

# Dimension reduction of dynamical systems on networks with leading and non-leading eigenvectors of adjacency matrices

Naoki Masuda<sup>1,2\*</sup> and Prosenjit Kundu<sup>1</sup>

<sup>1</sup>*Department of Mathematics, State University of New York at Buffalo, NY 14260-2900, USA and*

<sup>2</sup>*Computational and Data-Enabled Science and Engineering Program,  
State University of New York at Buffalo, Buffalo, NY 14260-5030, USA*

(Dated: June 16, 2022)

Dimension reduction techniques for dynamical systems on networks are considered to promote our understanding of the original high-dimensional dynamics. One strategy of dimension reduction is to derive a low-dimensional dynamical system whose behavior approximates the observables of the original dynamical system that are weighted linear summations of the state variables at the different nodes. Recently proposed methods use the leading eigenvector of the adjacency matrix of the network as the mixture weights to obtain such observables. In the present study, we explore performances of this type of one-dimensional reductions of dynamical systems on networks when we use non-leading eigenvectors of the adjacency matrix as the mixture weights. Our theory predicts that non-leading eigenvectors can be more efficient than the leading eigenvector and enables us to select the eigenvector minimizing the error. We numerically verify that the optimal non-leading eigenvector outperforms the leading eigenvector for some dynamical systems and networks. We also argue that, despite our theory, it is practically better to use the leading eigenvector as the mixture weights to avoid misplacing the bifurcation point too distantly and to be resistant against dynamical noise.

## I. INTRODUCTION

A variety of complex systems in the real world can be described by dynamical systems on networks [1–4]. This seems to be the case in particular when systems of question are composed of dynamical elements that are similar to each other except for the connectivity and some easily parameterizable heterogeneity across the individual elements. Examples include coupled oscillators on networks [5] including models of functioning of power grids [6], predator-prey, mutualistic, and other dynamics impacting community stability in ecological networks [7, 8], gene regulatory networks [9], and epidemic processes and ecological population dynamics considered on networks

---

\* naokimas@buffalo.edu

of habitat patches (called metapopulation models) [10, 11]. Analyses of dynamical systems on networks have clarified various collective phenomena on networks such as phase transitions and synchronous oscillations.

Dynamical systems on networks are necessarily high-dimensional because each node is assigned with one or more dynamical variables and those variables interact via edges of the given network. Even if the dynamics at each node is one-dimensional, the entire dynamical system on the network is  $N$ -dimensional, where  $N$  is the number of nodes in the network. Therefore, similar to various dimension reduction techniques for high-dimensional data [12–14], one may be tempted to map a dynamical system on networks into one in a low-dimensional space without losing much information. Then, by deploying analytical and numerical techniques suited to low-dimensional dynamical systems, we may be better able to understand the original high-dimensional network dynamics. Gao, Barzel, and Barabási developed a heterogeneous mean-field theory, assuming uncorrelated networks with general degree (i.e., number of edges that a node in the given network has) distributions, to reduce a class of dynamical systems on networks to one-dimensional dynamical systems [15]. Their method approximates a one-dimensional projection of the original high-dimensional dynamics on networks with a reasonably good accuracy in various cases. See Refs. [16, 17] for validation studies of this approach and Ref. [18] for an extension. Furthermore, to deal with general network structure, Laurence *et al.* developed a systematic method, which we call the spectral method, to use the eigenvalues and eigenvectors of the adjacency matrix of the network to reduce the same class of dynamics on networks into dynamics of small dimensions, such as one or two [19]. See Ref. [20] for a further advancement of the spectral method. In the case of one-dimensional reduction, which we focus on in the present study, the spectral method uses an observable that is a particular linear combination of the state variables on all nodes, denoted by  $R = \sum_{i=1}^N a_i x_i$ , where  $x_i$  is the dynamical state of the  $i$ th node, and  $a_i$  is the mixing weight. Then, one writes down a closed dynamical equation in terms of  $R$ . On a theoretical basis, they proposed to use the  $i$ th element of the leading eigenvector (i.e., the eigenvector associated with the largest eigenvalue) of the adjacency matrix as  $a_i$  [19].

