = -x_i + \gamma(1 - x_i) \sum_{j=1}^{N} w_{ij} x_j, \tag{2e}$$

with $\gamma \geq 0$ as the normalized infection rate.

III. ONE-DIMENSIONAL REDUCTION

The systems described by Eq. (2) are N dimensional, with their dynamics governed by N coupled differential equations. As the number of nodes N grows, the computational cost of solving N coupled equations increases, which raises a number of issues [17,18]. Moreover, the state of the original system given by the N-dimensional vector $\boldsymbol{x}$ becomes less intelligible and less insightful, and does not provide much into the general properties of the solutions.

Hence, we must rely on measures, or observables, to reduce N-dimensional systems to more practical and accessible objects. For instance, the unweighted average activity could be a measure on how dissimilar the system state is compared to a specifically chosen state. We also want to make predictions on those measures to anticipate dynamical breakdowns and locate the global state of the system on a standardized bifurcation diagram. However, when solely based on the unweighted average activity, the predictions are often nonrepresentative of the original system [10].

Alternatively, a *weighted* activity seems more reliable as we inject additional information on the importance of nodes, and it has already been proven to be a promising avenue of breakdown predictions [9,10]. In the next subsections, we introduce a general procedure to select a weighted activity and to predict its evolution.

A. Derivation of the reduction formalism

Let us consider a real linear observable R of the activity:

$$R = \sum_{i=1}^{N} a_i x_i = \boldsymbol{a}^T \boldsymbol{x}, \tag{3}$$

where $a_i \in \mathbb{R}$, the ith component of the column vector $\boldsymbol{a}$, is a normalized weight so that

$$\sum_{i=1}^{N} a_i = \mathbf{1}^T \boldsymbol{a} = 1. \tag{4}$$

In general, some components a_i can be positive or negative, and R represents a weighted activity. In many instances, however, a_i will be non-negative; i.e., the normalized vector $\boldsymbol{a}$ will be a probability vector, and our observable R could then justifiably be called a weighted *average* activity, where a_i is the relative contribution, or centrality, of node i to the observable.

The linear observable R is a function that takes the instantaneous activity of each node and returns a real number that describes the global state of the network. For instance, for $a_i = 1/N$, R describes the unweighted average activity. Although the average activity is attractive because of its simplicity, it may not be easy to predict its value using only the structure of the network and the nature of the dynamics. Thus, we hypothesize that $\boldsymbol{a}$ should be specific to the structure.

Let us explain how the weight vector $\boldsymbol{a}$ is constrained by the adjacency matrix $\boldsymbol{W}$. By taking the time derivative of Eq. (3) and using Eq. (1) (refer to Appendix A for complete derivation), we obtain that the dynamics of R—truncated up to second-order terms $O[(x_k - R)^2]$—is given by the one-dimensional equation

$$\dot{R} \approx F(R) + \alpha G(\beta R, R), \tag{5}$$

where β is a structural parameter given by

$$\beta = \frac{1}{\alpha} \frac{\boldsymbol{a}^T \boldsymbol{K} \boldsymbol{a}}{\boldsymbol{a}^T \boldsymbol{a}} \tag{6}$$

and $\boldsymbol{K}$ is an $N \times N$ diagonal matrix of diagonal elements $K_{ii} = k_i^{\text{in}} = \sum_{j=1}^{N} w_{ij}$, the in-degree of node i. The parameter α can be measured directly on the network as the weighted in-degree,

$$\alpha = \sum_{i=1}^{N} a_i k_i^{\text{in}} = \boldsymbol{a}^T \boldsymbol{k}^{\text{in}}. \tag{7}$$

Interestingly, we show, in Appendix A, that the closed form of Eq. (5) is satisfied only if $\boldsymbol{a}$ is a normalized eigenvector of the transposed adjacency matrix $\boldsymbol{W}^T$ with eigenvalue α,

$$\boldsymbol{W}^T \boldsymbol{a} = \alpha \boldsymbol{a}. \tag{8}$$

We have now obtained a single equation [Eq. (5)] that governs the evolution of the weighted activity of a complex network and constrained the weight vector $\boldsymbol{a}$ to be adapted to the structure under study.

Clearly, Eq. (5) shares similarities with Eq. (1). We can interpret the former as a reduced system of one dynamical node that interacts with itself. The nature of its dynamics is identical to the one from the original system, i.e., specified by $F(x_i)$ and $G(x_i, x_j)$, and the coupling is parametrized by α, β. The activity R of the single node describes the weighted activity of the original network and is obtained by solving Eq. (5), whose solutions are solely controlled by the nature of the dynamics and the structural parameters α, β.

B. Choice of a universal weight vector

We have seen that the weight vector $\boldsymbol{a}$ must be a normalized eigenvector of $\boldsymbol{W}^T$ so that the observable R satisfies Eq. (5). In principle, any eigenvectors of $\boldsymbol{W}^T$, except those that satisfy $\mathbf{1}^T \boldsymbol{a} = 0$, could be used for the dynamical reduction. However, the larger the modulus of α is, the stronger is the influence of the structure on the weighted activity. An eigenvector whose eigenvalue has a low modulus leads to a linear observable R that does not properly take into account the network structure, which in turn leads to correction terms $O[(x_k - R)^2]$ greater than those produced by eigenvectors with a higher modulus. The choice of $\boldsymbol{a}$ as the eigenvector with the largest eigenvalue modulus seems to impose itself: $\boldsymbol{a}$ is the dominant eigenvector.

For an arbitrary weighted adjacency matrix, the dominant eigenvalue and the components of the dominant eigenvector can be complex. In this case, the observable R, as well as the structural parameters α and β, is complex. The one-dimensional dynamical system of Eq. (5) becomes complex too and can be interpreted as a two-dimensional real dynamical system.

However, there is a large class of networks for which the dominant eigenvalue and the components of the dominant eigenvector are all real. For instance, strongly connected (in practice, sufficiently connected) undirected and directed networks with non-negative weights w_{ij} fall into this class. In fact, the Perron-Frobenius theorem guarantees that if the network is strongly connected, i.e., a path exists between each pair of nodes, and all edge weights satisfy $w_{ij} \geq 0$, then the dominant eigenvalue λ_D of $\boldsymbol{W}^T$ is non-negative $\lambda_D \geq 0$, and the dominant eigenvector is elementwise positive [19]. Moreover, in practice, the dominant eigenvector can be efficiently computed using the power method.

The procedure to apply this one-dimensional dimension reduction is straightforward. First, we compute the dominant eigenvalue α and the corresponding eigenvector $\boldsymbol{v}_D$ of $\boldsymbol{W}^T$. Second, we define the normalized eigenvector $\boldsymbol{a} = \boldsymbol{v}_D / (\mathbf{1}^T \boldsymbol{v}_D)$ and obtain β according to Eq. (6). In most cases, we want to determine the weighted activity at equilibrium R^*, determined by solving

$$0 = F(R^*) + \alpha G(\beta R^*, R^*). \tag{9}$$

This is a universal equation in the sense that α and β are independent of the dynamics, controlled by $F(x_i)$ and $G(x_i, x_j)$; α and β only depend upon the network structure, encoded in $\boldsymbol{W}$. The one-dimensional reduction process has been applied to different dynamics for small random uncorrelated networks and has led to surprisingly accurate predictions (Fig. 1). For larger networks, we expect the formalism to maintain a similar level of accuracy. However, since the number of nodes does not intervene explicitly in the formalism, we are generally unable to analytically describe how the quality of the reduction varies with the network size. We must rely on a numerical investigation.

Our numerical experiments indicate that the network size *by itself* has no significant impact on the quality of the reduction. Rather, the accuracy strongly depends on the network structure and, in particular, on the degree variance. In a nutshell, our findings can be summarized as follows:

(a) For a given N, the larger the average number of edges per node $\langle \boldsymbol{k}_B^{\text{in}} \rangle$ [21], the better the dimensional reduction will be.

(b) For a fixed value of $\langle \boldsymbol{k}_B^{\text{in}} \rangle$, the residual error of the reduction tends rapidly to a finite limit as N is increased.

(c) For large enough $\langle \boldsymbol{k}_B^{\text{in}} \rangle$, the reduction error is small and insensitive to the precise value of N [corollary of (a) and (b)].

In Fig. 2, we examine these conclusions by comparing different network ensembles. Although displayed for undirected Erdős-Rényi (ER) networks, $G(N, p)$, the results are representative of a larger set of calculations, and they nicely synthesize our general conclusions on this issue. In Fig. 2(a), dynamics on networks of different $\langle \boldsymbol{k}_B^{\text{in}} \rangle = p(N-1)$ but equal number of nodes $N = 200$ is reproduced differently by the one-dimensional reduction. Denser networks (large p, large $\langle \boldsymbol{k}_B^{\text{in}} \rangle$) are better represented by the reduction than sparser (small p, small $\langle \boldsymbol{k}_B^{\text{in}} \rangle$) networks. In Fig. 2(b), we compare networks of different sizes but equal $\langle \boldsymbol{k}_B^{\text{in}} \rangle$. For a fixed $\langle \boldsymbol{k}_B^{\text{in}} \rangle$ (here $\langle \boldsymbol{k}_B^{\text{in}} \rangle = 10$), the connection probability p is adjusted to compensate for the growing number of nodes. As the number of nodes is increased, the residual error on the

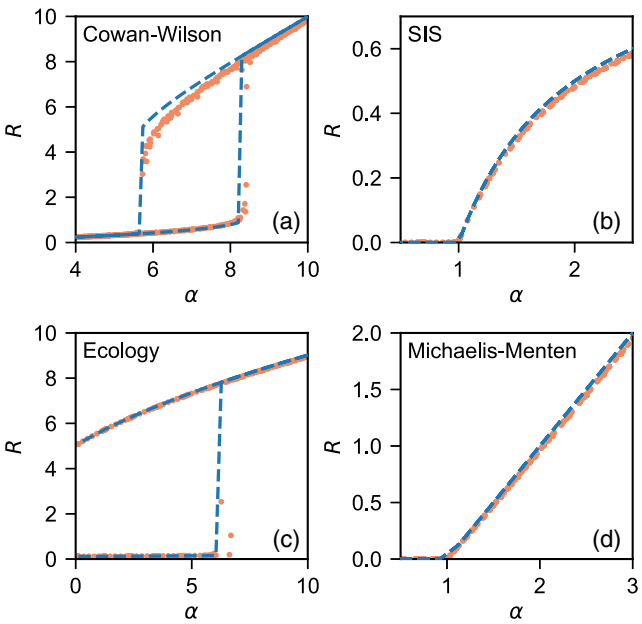

FIG. 1. Observable $R^* = \boldsymbol{a}^T \boldsymbol{x}^*$ at equilibrium as a function of the dominant eigenvalue α of $\boldsymbol{W}^T$, for different dynamics on Erdős-Rényi networks of $N = 100$ nodes and connection probability $p = 0.1$. (a) Cowan-Wilson dynamics [Eq. (2a)] with $\tau = 1$, $\mu = 3$, (b) SIS dynamics [Eq. (2e)] with $\gamma = 1$, (c) mutualistic ecological dynamics [Eq. (2c)] with $B_i = 0.1$, $C_i = 1$, $K_i = 5$, $D_i = 6$, $E_i = 0.9$, $H_i = 0.1$ [20], and (d) Michaelis-Menten dynamics [Eq. (2d)] with $a = 1$, $b = 1$, $c = 1$. Dashed lines are theoretical predictions obtained from Eq. (9), while dots are equilibrium states resulting from the evolution of the whole N-dimensional system. For each dynamics and network ensemble, 100 networks are generated. For each network, we scale the edge weights by a constant random factor as $w_{ij} \mapsto c w_{ij}$, so the dominant eigenvalue of the adjacency matrix is located in the region of interest. Then, the dynamics is integrated to equilibrium, and an orange dot is placed at the corresponding point (α, R^*). Next, the network is perturbed by removing an edge, the dynamics is brought back to equilibrium, and a new dot is placed at (α', R'^*). The perturbation step is repeated 50 times for each network.