The leading eigenvector of the adjacency matrix is a key descriptor of contagious processes on networks because the adjacency matrix tells us who can directly infect whom. In fact, the  $i$ th element of the leading eigenvector gives the likelihood that the  $i$ th node is infectious when the infection rate of an epidemic process model such as the susceptible-infectious-susceptible (SIS) model is poised near the epidemic threshold [21]. For example, scale-free networks (i.e., networks with power-law degree distributions) show eigenvector localization such that the leading eigenvector

has only a small fraction of considerably positive elements, which are at the largest-degree nodes and correspond to the presence of infection at these nodes [21–24]. In contrast, the other elements of the eigenvector are close to 0, corresponding to the scarcity of infection at small-degree nodes. These and other results lend support for the spectral method [19] and its extension called the dynamics approximate reduction technique (DART) [20], which use the leading eigenvectors of the adjacency matrix as mixing weights.

In the present study, we develop a theory to argue that it is often better to use a non-leading eigenvector of the adjacency matrix for the spectral method to realize a better accuracy at reducing the original  $N$ -dimensional dynamics into a one-dimensional dynamics. We derive optimization criteria under the assumption that  $\{x_1, \dots, x_N\}$  is not too heterogeneous, which numerically holds true for various dynamical systems on networks [17]. Our theoretical derivation suggests that the leading eigenvector does not necessarily yield the smallest error of the dimension reduction by the spectral method. For various networks, a non-leading eigenvector, which implies that we linearly combine  $x_i$  into one observable with some negative weights  $a_i$ , realizes a smaller error than the spectral method with the leading eigenvector. We verify our theory by numerical simulations of three dynamical systems on different networks. Finally, we argue that, despite our theory, the spectral method using the leading eigenvector as the mixing weights is better than with the non-leading eigenvector because of two factors that our theory does not address: precision in locating the bifurcation point and the robustness against dynamical noise. Our code for computing the optimal eigenvectors and reproducing the results in this article is available at <https://github.com/naokimas/nonleading-spectral>.

## II. SPECTRAL METHOD

Throughout the present study, we consider the following class of dynamical systems on networks [15, 19, 25]:

$$\frac{dx_i}{dt} = F(x_i) + \sum_{j=1}^N w_{ij} G(x_i, x_j), \quad (1)$$

where  $t$  is the time,  $x_i$  is the one-dimensional dynamical state of the  $i$ th node (with  $i \in \{1, \dots, N\}$ ),  $F(x)$  represents the intrinsic dynamics of the node,  $G(x_i, x_j)$  represents the influence of  $x_j$  on  $x_i$ , and  $w_{ij}$  is the strength of the influence of node  $j$  on node  $i$ , corresponding to the weighted adjacency matrix of the given network. We assume that the network is connected. We also assume that the network does not have self-loops, i.e.,  $w_{ii} = 0$  for  $i \in \{1, \dots, N\}$ . However, if all nodes have a self-

loop of the same edge weight (i.e.,  $w_{11} = \dots = w_{NN}$ ), one can include the effect of such self-loops into  $F(x_i)$  by replacing the original  $F(x_i)$  by  $F(x_i) + w_{ii}G(x_i, x_i)$ .

We describe the spectral method [19] in this section. With this method, one reduces the  $N$ -dimensional dynamical system given by Eq. (1) to an  $n$ -dimensional system, where  $n \ll N$ , by deriving an approximate  $n$ -dimensional dynamical system in terms of observables each of which is a linear combination of  $\{x_1, \dots, x_N\}$ . We focus on the case of  $n = 1$  in this paper. We consider an observable, which we denote by  $R$ , given by

$$R = \sum_{i=1}^N a_i x_i, \quad (2)$$

where  $\{a_1, \dots, a_N\}$  is normalized such that  $\sum_{i=1}^N a_i = 1$ .