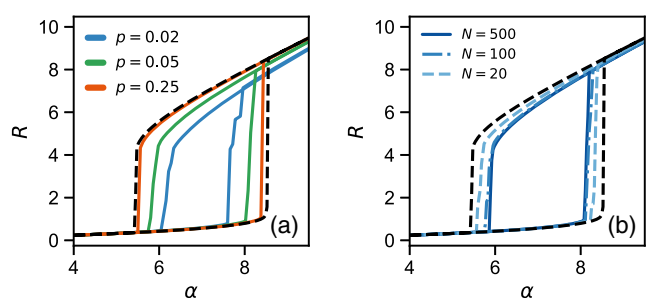

FIG. 2. (a) Comparison of the observable $R^* = \boldsymbol{a}^T \boldsymbol{x}^*$ at equilibrium as a function of the dominant eigenvalue α of $\boldsymbol{W}^T$ for different connection probabilities p of undirected Erdős-Rényi networks $G(N, p)$ and the Cowan-Wilson dynamics [Eq. (2a)] with $\tau = 1, \mu = 3$. For each value of p, 10 networks of 200 nodes are generated, and their dynamics are solved for $\alpha \in [4, 10]$. Lines are computed as a binned average over networks sharing the same probability of connection. (b) Comparison of Erdős-Rényi networks $G(N, p)$ of different sizes $N \in \{20, 100, 500\}$ but equal average numbers of edges per node $\langle \boldsymbol{k}_B^{\text{in}} \rangle = 10$. For each N, we adjust the connection probabilities $p \in \{0.50, 0.10, 0.02\}$ to match the expected value of $\langle \boldsymbol{k}_B^{\text{in}} \rangle$ and solve the dynamics on 10 network realizations. Lines are computed as a binned average over same-size networks. Dashed lines are theoretical predictions obtained from Eq. (9). To obtain a specific $\alpha = \boldsymbol{a}^T \boldsymbol{k}^{\text{in}}$, we multiply each edge weight by a constant scaling factor $w_{ij} \mapsto c w_{ij}$ so that the dominant eigenvalue $\alpha \mapsto c\alpha$ falls in the range [4,10] using the dominant eigenvector $\boldsymbol{a}$ as a weight vector. We then solve the dynamics at equilibrium and measure the observable $R^* = \boldsymbol{a}^T \boldsymbol{x}^*$.

prediction rapidly tends to a finite limit, and no further deterioration of the quality of the dimensional reduction is observed. In other words, the goodness of the reduction is practically invariant of the network size, if large enough, and is mostly governed by the average number of edges per node.

These observations extend to other types of networks and confirm that the quality of the reduction is more affected by the network connectivity than by the network size. A dynamical explanation goes as follows. In sparse networks (e.g., $N = 200$, $p = 0.02$), one must use large scaling factors $c > 1$ ($\boldsymbol{W} \mapsto c\boldsymbol{W}$, which implies $\boldsymbol{W}\mathbf{1} = \boldsymbol{k}^{\text{in}} \mapsto c\boldsymbol{k}^{\text{in}}$, $\alpha = \boldsymbol{a}^T\boldsymbol{k}^{\text{in}} \mapsto c\alpha$) to reach the desired range of the dominant eigenvalue $\alpha \in [4, 10]$ (see caption of Fig. 2; the scaling does not alter the topology, only the strength of what is injected in the dynamical equations through the parameter α). This accentuates the inequalities between the degrees of the nodes, $\text{Var}(\boldsymbol{k}^{\text{in}}) \mapsto c^2\text{Var}(\boldsymbol{k}^{\text{in}})$, and eventually splits the populations into active and inactive nodes. Hence, the reduction is unable to describe the two populations with a single variable. In the opposite extreme of dense networks (e.g., $N = 200$, $p = 0.25$), the edge weights must be scaled down, $c < 1$, to reach the same dominant eigenvalue. Therefore, the nodes follow a more global activation scheme that can be described with high accuracy by the one-dimensional reduction. This explanation matches the observed finite limit for the quality of the reduction [Fig. 2(b)]. For a fixed value of $\langle \boldsymbol{k}_B^{\text{in}} \rangle = p(N-1)$, when N grows larger, p tends to zero so that the variance $\text{Var}(\boldsymbol{k}_B^{\text{in}}) \approx (N-1)p(1-p)$ tends to a limit $\text{Var}(\boldsymbol{k}_B^{\text{in}}) \to \langle \boldsymbol{k}_B^{\text{in}} \rangle$ and no further deterioration of the quality is observed.

The quantitative impact of the degree variance is best explained by a simple example. Consider an ER network ($N = 200$, $p = 0.1$) with $\langle \boldsymbol{k}_B^{\text{in}} \rangle \approx 20$ and $c = 1$ so that the variance of the edge weights is $\text{Var}(\boldsymbol{k}^{\text{in}}) \approx 18$ and the dominant eigenvalue is $\alpha \approx 20$. For a denser network ($N = 200$, $p = 0.5$) with $\langle \boldsymbol{k}_B^{\text{in}} \rangle \approx 100$, one must use a small scaling factor, $\boldsymbol{W} \mapsto \boldsymbol{W}/5$, to have the same dominant eigenvalue $\alpha \approx 20$, and the scaled variance of the edge weights is now $\text{Var}(\boldsymbol{k}^{\text{in}}) \approx 2$. Therefore, the nodes in the denser networks are more uniformly activated than in the sparser networks and are better amenable to the one-dimensional reduction.

C. Choice of an approximate weight vector

Recently, Gao *et al.* [9] introduced a different one-dimensional reduction for dynamics of the form of Eq. (1). In this section, we show how their reduction is a special case of our one-dimensional reduction when applied to uncorrelated random networks.

Uncorrelated random networks are a family of networks for which the degree distribution can be arbitrary but the probability of the connection between two nodes is independent of the presence or absence of any other edge [22]. We generate our random networks using the configuration model [23]. We first sample the in and out expected node degrees, $\boldsymbol{\kappa}^{\text{in}}$ and $\boldsymbol{\kappa}^{\text{out}}$, from an arbitrary degree distribution. Then, we connect node j to node i with probability

$$P_{ij} = \frac{\kappa_i^{\text{in}}\kappa_j^{\text{out}}}{m}, \tag{10}$$

where $m = \sum_i \kappa_i^{\text{out}}$ is the expected total number of edges. If the resulting network is strongly connected, the Perron-Frobenius theorem guarantees that the dominant eigenvector $\boldsymbol{v}_D$ of $\boldsymbol{W}^T$ will have only non-negative elements. We may then use this dominant eigenvector to construct the observable

$$R = \boldsymbol{a}^T\boldsymbol{x} = \frac{\boldsymbol{v}_D^T\boldsymbol{x}}{\mathbf{1}^T\boldsymbol{v}_D}. \tag{11}$$

For networks that satisfy Eq. (10), spectral graph theory [24] informs us that the elements of the dominant eigenvector $\boldsymbol{v}_D$ of $\boldsymbol{W}^T$ (i.e., the weights of the reduced system) can be approximated by the vector of out-degrees $\boldsymbol{k}^{\text{out}}$ as

$$a_i = [\boldsymbol{v}_D]_i \approx \frac{k_i^{\text{out}}}{\sum_{j=1}^N k_j^{\text{out}}} \tag{12}$$

if the rather mild condition

$$\frac{\langle (k^{\text{out}})^2 \rangle}{\langle k^{\text{out}} \rangle} > \sqrt{\max[k_i^{\text{out}}]}\ln(N) \tag{13}$$

is satisfied. It then results from Eq. (7) that α measures the average neighbor in-degree, that is,

$$\alpha \approx \frac{\sum_{i=1}^N k_i^{\text{out}} k_i^{\text{in}}}{\sum_{i=1}^N k_i^{\text{out}}}, \tag{14a}$$

and Eq. (6) reduces to $\beta = 1$. Therefore, R is simply the average neighbor activity:

$$R \approx \frac{\sum_{i=1}^N k_i^{\text{out}} x_i}{\sum_{i=1}^N k_i^{\text{out}}}. \tag{14b}$$

It turns out that this special case is exactly the formalism proposed by Gao *et al.* [9], with $R = x_{\text{eff}}$ and $\alpha = \beta_{\text{eff}}$ in their notation. It also means that the formalism of Gao *et al.* is mostly appropriate for random networks [Eq. (10)] respecting Eq. (13). Moreover, a recent work [25] has introduced a corrected eigenvalue approximation for random networks with power-law degree distribution $p(k) \sim k^{-\gamma}$ with $\gamma > 5/2$. This may further limit the accuracy of the approach of Gao *et al.* with respect to our one-dimensional reduction scheme.

IV. MULTIDIMENSIONAL REDUCTION: DOMINANT EIGENVECTORS

In the one-dimensional reduction, we suppose that the dynamical global state of a network is dominated by the information contained in a single dominant eigenvalue and corresponding eigenvector. Therefore, we also presuppose that other eigenvalues can be safely neglected and do not provide relevant information about the dynamics on the network. But, if the network admits many eigenvalues of similarly large modulus, it is plausible to expect that all of these eigenvalues are important as well and should be included in an n-dimensional reduction. In this section, we address this problem by introducing an n-dimensional approach to predict the evolution of n coupled observables. We argue that if the spectrum $\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ of $\boldsymbol{W}^T$ satisfies

$$|\lambda_1| \approx |\lambda_2| \approx \ldots \approx |\lambda_n| \gg |\lambda_{n+1}| \geq \ldots \geq |\lambda_N|,$$

then n observables should be considered, leading to an n-dimensional reduced dynamical system.

A. Cycle reduction

Let us consider n observables R_j, $1 \leq j \leq n$, each being a different linear combination of the activity

$$R_j = \sum_{i=1}^{N} [\boldsymbol{a}_j]_i x_i, \tag{15}$$

where $\boldsymbol{a}_j$ is a real-value weight vector associated with the observable R_j and normalized $\sum_i [\boldsymbol{a}_j]_i = 1$. As in the one-dimensional reduction, $\boldsymbol{a}_j$ are yet undetermined. While there are several choices for $\boldsymbol{a}_j$ that are *a priori* plausible, we discuss the *cycle* reduction, which is natural when several dominant eigenvalues are approximately of equal modulus.

Using a similar approach as the one-dimensional reduction (see Appendix B), one finds that the evolution of the observables is given by

$$\dot{R}_j \approx \begin{cases} F(R_j) + \alpha_j G(\beta_j R_j, R_{j+1}) & j < n \\ F(R_j) + \alpha_j G(\beta_j R_j, R_1) & j = n, \end{cases} \tag{16}$$

where

$$\beta_j = \frac{1}{\alpha_j} \frac{\boldsymbol{a}_j^T \boldsymbol{K} \boldsymbol{a}_j}{\boldsymbol{a}_j^T \boldsymbol{a}_j} \tag{17}$$

and α_j is an observable of the weighted neighbor in-degree,

$$\alpha_j = \sum_{i=1}^{N} [\boldsymbol{a}_j]_i k_i^{\text{in}}. \tag{18}$$

To satisfy Eqs. (16), the weight vectors are constrained by the structure and must transform according to

$$\boldsymbol{a}_j = \frac{\boldsymbol{W}^T \boldsymbol{a}_{j-1}}{\alpha_{j-1}}, \tag{19}$$

which preserves the positiveness and the required normalization. Moreover, Eq. (19) needs to be a periodic application, i.e., $\boldsymbol{a}_{j+n} = \boldsymbol{a}_j$, in order to close the system to n observables. The initial choice of $\boldsymbol{a}_1$ is then highly constrained to satisfy this condition. In the following section, we explain how the weight vectors can be computed using the dominant eigenvectors of $\boldsymbol{W}^T$.

In contrast to Eq. (5) where a single observable is used, we have now developed a closed n-dimensional system of observables that are coupled by a set of structural parameters $\{\alpha_j, \beta_j\}_{j=1,\ldots,n}$.

B. Choice of the universal weight vectors

It is yet unclear if one should use an n-dimensional reduction or a one-dimensional reduction for a certain network structure. By answering this question, we also address how to set the weight vectors of the cycle reduction.

Recall that the Perron-Frobenius theorem guarantees that $\boldsymbol{W}^T$ has a non-negative dominant eigenvector $\boldsymbol{v}_D$ only if $\boldsymbol{W}$ is a connected graph. Therefore, we can rule out that it is always possible to construct a one-dimensional reduction relying on the dominant eigenvector.

But the same reasoning also implies that we can always construct an n-dimensional system using the dominant eigenvector $\boldsymbol{v}_D$ of $\boldsymbol{W}^T$. One could use $\boldsymbol{a}_1 = \boldsymbol{v}_D$ and iteratively apply Eq. (19) to obtain the set of weight vectors $\{\boldsymbol{a}_j\}_{1,\ldots,n}$, as prescribed. In doing so, the resulting weight vectors would all be identical, $\boldsymbol{a}_1 = \boldsymbol{a}_2 = \ldots = \boldsymbol{a}_n = \boldsymbol{v}_D$, as they obviously satisfy both Eq. (19) and the periodicity condition $\boldsymbol{a}_{j+n} = \boldsymbol{a}_j$. Hence, we find n identical observables $R_1 = R_2 = \ldots = R_n$, and the constructed n-dimensional system is no better than the one-dimensional system. For this reason, an n-dimensional cycle reduction is only advantageous if we can construct a set of *distinct* weight vectors $\boldsymbol{a}_1 \neq \boldsymbol{a}_2 \neq \ldots \neq \boldsymbol{a}_n$ from the dominant eigenvectors.

The maximum number of significant and distinct observables that we can construct is determined by the periodicity of the transposed adjacency matrix $\boldsymbol{W}^T$. The periodicity n is the number of eigenvalues λ_m of modulus equal to the spectral radius $r \geq 0$, i.e., $|\lambda_m| = r$. From the Perron-Frobenius theorem, they must be uniformly distributed on a circle, centered at the origin, in the complex plane. Thus, the mth dominant eigenvalue can be written as

$$\lambda_m = r e^{2\pi i m/n},$$

for a given periodicity n. Since λ_m is an eigenvalue of $\boldsymbol{W}^T$, it must have an eigenvector $\boldsymbol{v}_m$ that satisfies

$$\boldsymbol{W}^T\boldsymbol{v}_m = r\mathrm{e}^{2\pi i m/n}\boldsymbol{v}_m.$$

By multiplying both sides by $(\boldsymbol{W}^T)^{(n-1)}$, we find

$$(\boldsymbol{W}^T)^n\boldsymbol{v}_m = r^n\boldsymbol{v}_m.$$

Therefore, r^n is an eigenvalue of $(\boldsymbol{W}^T)^n$ that is positive and n-times degenerate. Since each eigenvector of $\boldsymbol{W}^T$ is also an eigenvector of $(\boldsymbol{W}^T)^n$, we can combine those eigenvectors to construct new distinct eigenvectors of $(\boldsymbol{W}^T)^n$ with eigenvalue r^n and use them as weight vectors. We construct the first weight vector as

$$\boldsymbol{a}_1 = \frac{\sum_{m=1}^n c_m\boldsymbol{v}_m}{\sum_{m=1}^n c_m\boldsymbol{1}^T\boldsymbol{v}_m}, \tag{20}$$

where $c_m \in \mathbb{C}$ are arbitrary coefficients. From Eq. (19), we iteratively compute $\boldsymbol{a}_j$ from $\boldsymbol{a}_{j-1}$. By doing so, we satisfy the periodic condition $\boldsymbol{a}_{j+n} = \boldsymbol{a}_j$ and construct distinct weight vectors.

The reduction only requires that we arbitrarily choose $\mathbf{c} = (c_1, \ldots, c_n)$ to construct $\boldsymbol{a}_1$. We propose selecting $\boldsymbol{c}$ by minimizing the scalar product of the first two weight vectors,

$$\boldsymbol{c} = \underset{\boldsymbol{c}}{\operatorname{argmin}}|\boldsymbol{a}_1^T\boldsymbol{a}_2|, \tag{21}$$

where $\boldsymbol{a}_2 = \alpha_1^{-1}\sum_{m=1}^n c_m\lambda_m\boldsymbol{v}_m$ from Eq. (19). In Appendix C, we give a general and exact solution of $\boldsymbol{c}$ for $n = 2$.

In summary, the cycle reduction method goes as follows. First, compute a set of n eigenvectors $\{\boldsymbol{v}_m\}_{m=1,\ldots,n}$ of $\boldsymbol{W}^T$ whose eigenvalues have a modulus equal to the spectral radius r. Second, obtain $\{c_i\}_{i=1,\ldots,n}$ by solving Eq. (21). Third, iteratively construct $\boldsymbol{a}_i$ from Eq. (19). Finally, compute α_i, β_i and solve R_i at equilibrium from Eqs. (16).

An interesting aspect of this method is that it allows us to combine the information of each observable to construct a global observable,

$$R_{\text{global}} = \sum_{j=1}^n \phi_j R_j, \tag{22}$$

where $\phi_j \in \mathbb{R}$. Since $R_j = \sum_i [\boldsymbol{a}_j]_i x_i$, the contribution p_j of node j to the global observable is

$$p_j = \sum_i \phi_i [\boldsymbol{a}_i]_j. \tag{23}$$

We can then tune ϕ_i to reach the desired node contributions. For instance, to access the unweighted average activity $R_{\text{global}} = \langle x\rangle$, one solves $\boldsymbol{A}\boldsymbol{\phi} = N^{-1}\boldsymbol{1}_N$, where $\boldsymbol{A} = [\boldsymbol{a}_1\boldsymbol{a}_2\ldots\boldsymbol{a}_n]$. The solution is $\boldsymbol{\phi} = N^{-1}\boldsymbol{A}^+\boldsymbol{1}_N$, where $\boldsymbol{A}^+$ is the Moore-Penrose pseudo-inverse of matrix $\boldsymbol{A}$ [19].

C. Examples: Star and bipartite networks

We give an example of the cycle reduction for a highly heterogeneous family of networks: star networks. We construct a star network of N nodes, where the strength of a directed edge to a periphery node is s_{pc} and s_{cp} for edges directed toward the central node. Hence, the adjacency matrix is

$$\boldsymbol{W} = \begin{bmatrix} 0 & s_{cp} & s_{cp} & \ldots & s_{cp} \\ s_{pc} & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{pc} & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

One finds that $\boldsymbol{W}^T$ has two eigenvalues of modulus equal to the spectral radius, $\lambda_+ = \sqrt{s_{pc}s_{cp}(N-1)}$ and $\lambda_- = -\sqrt{s_{pc}s_{cp}(N-1)}$. From the previous analysis, this result signals that we can construct a two-dimensional reduction. The associated eigenvectors are

$$\boldsymbol{v}_+^T = \begin{bmatrix} \frac{\lambda_+}{s_{cp}} & 1 & 1 & \ldots & 1 \end{bmatrix}, \tag{24a}$$

$$\boldsymbol{v}_-^T = \begin{bmatrix} \frac{\lambda_-}{s_{cp}} & 1 & 1 & \ldots & 1 \end{bmatrix}. \tag{24b}$$

We construct the two-dimensional reduction by combining the eigenvectors $\boldsymbol{v}_+, \boldsymbol{v}_-$ to minimize $|\boldsymbol{a}_1^T\boldsymbol{a}_2|$ and to satisfy the normalization $\boldsymbol{1}^T\boldsymbol{a}_1 = 1$. The following linear combinations fulfill the requirements

$$\boldsymbol{a}_1 = \frac{\boldsymbol{v}_+ - \boldsymbol{v}_-}{\boldsymbol{1}^T\boldsymbol{v}_+ - \boldsymbol{1}^T\boldsymbol{v}_-}, \tag{25}$$

$$\boldsymbol{a}_2 = \frac{1}{(N-1)s_{cp}}\frac{(\lambda_+\boldsymbol{v}_+ - \lambda_-\boldsymbol{v}_-)}{\boldsymbol{1}^T\boldsymbol{v}_+ - \boldsymbol{1}^T\boldsymbol{v}_-}, \tag{26}$$

where the second vector has been obtained from Eq. (19). Note that the overlap $\boldsymbol{a}_1^T\boldsymbol{a}_2 = 0$ exactly. Explicitly, in component form, we have

$$[\boldsymbol{a}_1]_i = \delta_{i,1}, \tag{27a}$$

$$[\boldsymbol{a}_2]_i = \frac{1}{N-1}(1-\delta_{i,1}). \tag{27b}$$

One also finds that $\beta_1 = 1, \beta_2 = 1, \alpha_1 = s_{cp}(N-1)$, $\alpha_2 = s_{pc}$. Hence, the two-dimensional reduction reads

$$\dot{R}_1 = F(R_1) + s_{cp}(N-1)G(R_1, R_2), \tag{28a}$$

$$\dot{R}_2 = F(R_2) + s_{pc}G(R_2, R_1). \tag{28b}$$

One notes that R_1 is exactly equal to the activity of the central node, and R_2 is the activity of a periphery node.