By combining Eqs. (1) and (2), one obtains

$$\frac{dR}{dt} = \sum_{i=1}^N a_i \left[ F(x_i) + \sum_{j=1}^N w_{ij} G(x_i, x_j) \right]. \quad (3)$$

By Taylor expanding  $F(x_i)$  around  $x_i = R$  to the first order, we obtain

$$F(x_i) = F(R) + (x_i - R)F'(R) + O((x_i - R)^2). \quad (4)$$

Similarly, we expand  $G(x_i, x_j)$  around  $x_i = \beta R$  and  $x_j = \gamma R$ , where  $\beta$  and  $\gamma$  are constants to be determined, to obtain

$$G(x_i, x_j) = G(\beta R, \gamma R) + (x_i - \beta R)G_1(\beta R, \gamma R) + (x_j - \gamma R)G_2(\beta R, \gamma R) + O((x_i - \beta R)^2) + O((x_j - \gamma R)^2), \quad (5)$$

where  $G_1$  and  $G_2$  are the partial derivatives of  $G$  with respect to the first and second argument, respectively. By substituting Eqs. (4) and (5) into Eq. (3), one obtains

$$\begin{aligned} \frac{dR}{dt} = & F(R) + \alpha G(\beta R, \gamma R) + G_1(\beta R, \gamma R) \sum_{i,j=1}^N a_i w_{ij} (x_i - \beta R) + G_2(\beta R, \gamma R) \sum_{i,j=1}^N a_i w_{ij} (x_j - \gamma R) \\ & + O((x - R)^2) + O((x - \beta R)^2) + O((x - \gamma R)^2), \end{aligned} \quad (6)$$

where

$$\alpha = \sum_{i,j=1}^N a_i w_{ij}, \quad (7)$$

and  $O((x - R)^2)$  is a short-hand notation for  $\sum_{i=1}^N O((x_i - R)^2)$  and similar for  $O((x - \beta R)^2)$  and  $O((x - \gamma R)^2)$ . By requiring that the first-order terms in Eq. (6), i.e., those containing  $G_1$

and  $G_2$ , disappear for any  $\{x_1, \dots, x_N\}$ , one obtains

$$\alpha\beta R = \sum_{i,j=1}^N a_i w_{ij} x_i = \mathbf{x}^\top K \mathbf{a}, \quad (8)$$

$$\alpha\gamma R = \sum_{i,j=1}^N a_i w_{ij} x_j = \mathbf{x}^\top W^\top \mathbf{a}, \quad (9)$$

where  $\mathbf{x} = (x_1, \dots, x_N)^\top$ ,  $\mathbf{a} = (a_1, \dots, a_N)^\top$ ,  $K$  is the  $N \times N$  diagonal matrix whose  $i$ th diagonal entry is equal to the weighted in-degree of the  $i$ th node, i.e.,  $k_i^{\text{in}} \equiv \sum_{j=1}^N w_{ij}$ ,  $W = (w_{ij})$  is the  $N \times N$  adjacency matrix, and  $^\top$  represents the transposition. By substituting Eq. (2) in Eqs. (8) and (9), one obtains

$$\mathbf{x}^\top (\alpha\beta \mathbf{a} - K \mathbf{a}) = 0, \quad (10)$$

$$\mathbf{x}^\top (\alpha\gamma \mathbf{a} - W^\top \mathbf{a}) = 0. \quad (11)$$

Because Eqs. (10) and (11) ideally hold true for any  $\mathbf{x}$ , one obtains

$$K \mathbf{a} = \alpha\beta \mathbf{a}, \quad (12)$$

$$W^\top \mathbf{a} = \alpha\gamma \mathbf{a}. \quad (13)$$

Equations (12) and (13) indicate that  $\mathbf{a}$  is a right eigenvector of both  $K$  and  $W^\top$ . However,  $K$  and  $W^\top$  do not share the eigenspace in general. In particular, the eigenvectors of  $K$  are the standard unit vectors because  $K$  is a diagonal matrix. The standard unit vector  $\mathbf{a} = (1, 0, \dots, 0)^\top$ , for example, satisfies Eq. (12), but it implies that  $R = x_1$  so that we only observe the first node. Note that  $\mathbf{a} = (1, 0, \dots, 0)^\top$  satisfies Eq. (13) if and only if  $k_1^{\text{in}} = 0$  such that the first node is not influenced by any other node.

Laurence *et al.* proposed to set  $\mathbf{a}$  to be a right eigenvector of  $W^\top$ . Because  $\mathbf{a}$  is normalized such that  $\mathbf{1}^\top \mathbf{a} = 1$ , where  $\mathbf{1} = (1, \dots, 1)^\top$ , we left-multiply  $\mathbf{1}^\top$  by Eq. (13) and use  $\mathbf{1}^\top W^\top = \mathbf{k}^{\text{in}}$ , where  $\mathbf{k}^{\text{in}} = (k_1^{\text{in}}, \dots, k_N^{\text{in}})^\top$ , and  $\alpha = \mathbf{a}^\top \mathbf{k}^{\text{in}}$ , which originates from Eq. (7), to obtain  $\alpha = \alpha\gamma$ , i.e.,  $\gamma = 1$ . Substitution of  $\gamma = 1$  in Eq. (13) implies that  $\alpha$  is the eigenvalue of  $W^\top$  associated with eigenvector  $\mathbf{a}$ . One can normalize  $\mathbf{a}$  as  $\mathbf{1}^\top \mathbf{a} = 1$  unless the associated eigenvalue is 0. Finally, because Eq. (12) no longer holds true in general, one uses the  $\beta$  value that minimizes the approximation error, i.e.,

$$\beta^* = \operatorname{argmin}_\beta \|K \mathbf{a} - \alpha\beta \mathbf{a}\|^2 = \frac{\mathbf{b}^\top \mathbf{k}^{\text{in}}}{\mathbf{a}^\top \mathbf{k}^{\text{in}}}, \quad (14)$$

where  $\mathbf{b} = (b_1, \dots, b_N)^\top$ , and  $b_i = a_i^2 / \sum_{\ell=1}^N a_\ell^2$  with  $i \in \{1, \dots, N\}$ .

By neglecting the second and higher order terms in Eq. (6), one obtains the spectral reduction given by

$$\frac{dR}{dt} = F(R) + \alpha G(\beta^* R, R). \quad (15)$$

### III. MODIFIED SPECTRAL METHOD

#### A. Eigenvector minimizing the approximation error

With the optimal  $\beta$  value given by Eq. (14), we obtain

$$e_1 \equiv \|K\mathbf{a} - \alpha\beta^*\mathbf{a}\|^2 = \frac{\sum_{i=1}^N \sum_{j=1}^{i-1} a_i^2 a_j^2 (k_i^{\text{in}} - k_j^{\text{in}})^2}{\sum_{\ell=1}^N a_\ell^2}. \quad (16)$$

We have two remarks on Eq. (16). First, although the use of the leading eigenvector as  $\mathbf{a}$  has been recommended for the spectral method [19] and the DART [20],  $e_1$  may be smaller for other right eigenvectors of  $W^\top$ . Second, Eq. (16) implies that  $e_1 = 0$  if all the nodes have the same in-degree, regardless of which right eigenvector of  $W^\top$  we use as  $\mathbf{a}$ . However, numerical simulations indicate that the spectral method is not exact even for random regular graphs [17]. The approximation error may be due to the higher-order terms in the Taylor expansion in Eq. (6) or the fact that we have expanded Eq. (3) around both  $x_i = R$  and  $x_i = \beta^*R$ . (Although we also expanded  $x_i$  around  $\gamma R$ , we found that  $\gamma = 1$ .) In fact, our numerical test suggests that  $\beta^*$  is often far from 1 both for the leading eigenvector and the eigenvector minimizing  $\epsilon_1$ , as we show for scale-free networks in Fig. 1(a) and Fig. 1(b), respectively.