Thus, the two-dimensional formalism is an exact reduction in this example. This idea is confirmed in Fig. 3 where simulations and predictions are compared for the Gao *et al.*, the one-dimensional reduction, and the two-dimensional reduction for the Cowan-Wilson dynamics.

Star networks are not the only systems conforming to the two-dimensional reduction; all bipartite networks also conform. Bipartite networks have nodes that can be separated into two groups such that connections only exist between nodes of different groups [23]. This network architecture is common in many real systems such as plant-pollinator interactions [26], scientific collaborations [27], and actor-film networks [28].

Bipartite networks exhibit a remarkable and useful spectral property. Each eigenvalue is paired, i.e.,

$$\lambda_j = -\lambda_{N-j+1}, \tag{29}$$

for all $j = 1, 2, \ldots, N$, assuming that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ [29]. Thus, a bipartite graph contains two eigenvalues λ_1, λ_N of modulus equal to the spectral radius. This suggests that a two-dimensional representation could always be constructed out of these two eigenvalues, following the prescription of Sec. IV A.

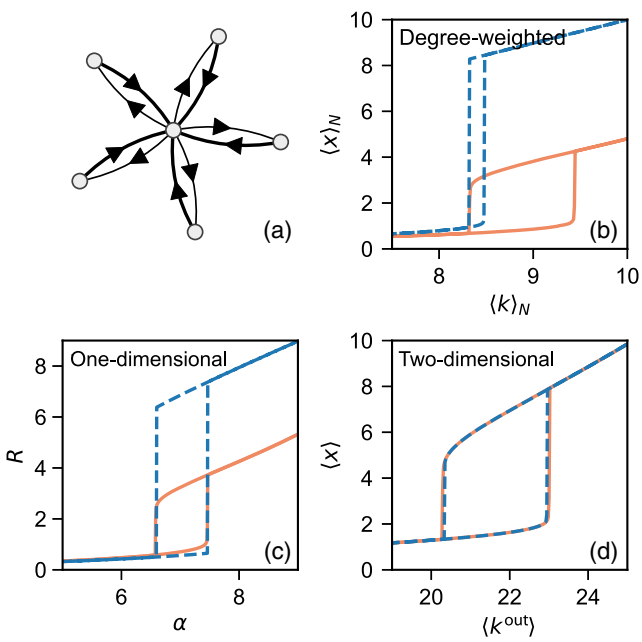

FIG. 3. (a) Schematization of the star network of $N = 6$ nodes where the edge weight toward the core is twice the weight of an edge toward the periphery. (b) Average neighbor activity at equilibrium as a function of its structural parameter using the degree-weighted reduction of Gao *et al.* [9]. (c) Dominant eigenvector weighted activity at equilibrium $R^* = \boldsymbol{v}_D^T \boldsymbol{x}^*$ as a function of the dominant eigenvalue α for the one-dimensional reduction [Eq. (5)]. (d) Average activity at equilibrium obtained by a combination of two observables $\langle x \rangle^* = N^{-1}[R_1^* + (N-1)R_2^*]$, as a function of the average out-degree $\langle k^{\text{out}} \rangle = N^{-1}[\alpha_1 + (N-1)\alpha_2]$ computed using the two-dimensional reduction formalism [Eq. (28b)]. Solid lines are results from simulations, and dashed lines are theoretical predictions. The network dynamics is the Cowan-Wilson model with $\tau = 1$, $\mu = 3$ [Eq. (2a)].

We therefore gain a clear understanding of the results of Jiang *et al.* [10], where they present numerical evidence on real bipartite networks, suggesting that two-dimensional reductions are better predictors of critical points for bipartite networks than the one-dimensional approaches. Moreover, our work solves the problem of selecting the right weighted combination: The eigenvector-weighted combination should always be favored over others.

V. MULTIDIMENSIONAL REDUCTION: INCLUDING SUBDOMINANT EIGENVECTORS

Until now, we have developed a direct method to construct n-dimensional reduced systems. Using only the network structure, we can first identify the number of dimensions, i.e., the number of eigenvalues of modulus equal to the spectral radius, and then construct the weight vectors $\boldsymbol{a}_j$ to predict the observables R_j.

Yet, our method is compelling only if the observables R_j are good indicators of the global states of the network, which requires that each region of the network contributes significantly to at least one observable R_j. Since we use the dominant eigenvectors, we do not control the contribution of each node. Thus, if the dominant eigenvector assigns negligible weights to some nodes, it may result in an incomplete description of the network. Modular networks fall into this category. Let us introduce the stochastic block model (SBM) to understand the underlying problem of misrepresentation.

The SBM is a generative model of modular networks [30]. Nodes are first assigned to modules. Then, we connect a node from module s to a node from module r with probability p_{rs}. This simple method accurately generates modular random networks.

The spectrum of a SBM network is rather different than the spectrum of a random network (Fig. 5). We first note that we only have a single eigenvalue of modulus equal to the spectral radius, indicating that we should use a one-dimensional reduction. However, the eigenvalues are distributed in a multimodal distribution with as many dominant and subdominant eigenvalues as there are modules.

For instance, let us consider a network with two communities of equal size $N/2$. Using the spectral theory of random matrices, we estimate the two dominant eigenvalues

$$\lambda_1 = \frac{(p_{11} + p_{22}) + [(p_{11} - p_{22})^2 + 4p_{12}p_{21}]^{1/2}}{2},$$
$$\lambda_2 = \frac{(p_{11} + p_{22}) - [(p_{11} - p_{22})^2 + 4p_{12}p_{21}]^{1/2}}{2},$$

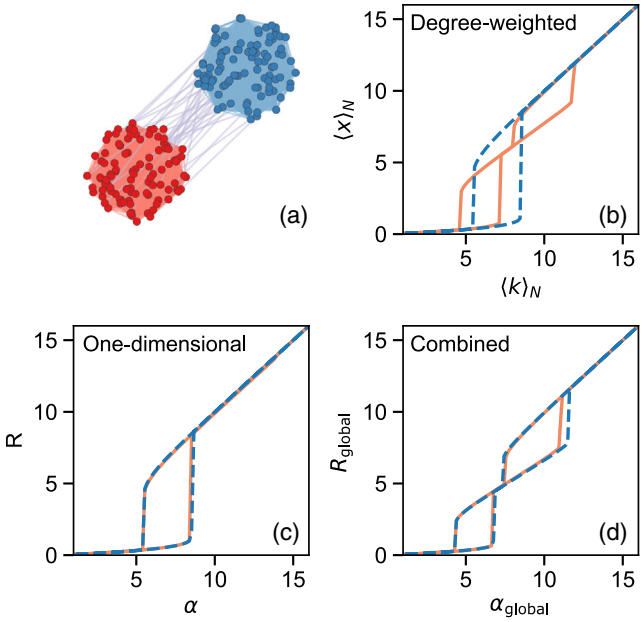

FIG. 4. (a) Schematization of the undirected planted partition network of two communities of 100 nodes each with in-densities $p_{\text{in}} = 0.4$ and $p_{\text{in}} = 0.7$, and out-density $p_{\text{out}} = 3 \times 10^{-3}$. (b) Average neighbor activity at equilibrium as a function of its structural parameter using the degree-weighted reduction of Gao *et al.* [9]. (c) Dominant eigenvector-weighted activity $R = \boldsymbol{v}_D^T \boldsymbol{x}$ at equilibrium as a function of the associated eigenvalue α for the one-dimensional reduction [Eq. (9)]. (d) Combination of the uncoupled observables $R_{\text{global}} = (R_D + R_{SD})/2$ at equilibrium as a function of the structural parameter α_{global} [see Eq. (30b)]. Solid lines result from simulations, and dashed lines are theoretical predictions. The dynamics is the Cowan-Wilson model with $\tau = 1$, $\mu = 3$ [Eq. (2a)].

and their corresponding eigenvectors

$$\boldsymbol{v}_1 = \begin{bmatrix} \frac{p_{12}}{p_{12}+\lambda_1-p_{11}} \mathbf{1}_{N/2} \\ \frac{\lambda_1-p_{11}}{p_{12}+\lambda_1-p_{11}} \mathbf{1}_{N/2} \end{bmatrix},$$

$$\boldsymbol{v}_2 = \begin{bmatrix} \frac{\lambda_2-p_{22}}{\lambda_2-p_{22}+p_{21}} \mathbf{1}_{N/2}, \\ \frac{p_{21}}{\lambda_2-p_{22}+p_{21}} \mathbf{1}_{N/2} \end{bmatrix}.$$

If $p_{rs}/p_{rr} \approx 0 \;\forall\; r \neq s$, the dominant eigenvector $\boldsymbol{v}_1$ assigns a negligible weight to the nodes in the community $r = 1$ [31]. Thus, if we solely use a one-dimensional reduction with $\boldsymbol{a} = \boldsymbol{v}_1$, the observable R will not take into account the activity of half the network. Fortunately, the second-dominant eigenvector accounts for the remaining nodes. Thus, we must apply the one-dimensional reduction of Sec. III A twice and construct two uncoupled observables, one with the dominant eigenvector $R_D = \boldsymbol{v}_1^T \boldsymbol{x}$, and one for the subdominant eigenvector $R_{SD} = \boldsymbol{v}_2^T \boldsymbol{x}$, for which the dynamics follows,

$$\dot{R}_D = F(R_D) + \alpha_D G(\beta_D R_D, R_D), \tag{30a}$$

$$\dot{R}_{SD} = F(R_{SD}) + \alpha_{SD} G(\beta_{SD} R_{SD}, R_{SD}). \tag{30b}$$

To make a global prediction, we simply combine the observables

$$R_{\text{global}} = \frac{R_D + R_{SD}}{2}. \tag{31}$$

It follows that the global structural parameter is also a linear composition

$$\alpha_{\text{global}} = \frac{\alpha_D + \alpha_{SD}}{2}. \tag{32}$$

In general, we can construct as many uncoupled observables as the number of modules in the network, by using the eigenvectors associated with the eigenvalues detached from the bulk of the spectrum.

We show a numerical example of this method for a network of two communities (Fig. 4). The Gao *et al.* formalism predicts a single bifurcation, which almost coincides with the transition of the densest community. In Fig. 4(b), the one-dimensional reduction, using the first dominant eigenvector, predicts the activity accurately. However, the first dominant eigenvector omits half the network, and we only see a single bifurcation. Thus, even if we are highly accurate, we miss characterizing the distinctive multistep bifurcation, a prominent feature of interacting

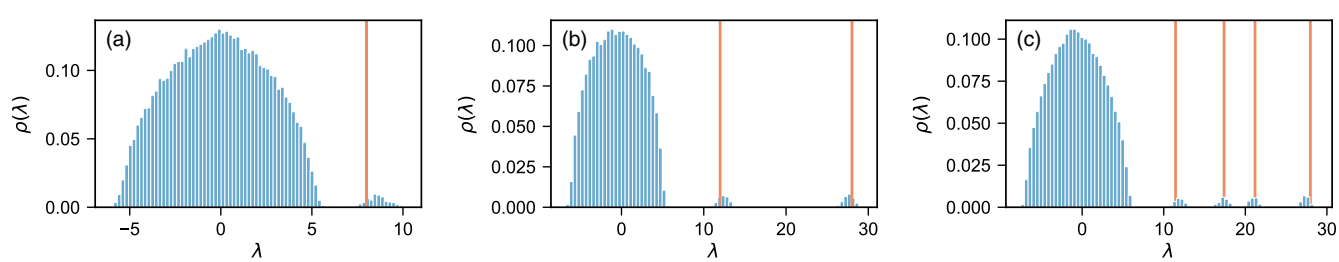

FIG. 5. Spectral density $\rho(\lambda)$ for (a) random networks of 80 nodes with connection probability $p = 0.1$, (b) the SBM with two communities of 40 nodes with in-connection probabilities $p_{11} = 0.3$ and $p_{22} = 0.7$ and out-connection probabilities $p_{12} = p_{21} = 0.01$, and (c) the SBM with four communities of 40 nodes with in-connection probabilities $p_{11} = 0.4$, $p_{22} = 0.6$, $p_{33} = 0.7$, and $p_{44} = 0.9$ and out-connection probability $p_{rs} = 0.05 \;\forall\; r \neq s$. The orange lines highlight the position of the dominant and subdominant eigenvalues in the expected network. Each spectral density is produced by collecting spectra from 500 network instances.

networks [32]. Finally, combining observables as in Eq. (31), we recover the bifurcations of both modules [Fig. 4(d)].

VI. GOODNESS OF REDUCTION

The goodness of the reduction method, i.e., how accurate the predictions of the low-dimensional representation are compared with the observations on the original network, depends on the nature of the dynamics and on the network structure. For instance, some complex patterns of interactions may be less amenable to a low-dimensional formalism, resulting in disparities between the predicted and exact values of the observables. In this section, we explore the impact of the structure and the dynamics on the goodness of the reduction methods.

A. Impact of the structure

To measure the impact of the structure, we introduce a generative model of networks called the generalized preferential attachment model [33]. Parameters of this model can be continuously tuned to obtain networks ranging from chainlike networks to star networks, with scale-free networks as an intermediate state. Scale-free systems are an important family of networks, recognizable by their power-law degree distribution $p(k) \sim k^{-\gamma}$ [20]. Because of their lack of well-defined characteristic scale, it is *a priori* unclear whether these systems can be efficiently reduced.

The growth process of the generalized preferential attachment goes as follows. We initialize the network with two connected nodes. Then, at each time step t, we add a new node to the network. It is connected to an existing node chosen with probability

$$w_i(\nu, t) = \frac{s_i^\nu(t)}{\sum_{j=1}^{N(t)} s_j^\nu(t)}, \tag{33}$$

where $\nu \in \mathbb{R}$ is the exponent of the attachment kernel, $s_j(t)$ is the number of connections of node j at time t, and $N(t)$ is the number of nodes at time t.

The generative model is solely tuned by the kernel parameter $\nu \in \mathbb{R}$. It controls the inequalities of the attachment probability. On the one hand, if $\nu \gg 1$, the generated networks are starlike as we always attach new nodes to the richest node. On the other hand, if $\nu \ll 0$, the networks are more chainlike as we always connect to the least-connected node [33]. The classic preferential attachment model is found for $\nu = 1$. Therefore, for $0 < \nu < 1$, we observe a continuum of network organizations that gradually become more scale-free as the parameter ν is increased. Examples of networks generated from this model are illustrated in Fig. 6.

From now on, we distinguish the predicted observable from the reduced system, denoted $\tilde{R}(\alpha)$, and the measured observable $R(\alpha) = \boldsymbol{a}^T\boldsymbol{x}$ from the original network.

We have applied the degree-weighted, the one-dimensional, and the two-cycle reductions to networks generated with the generalized preferential attachment model for $\nu \in [-1, 2]$. For each network, we have computed the total error Δ_R between the measured activity $R^*(\alpha) = \boldsymbol{a}^T\boldsymbol{x}^*$ at equilibrium on the original network and the predicted activity $\tilde{R}^*(\alpha)$ by the reduction system,

$$\Delta_R = \int_0^\infty |\tilde{R}^*(\alpha) - R^*(\alpha)| d\alpha. \tag{34}$$

We have found a transition in the dimension reduction accuracy for all methods at $\nu = 1$, corresponding to the preferential attachment model (Fig. 7). As we enter the starlike region $\nu > 1$, the average error reaches a plateau to specific values for the one-dimensional reductions, while the two-dimensional reduction remains highly effective.

We argue that this transition is not dynamics specific but mostly due to the network architecture. The nature of the error can be interpreted by a careful examination of the generative model. First, negative values of ν tend to homogenize the degree of the nodes. Thus, the more uniform the network is, the easier it is to capture its behavior in a one-dimensional reduction. For positive values of ν, the reduced model tends to favor degree inequalities, which are best achieved when the networks are starlike.

For $0 < \nu < 1$, the degree distribution resembles a power law with exponential cutoffs [33]. However, a pure power-law distribution is only achieved precisely at $\nu = 1$. Thus, this transition in the degree distribution forces the reduction of the degree-weighted reduction to predict inaccurate

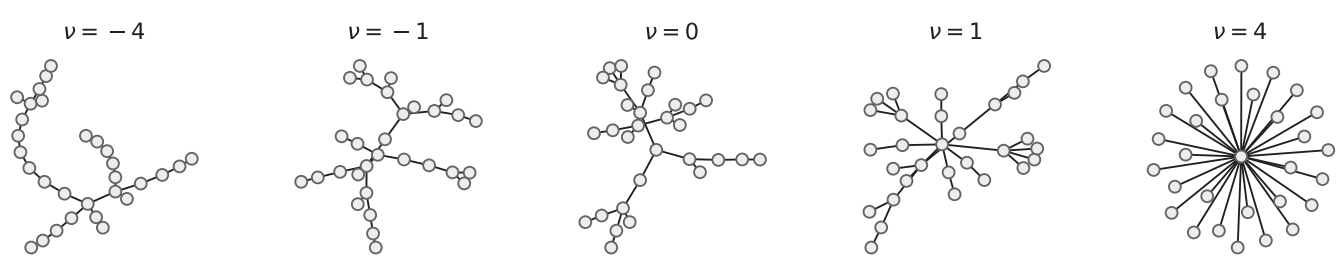

FIG. 6. Network instances produced using the generalized preferential attachment model. The model tends to generate chainlike networks for negative values of ν and star networks for positive values. The figure was inspired by Ref. [34].

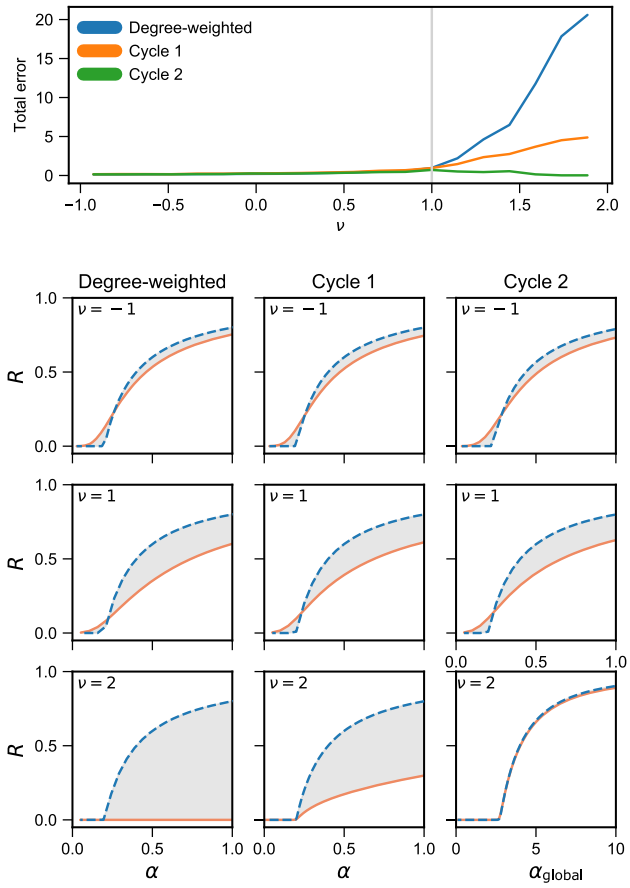

FIG. 7. (Top) Total error on the prediction of the global state activity at equilibrium for networks obtained from the generalized preferential attachment model [see Eq. (34)]. A gray line indicates the classical preferential attachment model at $\nu = 1$ where a critical transition in the reduced descriptions is found. The total error is averaged from 300 networks of $N = 200$ generated uniformly on the domain $\nu \in [-1, 2]$. The activity on the network is the SIS model with $\gamma = 1$ [see Eq. (2e)]. (Bottom) Instances of bifurcation diagrams for the three reduction schemes (columns) and for $\nu = \{-1, 1, 2\}$ (rows). For the two-dimensional cycle reduction, the x axis is the average of the structural parameters: $\alpha_{\text{global}} = (\alpha_1 + \alpha_2)/2$. The blue dashed lines are predictions from reduced systems, and the orange lines are the measured activities on the original networks. The gray regions indicate the absolute errors [see Eq. (34)].

observables. Finally, the region $\nu > 1$ is dominated by starlike networks, which has previously been shown to be better represented by the two-dimensional reduction than any one-dimensional reduction (see Sec. IV C).