For these reasons, we propose to expand Eq. (3) only around  $x_i = R$  as follows. Let us set  $x_i = R + \Delta x_i$  (with  $i \in \{1, \dots, N\}$ ), where  $|\Delta x_i| \ll 1$ . Then, we obtain

$$\begin{aligned} & \sum_{i,j=1}^N a_i w_{ij} G(x_i, x_j) \\ &= \sum_{i,j=1}^N a_i w_{ij} G(R + \Delta x_i, R + \Delta x_j) \\ &= \sum_{i,j=1}^N a_i w_{ij} [G(R, R) + G_1(R, R)\Delta x_i + G_2(R, R)\Delta x_j] + O((\Delta x)^2) \\ &= \alpha G(R, R) + G_1(R, R) \sum_{i,j=1}^N a_i w_{ij} (x_i - R) + G_2(R, R) \sum_{i,j=1}^N a_i w_{ij} (x_j - R) + O((\Delta x)^2) \\ &= \alpha G(R, R) + G_1(R, R) \left( \sum_{i,j=1}^N a_i w_{ij} x_i - \alpha R \right) + G_2(R, R) \left( \sum_{i,j=1}^N a_i w_{ij} x_j - \alpha R \right) + O((\Delta x)^2). \end{aligned} \quad (17)$$

By imposing that the first-order terms on the right-hand side of Eq. (17) disappear, we obtain

$$\alpha R = \mathbf{x}^\top K \mathbf{a}, \quad (18)$$

$$\alpha R = \mathbf{x}^\top W^\top \mathbf{a}. \quad (19)$$

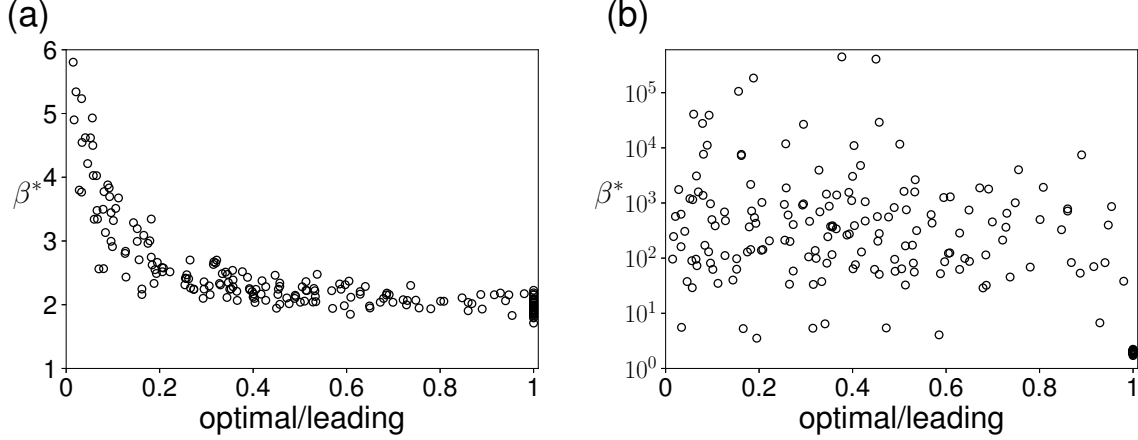


FIG. 1. The  $\beta^*$  value for 200 scale-free networks (i.e., networks with power-law degree distributions) with  $N = 1000$  nodes. Each circle represents a network. The horizontal axis represents the  $\epsilon_1$  value for the minimizer of  $\epsilon_1$  divided by the  $\epsilon_1$  value for the leading eigenvector. A small value on the horizontal axis implies that the minimizer of  $\epsilon_1$  is efficient at reducing  $\epsilon_1$  relative to the leading eigenvector. The value on the horizontal axis being equal to 1 implies that the leading eigenvector minimizes  $\epsilon_1$ . (a)  $\beta^*$  for the leading eigenvector. (b)  $\beta^*$  for the eigenvector minimizing  $\epsilon_1$ . We use the configuration model to generate the 200 scale-free networks with the power-law exponent of the degree distribution equal to  $\tilde{\gamma} = 3.5$ . See Section IV A for details of the degree distribution used. Note that one can rewrite Eq. (14) as  $\beta^* = \mathbf{b}^\top \mathbf{k}^{\text{in}} / \alpha$ . Therefore, if the minimizer of  $\epsilon_1$  is associated with a small eigenvalue,  $\alpha$ , then  $\beta^*$  tends to be large. This explains why the  $\beta^*$  value tends to be much larger in (b) than (a).