B. Detection of transitions

Most dynamical systems exhibit activity bifurcations when a certain structural threshold is reached [35]. Therefore, the goodness of the reduction should display at least qualitative changes as the structural threshold is crossed. We investigate this kind of prediction using the SIS model.

The SIS model has been thoroughly studied in the last decade [36]. In the SIS model, nodes reversibly switch from susceptible to infected states with a certain probability that depends on their neighborhood. We can formulate this dynamics using a mean-field approach,

$$\dot{x}_i = -x_i + \gamma(1 - x_i)\sum_{j=1}^{N} w_{ij}x_j, \tag{35}$$

where x_i is the probability that node i is infected and $\gamma \geq 0$ is the normalized infection rate. In this model, the average fraction of infected node $\langle x \rangle = N^{-1}\sum_i x_i$ undergoes a critical transition at a certain threshold γ_C. The classical problem in the study of the SIS model consists in estimating the value of γ_C above which a significant fraction of the whole system is infected [37].

However, we study a related problem: The parameter γ is fixed, and the structure is evolving. Using the dimension reduction procedure, we investigate the critical structural parameter α_{global}, or an equivalent parameter depending on the reduction approach, for which the global state at equilibrium R^* undergoes a critical transition characterized by

$$\frac{d^2R^*(\alpha_C)}{d\alpha^2} = 0. \tag{36}$$

In Fig. 8, we investigate the errors on the position of the critical transition for the degree-weighted one-dimensional approach, the eigenvector-weighted one-dimensional approach, and the two-dimensional *cycle* reduction. We use a network of $N = 60$ nodes generated from the generalized preferential attachment model with $\nu = 1.8$. We observe that the two proposed approaches based on dominant eigenvectors are able to accurately predict the critical transition, while the degree-weighted approach does not. This behavior is typical for reductions of networks generated from $\nu \in [1, \infty)$. Our results indicate that even if the one-dimensional observable fails to predict the true level of activity R, it still predicts, with high accuracy, the onset of the epidemy. We conclude that the largest eigenvalue is a reliable indicator of the onset for correlated networks. Perhaps this conclusion is not surprising as it has previously been discovered under a different approach [38]. Nonetheless, it supports the proposed reductions as valuable candidates for predicting the onset of critical transitions.

C. Impact of the dynamics

As previously discussed, the goodness of the reduction is highly dependent on the network structure. The other element that impacts the goodness of the reduction is the nature of the dynamics. For instance, linear dynamics such as $F(x_i) = x_i$, $G(x_i, x_j) = x_j$ lead to exact one-dimensional reduction. However, typical dynamics are nonlinear and

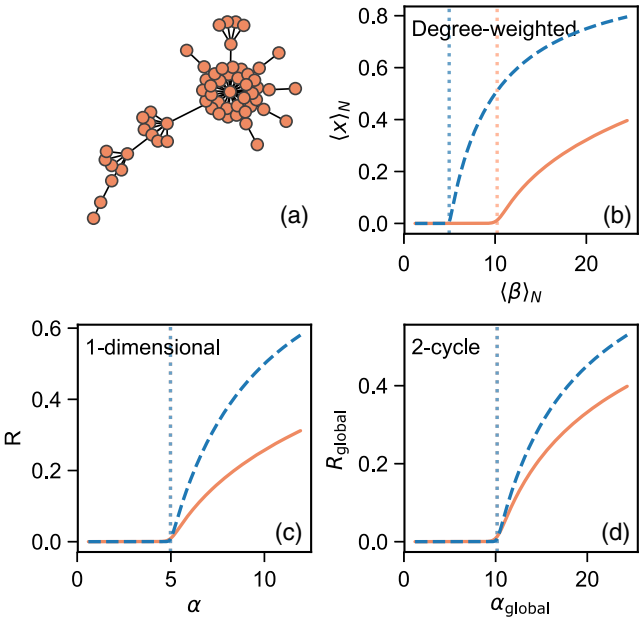

FIG. 8. (a) Schematization of the undirected network generated using the generalized preferential attachment model with $\nu = 1.8$ and $N = 60$. (b) Average neighbor activity at equilibrium as a function of its structural parameter using the degree-weighted reduction [9]. (c) Dominant eigenvector-weighted activity at equilibrium $R^* = \boldsymbol{v}_D^T \boldsymbol{x}^*$ as a function of the associated eigenvalue α for the one-dimensional reduction [Eq. (5)]. (d) Solution at equilibrium of the two-dimensional system $R^*_{\text{global}} = (R_1^* + R_2^*)/2$ as a function of the structural parameter $\alpha_{\text{global}} = (\alpha_1 + \alpha_2)/2$. Solid lines are results from simulations, and dashed lines are theoretical predictions. Dotted lines indicate the position of the transition α_{global}. The dynamics is the SIS model with $\gamma = 0.2$ [Eq. (2e)].

may add significant contributions to the quadratic terms that have been neglected (see Appendix A). We investigate this aspect by looking at two contrasting dynamics: the Cowan-Wilson model and the Lotka-Volterra model.

First, let us introduce the relative error Δ_α on the structural parameter that predicts the original network activity. We compare the measured structural parameter $\alpha = \boldsymbol{a}^T \boldsymbol{k}^{\text{in}}$ on the original network with the structural parameter $\tilde{\alpha}$ that matches, from the one-dimensional reduction, the measured activity $R^* = \boldsymbol{a}^T \boldsymbol{x}^*$ on the original network. The relative error can be written as

$$\Delta_\alpha = \frac{\alpha - \tilde{\alpha}(R^*)}{\alpha}.$$

Notice that this is different from Eq. (34): Δ_α is the horizontal error in the space (α, R), while Δ_R measures the vertical error. Both errors convey distinct information and are complementary. However, we use Δ_α since it can be written as a function that depends explicitly on the nature of the dynamics for the one-dimensional reduction.

From Eq. (5), we obtain $\tilde{\alpha}$ at the dynamical equilibrium $\dot{R} = 0$. This result leads to

$$\Delta_\alpha = 1 + \frac{F(R^*)}{G(R^*)\alpha}. \tag{37}$$

One can then evaluate the error for specific dynamics as in Ref. [39]. In the following paragraphs, we give two examples of dynamics and compare the errors for the Gao *et al.* formalism and our one-dimensional reduction.

1. Error on the Cowan-Wilson dynamics

The Cowan-Wilson dynamics describes the firing-rate activity of populations of neurons. The evolution of a node activity is given by Eq. (2a) and repeated here as

$$\dot{x}_i = -x_i + \sum_{j=1}^{N} w_{ij} \frac{1}{1 + \exp[-\tau(x_j - \mu)]}, \tag{38}$$

where $\tau > 0$. The equilibrium solution $\boldsymbol{x}^*$ cannot be found analytically, so it must be evaluated numerically, even for $N = 1$. However, $\boldsymbol{x}^*$ is well approximated for extreme values of activities. We derive error estimates for extreme regimes $x_j^* \gg \mu$ and $x_j^* \ll \mu$.

In general, from Eq. (37), the error can be written as

$$\Delta_\alpha = 1 - \frac{[1 + \exp[-\tau(R^* - \mu)]]R^*}{\alpha}. \tag{39}$$

In the limit of high levels of activity $x_j^* \gg \mu$, the exponential vanishes, and one finds that $\boldsymbol{x}^* \approx \boldsymbol{W}\boldsymbol{1} = \boldsymbol{k}^{\text{in}}$. Therefore, the error is approximately

$$\Delta_\alpha \approx 1 - \frac{R^*}{\alpha} \approx 1 - \frac{\boldsymbol{a}^T \boldsymbol{W}\boldsymbol{1}}{\alpha}.$$

For the one-dimensional reduction, $\boldsymbol{a}$ is an eigenvector of $\boldsymbol{W}^T$ such that $\boldsymbol{a}^T \boldsymbol{W} = \alpha \boldsymbol{a}^T$. Since $\boldsymbol{a}$ is normalized, $\boldsymbol{a}^T \boldsymbol{1} = 1$, it follows that $\boldsymbol{a}^T \boldsymbol{W}\boldsymbol{1} = \alpha$. Thus, the error Δ_α vanishes in the limit of large activity.

The same applies for Gao *et al.* formalism for high activity. By using $\boldsymbol{a} = \boldsymbol{k}^{\text{in}}$, one finds that $\alpha = (\boldsymbol{k}^{\text{in}})^T \boldsymbol{k}^{\text{out}}$ and $R^* = (\boldsymbol{k}^{\text{out}})^T \boldsymbol{k}^{\text{in}}$ so that $\Delta_\alpha \to 0$.

Using a similar procedure for $x_j \ll \mu$, one can also show that $\Delta_\alpha \approx 0$ for both methods.

We conclude that in the extreme regimes of high and low activities, both reduction methods provide a practically exact solution. However, we are more often interested in the hysteresis region where the activity collapses rapidly. Unfortunately, analytic error estimates are lacking in this regime. Still, numerical results and theoretical insights suggest that the proposed one-dimensional reduction should always be favored over the degree-weighted

reduction. This idea is confirmed below for a more tractable dynamics.

2. Error on the Lotka-Volterra dynamics

Let us consider the Lotka-Volterra dynamics governing the evolution of species populations. The N-dimensional system goes as

$$\dot{\boldsymbol{x}} = \omega \boldsymbol{x} + \boldsymbol{x} \circ \boldsymbol{W}\boldsymbol{x}, \tag{40}$$

where $\circ$ denotes an elementwise multiplication. At equilibrium, $\boldsymbol{x}^*$ satisfies

$$-\omega \mathbf{1} = \boldsymbol{W}\boldsymbol{x}^*. \tag{41}$$

With Eq. (37), we write the expected error as

$$\Delta_\alpha = 1 + \frac{\omega}{\beta \alpha R^*} = 1 + \frac{\omega}{\beta \alpha \boldsymbol{a}^T \boldsymbol{x}^*}.$$

Using the one-dimensional reduction,

$$\boldsymbol{W}^T \boldsymbol{a} = \alpha \boldsymbol{a},$$

or $\boldsymbol{a}^T \boldsymbol{W} = \alpha \boldsymbol{a}^T$. Furthermore,

$$\boldsymbol{a}^T \boldsymbol{W}\boldsymbol{x}^* = -\omega \boldsymbol{a}^T \mathbf{1} = \alpha \boldsymbol{a}^T \boldsymbol{x}^*. \tag{42}$$

Thus, the error depends only on β:

$$\Delta_\alpha = 1 - \frac{1}{\beta}. \tag{43}$$

Therefore, the difference between the exact value for $\boldsymbol{a}^T \boldsymbol{x}^*$ and the approximate value derived from the one-dimensional reduced system is only $1 - 1/\beta$. Given the expression for β [Eq. (A5)], it should be close to $\beta \approx 1$, so the error goes to zero, $\Delta_\alpha \approx 0$. This result contrasts with results derived by Tu *et al.* [39] for the method of Gao *et al.*, where they reported nonvanishing error averages and variances.