By substituting  $R = \mathbf{x}^\top \mathbf{a}$  in Eqs. (18) and (19) and imposing that Eqs. (18) and (19) hold true for any  $\mathbf{x}$ , we obtain

$$K\mathbf{a} = \alpha\mathbf{a}, \quad (20)$$

$$W^\top \mathbf{a} = \alpha\mathbf{a}. \quad (21)$$

For the same reason as that for the original spectral method, Eqs. (20) and (21) do not simultaneously hold true in general. Therefore, by following the recommendation for DART [20], we require Eq. (21) to hold exactly and let Eq. (20) be satisfied only approximately. The opposite case, i.e., to impose Eq. (20) strictly and Eq. (21) only approximately, implies that we only observe a single node as a representative of the entire system (see Appendix A for the derivation). Equation (21) implies that, as in the case of the original spectral method,  $\mathbf{a}$  must be a right eigenvector of  $W^\top$  and that  $\alpha$  is the associated eigenvalue. Therefore, we select the eigenvalue  $\alpha$  and the associated

eigenvector  $\mathbf{a}$  that minimize the the approximation error for Eq. (20) defined by

$$\epsilon_2 \equiv \|K\mathbf{a} - \alpha\mathbf{a}\|^2 = \sum_{i=1}^N (k_i^{\text{in}} - \alpha)^2 a_i^2. \quad (22)$$

We remind that  $\mathbf{a}$  is normalized such that  $\sum_{i=1}^N a_i = 1$ . The modified one-dimension reduction reads

$$\frac{dR}{dt} = F(R) + \alpha G(R, R). \quad (23)$$

A few remarks are in order. First, as in the case of the original spectral method, the leading eigenvector of  $W^\top$  may not minimize  $\epsilon_2$ . Second,  $\epsilon_2 \geq \epsilon_1$  holds true because our one-dimensional reduction scheme corresponds to  $\beta = 1$ , whereas the original spectral method also optimizes the  $\beta$  value. Nevertheless, which reduction method is more accurate than the other is a nontrivial question because the original spectral method uses the Taylor expansion of each  $x_i$  around two reference values (i.e.,  $R$  and  $\beta^*R$ ). Third, the Perron-Frobenius theorem guarantees that  $a_i > 0$ ,  $\forall i \in \{1, \dots, N\}$  when  $\mathbf{a}$  is the right leading eigenvector of  $W^\top$ . Therefore, the observable,  $R$ , when the spectral method uses the leading eigenvector, is a weighted average of all the dynamical variables,  $x_i$ , with positive weights. The Perron-Frobenius theorem also guarantees that all the entries of the left leading eigenvector of  $W^\top$ , which we denote by  $\mathbf{v} \in \mathbb{R}^{1 \times N}$ , are also positive. If  $W^\top$  is diagonalizable and the leading eigenvalue is not repeated, we obtain  $\mathbf{v}\mathbf{a} = 0$  for any right eigenvector  $\mathbf{a}$  of  $W^\top$  other than the leading one. Therefore, the sign of at least one  $a_i$  must be opposite to the sign of some other  $a_i$ . This fact implies that  $R = \mathbf{a}^\top \mathbf{x}$  is a weighted average of  $\{x_1, \dots, x_N\}$  including negative weights.

## B. Case of regular graphs

In the case of the complete graph,  $K$  is the diagonal matrix whose all diagonal entries are equal to  $N - 1$ . Matrix  $W$  is given by  $W = (w_{ij})$ , where  $w_{ij} = 1 - \delta_{ij}$ , and  $\delta_{ij}$  is the Kronecker delta. In this case,  $W$  has the leading eigenvalue  $\alpha = N - 1$  with the associated eigenvector  $\mathbf{a} = (1, \dots, 1)^\top$  and the  $(N - 1)$ -fold eigenvalue  $\alpha = -1$  whose associated eigenvectors can be chosen as  $\mathbf{a} = (1, -1, 0, \dots, 0)^\top, (1, 0, -1, 0, \dots, 0)^\top, (1, 0, 0, -1, 0, \dots, 0)^\top, \dots, (1, 0, \dots, 0, -1)^\top$ . We find that the combination of  $\alpha = N - 1$  and  $\mathbf{a} = (1, \dots, 1)^\top$  satisfies Eq. (20), whereas the combination of  $\alpha = -1$  and any of its eigenvector does not. Therefore, only the leading eigenvalue and eigenvector realize  $\epsilon_2 = 0$ . Furthermore, due to the symmetry, all the  $x_i$  values are same, such that the  $O((\Delta x)^2)$  in Eq. (17) disappears. Therefore, the spectral method is exact. Because the



complete graph is a regular graph (i.e., a network in which all the nodes have the same degree), we also obtain  $\epsilon_1 = 0$  for any eigenvector.

Similar results hold true for regular graphs in general. Specifically, for any undirected regular graphs with degree  $k$ , the leading eigenvalue of  $W$  is equal to  $k$ , and the associated eigenvector is  $\mathbf{a} = (1, \dots, 1)^\top$ . We find that this pair of eigenvalue and eigenvector, which is known to be of multiplicity 1 [26], satisfies Eq. (20). No other pair of eigenvalue and the associated eigenvector of  $W$  does not satisfy Eq. (20) because Eq. (20) implies that the eigenvalue needs to be equal to  $k$ . With the leading eigenvalue and eigenvector of  $W^\top$ , if the network is vertex-transitive [26] (e.g., square lattice with periodic boundary conditions),  $O((\Delta x)^2)$  in Eq. (17) disappears due to the symmetry (i.e.,  $x_i = x_j$  for all  $i, j \in \{1, \dots, N\}$ ) such that the spectral method is exact. Otherwise, the  $O((\Delta x)^2)$  term may cause the discrepancy between the spectral method and the numerical results, as is the case for random regular graphs [17].

### C. Case of degree-heterogeneous random graphs

Consider undirected configuration models, i.e., uniformly random undirected networks with a given degree sequence. In the limit of  $N \rightarrow \infty$ , the normalized leading eigenvector of the adjacency matrix is approximated by  $a_i = k_i / \sum_{\ell=1}^N k_\ell$  [23], and the leading eigenvalue is approximated by  $\alpha = \langle k^2 \rangle / \langle k \rangle$ , where  $\langle \cdot \rangle$  represents the average over the  $N$  nodes [27]. By substituting these relationships in Eq. (22), we obtain

$$\epsilon_2 \approx \sum_{i=1}^N \left( k_i - \frac{\langle k^2 \rangle}{\langle k \rangle} \right)^2 \left( \frac{k_i}{\sum_{\ell=1}^N k_\ell} \right)^2 = \frac{\langle k \rangle^2 \langle k^4 \rangle - 2 \langle k \rangle \langle k^2 \rangle \langle k^3 \rangle + \langle k^2 \rangle^3}{N \langle k \rangle^4}, \quad (24)$$

where  $\approx$  represents “approximately equal to”. If we instead use a non-leading eigenvector whose associated eigenvalue is of  $O(1)$  and assume that  $a_i$  and  $k_i$  are uncorrelated and that  $a_i = O(N^{-1})$ , we obtain

$$\epsilon_2 = \sum_{i=1}^N (k_i - \alpha)^2 \cdot O(N^{-2}) = O(\langle k^2 \rangle N^{-1}). \quad (25)$$

In degree-heterogeneous networks, the leading term in Eq. (24) is  $\langle k^4 \rangle \langle k \rangle^{-2} N^{-1}$ , which is expected to be much larger than  $O(\langle k^2 \rangle N^{-1})$  in Eq. (25). Therefore, if  $a_i = O(N^{-1})$  for various non-leading eigenvectors, we expect that the spectral method is likely to be optimized by a non-leading eigenvector in degree-heterogeneous networks.