VII. CONCLUSION

We have built systematic methods of dimension reduction adapted to different families of networks (random, starlike, bipartite, SBM). The activity of the reduced systems is used as an indicator of the global activity of large networks. Without a restriction further than imposing a linear form of the global activity, we have found that the dominant eigenvectors of the adjacency matrix are central to the global states' evolution. Moreover, when considering the *cycle* reduction, the dimension of the reduced systems corresponds to the periodicity of the adjacency matrix.

We have further shown that the proposed reduction of Gao *et al.* is a special case of the general scheme when applied to uncorrelated random networks. Moreover, the range of applicability of our method extends to modular, heterogeneous, and bipartite networks.

Our results suggest, both numerically and theoretically, that the eigenvector-weighted reduction should be preferred over the degree-weighted reduction. Originally, the degree-weighted reduction has been used to approximate the state of a network by the state of the average neighbor. However, the degree of a node is only a local centrality measure since it does not provide information about whom a node is connected to. In contrast, in the eigenvector-weighted reduction, the dominant eigenvector yields a more global node centrality since it contains the information on how each node is connected with the rest of the network [23]. Therefore, the eigenvector-based reduction brings a new light on the influence of each node on the global states of a network.

On a more practical side, our general method is able to predict the correct number of bifurcation points. The expected number of predicted bifurcation points depends on the dimension of the reduction. Intuitively, and confirmed by our investigations, a single linear observable provides a good reduction if the network is homogeneous or if, for instance, the degree variance is small. When this is not the case, however, different parts of the network behave differently, and a single observable is no longer sufficient to capture the characteristics of the global dynamics, which effectively becomes multidimensional. In the SBM case, the two-dimensional reduction reveals additional bifurcation points that are missed altogether by all one-dimensional reductions.

As a closing remark, although our reduction method has been designed to access large dynamical networks through low-dimensional formalisms, it was not clear from the outset how the dimensional reductions would fare with respect to the size of the networks. Our findings on the matter have been comforting since size by itself has a secondary effect on the quality of the reduction procedure, leaving precedence to connectivity and dynamics. Hence, beyond the addition to the theoretical arsenal, our systematic and versatile approach can now be used to address concrete problems of real-world systems. To name a few, it could be used to describe, with high accuracy, the bifurcation patterns, to identify dynamical vulnerabilities, to suggest intervention strategies to prevent dynamical breakdowns, or to classify networks on a standardized diagram.

ACKNOWLEDGMENTS

We are thankful to Charles Murphy, Guillaume St-Onge, Vincent Thibeault, and Jean-Gabriel Young for useful comments and suggestions. This work was funded by the Fonds de recherche du Québec - Nature et technologies (E. L., P. D.), the Natural Sciences and Engineering Research Council of Canada (L. J. D.), and the program Sentinel North, financed by the Canada First Research Excellence Fund (E. L., P. D., L. J. D., N. D.).

APPENDIX A: DERIVATION OF ONE-DIMENSIONAL FORMALISM

In this section, we detail the analytical derivation of the evolution of R for the one-dimensional reduction. We consider the observable

$$R = \sum_{i=1}^{N} a_i x_i$$

with $a_i \in \mathbb{R}$ and $\sum_i a_i = \mathbf{1}^T \boldsymbol{a} = 1$. We first take the time derivative and insert Eq. (1), which leads to

$$\begin{aligned}\dot{R} &= \sum_{i=1}^{N} a_i \dot{x}_i \\ &= \sum_{i=1}^{N} a_i \left[F(x_i) + \sum_{j=1}^{N} w_{ij} G(x_i, x_j) \right].\end{aligned}$$

We wish to show that if $\boldsymbol{a}$ is chosen correctly, then the right-hand side can be written, up to second-order corrections, in terms of R only. To do so, we develop each function around the observable:

$$F(x_i) = F(R) + (x_i - R)F'(R) + \mathcal{O}[(x_i - R)^2]. \quad \text{(A1)}$$

Thus,

$$\begin{aligned}\sum_{i=1}^{N} a_i F(x_i) &= F(R) + F'(R) \sum_{i=1}^{N} a_i (x_i - R) + \mathcal{O}[(x_i - R)^2] \\ &= F(R) + \mathcal{O}[(x_i - R)^2].\end{aligned}$$

This result means that $F(x_i)$ does not impose any constraint on $\boldsymbol{a}$. Now, for the function $G(x_i, x_j)$, we develop around $x_i = \beta R$ and $x_j = \gamma R$:

$$\begin{aligned}G(x_i, x_j) \approx\ & G(\beta R, \gamma R) + (x_i - \beta R) G_1(\beta R, \gamma R) \\ & + (x_j - \gamma R) G_2(\beta R, \gamma R),\end{aligned}$$

where second-order terms have been neglected. Letting $\alpha = \sum_{i,j} a_i w_{ij}$, we find that $\sum_{i,j} a_i w_{ij} G(x_i, x_j)$ is given by

$$\begin{aligned}\sum_{i,j} a_i w_{ij} G(x_i, x_j) \approx\ & \alpha G(\beta R, \gamma R) \\ & + G_1(\beta R, \gamma R) \sum_{i,j} a_i w_{ij} (x_i - \beta R) \\ & + G_2(\beta R, \gamma R) \sum_{i,j} a_i w_{ij} (x_j - \gamma R).\end{aligned}$$

The left-hand side is a function of R only if the linear terms cancel out exactly, which is possible if and only if

$$\alpha \beta R = \sum_{i,j} a_i w_{ij} x_i = \boldsymbol{x}^T \boldsymbol{K} \boldsymbol{a}, \quad \text{(A2a)}$$

$$\alpha \gamma R = \sum_{i,j} a_i w_{ij} x_j = \boldsymbol{x}^T \boldsymbol{W}^T \boldsymbol{a}. \quad \text{(A2b)}$$

Since $R = \boldsymbol{x}^T \boldsymbol{a}$, we conclude that the last two equations are satisfied for all $\boldsymbol{x} \in \mathbb{R}^N$ only if $\boldsymbol{a}$ is an eigenvector of both matrices $\boldsymbol{K}$ and $\boldsymbol{W}^T$, with corresponding eigenvalues $\beta\alpha$ and $\gamma\alpha$. Although we cannot solve these two equations simultaneously, in general, we can enforce that at least one equation is satisfied exactly. Choosing $\boldsymbol{a}$ as an eigenvector of $\boldsymbol{W}^T$, we can prove that $\gamma = 1$ if the vector $\boldsymbol{a}$ is normalized $\mathbf{1}^T \boldsymbol{a} = 1$: If $\boldsymbol{W}^T \boldsymbol{a} = \lambda \boldsymbol{a}$ and $\lambda = \alpha\gamma$, then $\lambda = \mathbf{1}^T \boldsymbol{W}^T \boldsymbol{a} = \boldsymbol{a}^T \boldsymbol{W} \mathbf{1} = \boldsymbol{a}^T \boldsymbol{k}^{\text{in}} = \alpha$, so $\gamma = 1$.

We then choose β to best satisfy Eq. (A2a) by minimizing the mean squared error (MSE):

$$\beta^* = \underset{\beta}{\operatorname{argmin}} ||\boldsymbol{K}\boldsymbol{a} - \beta\alpha\boldsymbol{a}||^2, \quad \text{(A3)}$$

where the symbol $||\cdot||$ denotes the standard Euclidean norm. Basic calculus leads to

$$\beta^* = \frac{1}{\alpha} \frac{\boldsymbol{a}^T \boldsymbol{K} \boldsymbol{a}}{\boldsymbol{a}^T \boldsymbol{a}} = \frac{1}{\alpha} \frac{\sum_i a_i^2 k_i^{\text{in}}}{\sum_i a_i^2}. \quad \text{(A4)}$$

Note that β^* is a ratio of weighted averages,

$$\beta^* = \frac{\boldsymbol{b}^T \boldsymbol{k}^{\text{in}}}{\boldsymbol{a}^T \boldsymbol{k}^{\text{in}}}, \quad \text{(A5)}$$

where $\boldsymbol{b}$ is normalized, $\mathbf{1}^T \boldsymbol{b} = 1$, and has elements $b_i = a_i^2 / \sum_{i=1}^{N} a_i^2$. From the construction of $\boldsymbol{b}$, we deduce that $\boldsymbol{b}$ must be similar to $\boldsymbol{a}$ and β^* close to 1, which has been confirmed throughout most of the simulations.

APPENDIX B: DERIVATION OF THE MULTIDIMENSIONAL CYCLE FORMALISM

For the cycle reduction, we construct n observables,

$$R_k = \sum_{i=1}^{N} [\boldsymbol{a}_k]_i x_i, \quad \text{(B1)}$$

with normalized weights $\mathbf{1}^T \boldsymbol{a}_j = 1$. Using Eq. (1), we find that the dynamics of R_k is equal to

$$\dot{R}_k = \sum_i [\boldsymbol{a}_k]_i F(x_i) + \sum_{i,j} [\boldsymbol{a}_k]_i w_{ij} G(x_i, x_j). \quad \text{(B2)}$$

As for the one-dimensional reduction, one finds that

$$\sum_i [\boldsymbol{a}_k]_i F(x_i) \approx F(R_k) \quad \text{(B3)}$$

up to the second order of corrections. We then develop $G(x_i, x_j)$ around $x_i = \beta_k R_k$ and $x_j = \gamma_k R_{k+1}$, which yields

$$\begin{aligned} G(x_i, x_j) \approx\, & G(\beta_k R_k, \gamma_k R_{k+1}) \\ & + (x_i - \beta_k R_k) G_1(\beta_k R_k, \gamma_k R_{k+1}) \\ & + (x_j - \gamma_k R_{k+1}) G_2(\beta_k R_k, \gamma_k R_{k+1}). \end{aligned}$$

Using the same arguments as in the one-dimensional reduction, one can prove that $\sum_{i,j} [\boldsymbol{a}_k]_i w_{ij} G(x_i, x_j) \approx \alpha_k G(\beta_k R_k, \gamma_k R_{k+1})$, with $\alpha_k = \boldsymbol{a}^T \boldsymbol{k}^{\mathrm{in}}$, only if the following equations are satisfied simultaneously:

$$\boldsymbol{x}^T \boldsymbol{K} \boldsymbol{a}_k = \beta_k R_k \alpha_k, \tag{B4}$$

$$\boldsymbol{x}^T \boldsymbol{W}^T \boldsymbol{a}_k = \gamma_k R_{k+1} \alpha_k. \tag{B5}$$

The second equation is satisfied if

$$\boldsymbol{W}^T \boldsymbol{a}_k = \alpha_k \boldsymbol{a}_{k+1}, \tag{B6}$$

with $\gamma_k = 1$. After n applications of Eq. (B6), we close the system with $a_{n+1} = a_1$, which is the expected result if $\boldsymbol{a}_1$ is an eigenvector of $(\boldsymbol{W}^T)^n$.

As for the parameter β_k, we minimize the MSE,

$$\beta_k^* = \underset{\beta_k}{\operatorname{argmin}} ||\boldsymbol{K} \boldsymbol{a}_k - \beta_k \alpha_k \boldsymbol{a}_k||^2, \tag{B7}$$

and find

$$\beta_k = \frac{1}{\alpha_k} \frac{\boldsymbol{a}_k^T \boldsymbol{K} \boldsymbol{a}_k}{\boldsymbol{a}_k^T \boldsymbol{a}_k}. \tag{B8}$$

APPENDIX C: COMBINATION FOR TWO-CYCLE REDUCTION

In this Appendix, we show that, for an adjacency matrix $\boldsymbol{W}^T$, non-negative and of periodicity 2, the first weight vector $\boldsymbol{a}_1$ of the reduction is an equipartition of the two dominant eigenvectors.

For such an adjacency matrix, the Perron-Frobenius theorem states that $\boldsymbol{W}^T$ admits two eigenvalues of modulus equal to the spectral radius. The eigenvectors satisfy

$$\boldsymbol{W}^T \boldsymbol{v}_1 = r \boldsymbol{v}_1; \qquad \boldsymbol{W}^T \boldsymbol{v}_2 = -r \boldsymbol{v}_2, \tag{C1}$$

where r is the spectral radius and $\boldsymbol{v}_i$ are the eigenvectors normalized as $\boldsymbol{v}_i^T \boldsymbol{v}_i = 1$.

Now, let us consider the first weight vector as a linear combination of the dominant eigenvectors:

$$\boldsymbol{a}_1 = \frac{c_1 \boldsymbol{v}_1 + c_2 \boldsymbol{v}_2}{c_1 \boldsymbol{1}^T \boldsymbol{v}_1 + c_2 \boldsymbol{1}^T \boldsymbol{v}_2}. \tag{C2}$$

From transformation (19), we get the equation

$$\boldsymbol{a}_2 = \frac{1}{\alpha_1} \frac{c_1 r \boldsymbol{v}_1 - c_2 r \boldsymbol{v}_2}{c_1 \boldsymbol{1}^T \boldsymbol{v}_1 + c_2 \boldsymbol{1}^T \boldsymbol{v}_2}. \tag{C3}$$

We want to find c_1, c_2 such that $c_1 + c_2 = 1$ and $S = |\boldsymbol{a}_1^T \boldsymbol{a}_2|$ is minimized. The former condition is chosen for definiteness, while the latter condition favors the weight vectors that represent almost exclusive groups of nodes. The scalar product is then simply

$$S \propto |c_1^2 - c_2^2|,$$

and it is minimized with $c_1 = c_2 = 1/2$.

[1] M. Scheffer, *Critical Transitions in Nature and Society* (Princeton University, Princeton, NJ, 2009).

[2] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D. R. White, *Economic Networks: The New Challenges*, Science **325**, 422 (2009).

[3] S. Gualdi, M. Tarzia, F. Zamponi, and J.-P. Bouchaud, *Tipping Points in Macroeconomic Agent-Based Models*, J. Econ. Dyn. Control **50**, 29 (2015).

[4] M. P. Richardson, *Large Scale Brain Models of Epilepsy: Dynamics Meets Connectomics*, J. Neurol. Neurosurg. Psychiatry **83**, 1238 (2012).

[5] R. V. Sole and M. Montoya, *Complexity and Fragility in Ecological Networks*, Proc. R. Soc. B **268**, 2039 (2001).

[6] C. Boettiger and A. Hastings, *Tipping Points: From Patterns to Predictions*, Nature (London) **493**, 157 (2013).

[7] F. Morone and H. A. Makse, *Influence Maximization in Complex Networks through Optimal Percolation*, Nature (London) **524**, 65 (2015).

[8] L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y. C. Zhang, and T. Zhou, *Vital Nodes Identification in Complex Networks*, Phys. Rep. **650**, 1 (2016).

[9] J. Gao, B. Barzel, and A.-L. Barabási, *Universal Resilience Patterns in Complex Networks*, Nature (London) **530**, 307 (2016).

[10] J. Jiang, Z.-G. Huang, T. P. Seager, W. Lin, C. Grebogi, A. Hastings, and Y.-C. Lai, *Predicting Tipping Points in Mutualistic Networks through Dimension Reduction*, Proc. Natl. Acad. Sci. U.S.A. **115**, E639 (2018).

[11] H. R. Wilson and J. D. Cowan, *Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons*, Biophys. J. **12**, 1 (1972).

[12] This is actually a modified version of the original Cowan-Wilson model, which describes output activities of the specific form of Eq. (1). In Eq. (2a), x_i must be interpreted as the input activity to neuron i.

[13] J. Grilli, M. Adorisio, S. Suweis, G. Barabás, J. R. Banavar, S. Allesina, and A. Maritan, *Feasibility and Coexistence of Large Ecological Communities*, Nat. Commun. **8**, 14389 (2017).

[14] J. N. Holland, D. L. DeAngelis, and J. L. Bronstein, *Population Dynamics and Mutualism: Functional Responses of Benefits and Costs*, Am. Nat. **159**, 231 (2002).

[15] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, Boca Raton, FL, USA, 2006).

[16] R. Pastor-Satorras and A. Vespignani, *Epidemic Dynamics and Endemic States in Complex Networks*, Phys. Rev. E **63**, 066117 (2001).

[17] C. Eliasmith, T. C. Stewart, X. Choo, T. Bekolay, T. DeWolf, Y. Tang, and D. Rasmussen, *A Large-Scale Model of the Functioning Brain*, Science **338**, 1202 (2012).

[18] E. Laurence, J.-G. Young, S. Melnik, and L. J. Dubé, *Exact Analytical Solution of Irreversible Binary Dynamics on Networks*, Phys. Rev. E **97**, 032302 (2018).

[19] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University, Cambridge, England, 1990).

[20] A.-L. Barabási and R. Albert, *Emergence of Scaling in Random Networks*, Science **286**, 509 (1999).

[21] The average number of edges per node $\langle \boldsymbol{k}_B^{\mathrm{in}} \rangle = N^{-1} \langle \mathbf{1}^T \boldsymbol{A} \mathbf{1} \rangle$ is computed from the binary adjacency matrix $\boldsymbol{A}$ and must be distinguished from the average degree $\langle \boldsymbol{k}^{\mathrm{in}} \rangle = N^{-1} \langle \mathbf{1}^T \boldsymbol{W} \mathbf{1} \rangle$ computed from the weighted adjacency matrix $\boldsymbol{W}$. In the former equations, $\langle \mathbf{1}^T \boldsymbol{A} \mathbf{1} \rangle$ either denotes the expected value of $\mathbf{1}^T \boldsymbol{A} \mathbf{1} = \sum_{i,j} A_{ij}$ over a random network ensemble (in the case of random networks such as ER networks) or simply $\sum_{i,j} A_{ij}$ (in the case of a particular network), and similarly for $\langle \mathbf{1}^T \boldsymbol{W} \mathbf{1} \rangle$.

[22] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Random Graphs with Arbitrary Degree Distributions and Their Applications*, Phys. Rev. E **64**, 026118 (2001).

[23] M. E. J. Newman, *Networks: An Introduction* (Oxford University, New York, NY, USA, 2010).

[24] F. Chung, L. Lu, and V. Vu, *Spectra of Random Graphs with Given Expected Degrees*, Proc. Natl. Acad. Sci. U.S.A. **100**, 6313 (2003).

[25] C. Castellano and R. Pastor-Satorras, *Relating Topological Determinants of Complex Networks to Their Spectral Properties: Structural and Dynamical Effects*, Phys. Rev. X **7**, 041024 (2017).

[26] U. Bastolla, M. A. Fortuna, A. Pascual-García, A. Ferrera, B. Luque, and J. Bascompte, *The Architecture of Mutualistic Networks Minimizes Competition and Increases Biodiversity*, Nature (London) **458**, 1018 (2009).

[27] M. E. J. Newman, *Scientific Collaboration Networks. I. Network Construction and Fundamental Results*, Phys. Rev. E **64**, 016131 (2001).

[28] D. J. Watts and S. H. Strogatz, *Collective Dynamics of Small-World Networks*, Nature (London) **393**, 440 (1998).

[29] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Application*, Vol. 87 (Academic Press, New York, 1980).

[30] P. W. Holland, K. B. Laskey, and S. Leinhardt, *Stochastic Blockmodels: First Steps*, Soc. Networks **5**, 109 (1983).

[31] Note that if $(p_{11} - p_{22})^2 \ll 4 p_{12} p_{21}$, then the problem of misrepresentation is absent, and it reduces to a single one-dimensional representation.

[32] A. Majdandzic, L. A. Braunstein, C. Curme, I. Vodenska, S. Levy-Carciente, H. E. Stanley, and S. Havlin, *Multiple Tipping Points and Optimal Repairing in Interacting Networks*, Nat. Commun. **7**, 10850 (2016).

[33] P. L. Krapivsky, S. Redner, and F. Leyvraz, *Connectivity of Growing Random Networks*, Phys. Rev. Lett. **85**, 4629 (2000).

[34] J.-G. Young, L. Hébert-Dufresne, E. Laurence, C. Murphy, G. St-Onge, and P. Desrosiers, *Network Archaeology: Phase Transition in the Recoverability of Network History*, arXiv:1803.09191.

[35] G. Weisbuch, *Complex Systems Dynamics* (CRC Press, Boca Raton, FL, USA, 2018).

[36] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Epidemic Processes in Complex Networks*, Rev. Mod. Phys. **87**, 925 (2015).

[37] G. St-Onge, J.-G. Young, E. Laurence, C. Murphy, and L. J. Dubé, *Phase transition of the Susceptible-Infected-Susceptible Dynamics on Time-Varying Configuration Model Networks*, Phys. Rev. E **97**, 022305 (2018).

[38] M. Boguná and R. Pastor-Satorras, *Epidemic Spreading in Correlated Complex Networks*, Phys. Rev. E **66**, 047104 (2002).

[39] C. Tu, J. Grilli, F. Schuessler, and S. Suweis, *Collapse of Resilience Patterns in Generalized Lotka-Volterra Dynamics and Beyond*, Phys. Rev. E **95**, 062307 (2017).