#### IV. NUMERICAL RESULTS

In this section, we numerically compare the spectral method with the leading eigenvector and that with the error-minimizing eigenvectors.

##### A. Estimated performance of non-leading eigenvectors for various networks

We first examine the performance of the optimal non-leading eigenvectors in reducing  $\epsilon_1$  and  $\epsilon_2$  in comparison with the leading eigenvector for several networks.

First, we consider networks with  $N$  nodes by the Erdős-Rényi (ER) random graph in which each node pair is adjacent with probability  $\langle k \rangle / (N - 1)$ ; we remind that  $\langle k \rangle$  is the mean degree. We generate 200 networks from the ER random graph and use the largest connected component of each network for each pair of  $N \in \{100, 1000\}$  and  $\langle k \rangle \in \{4, 10\}$ . For the other network models that we consider in the following text, we also generate 200 networks and use the largest connected component. The largest connected component of the networks generated by the ER random graph model contains at least 90% of the nodes. We then compute the minimizer of  $\epsilon_1$  and that of  $\epsilon_2$  for each network. We find that the leading eigenvector is the minimizer of  $\epsilon_1$  and  $\epsilon_2$  in all networks. Therefore, we should use the spectral method with the leading eigenvector [19, 20] for the ER random graph.

The results are similar for the Watts-Strogatz model of small-world networks [28], whereas the leading eigenvector does not minimize  $\epsilon_1$  or  $\epsilon_2$  for a small fraction of network instances. Specifically, we set the rewiring probability to 0.1 and generate networks for each pair of  $N \in \{100, 1000\}$  and  $\langle k \rangle \in \{4, 10\}$ . We have confirmed that the largest connected component of the network always contains all nodes. With  $\langle k \rangle = 10$ , the leading eigenvector minimizes both  $\epsilon_1$  and  $\epsilon_2$  for all networks. With  $\langle k \rangle = 4$ , the minimizer of each type of error (i.e.,  $\epsilon_1$  or  $\epsilon_2$ ) is not the leading eigenvector only for 1.5% and 6% of the networks with  $N = 100$  and  $N = 1000$  nodes, respectively.

Next, we generate networks with power-law degree distributions, which we refer to as scale-free networks, using the configuration model. To this end, we draw the degree of each node, denoted by  $k$ , by independently sampling  $k$  from a power-law distribution given by  $p(k) = \kappa(\tilde{\gamma} - 1) / [1 + \kappa(k - 1)]^{\tilde{\gamma}}$ , where  $k \geq 1$ ,  $\kappa > 0$ , and  $\tilde{\gamma} > 1$ . Because  $p(k)$  is a probability density function, we round the sampled  $k$  to the nearest integer. The mean of the probability density  $p(k)$  is  $1 + [\kappa(\tilde{\gamma} - 2)]^{-1}$ , we equate it to the mean degree,  $\langle k \rangle$ . With the given values of  $\langle k \rangle$  and  $\tilde{\gamma}$  values, we uniquely determine the  $\kappa$  value. We examine combinations of  $N \in \{100, 1000\}$ ,  $\langle k \rangle \in \{4, 10\}$ ,

and  $\tilde{\gamma} \in \{2.5, 3.5\}$ . The largest connected component contains at least 91.5% of the nodes except when  $N = 100$ ,  $\langle k \rangle = 4$ , and  $\tilde{\gamma} \in \{2.5, 3.5\}$ , in which case it contains at least 79% of the nodes (i.e., 79 nodes). We calculate the ratio of the  $\epsilon_1$  value for its minimizer to the  $\epsilon_1$  value for the leading eigenvector. By definition, this ratio ranges between 0 and 1. If the ratio is small, the minimizer of  $\epsilon_1$  is expected to be better at approximating a one-dimensional projection of the original  $N$ -dimensional dynamics. In contrast, the ratio value equal to 1 implies that the leading eigenvector minimizes  $\epsilon_1$ . We similarly measured the ratio of  $\epsilon_2$  for its minimizer to  $\epsilon_2$  for the leading eigenvector. For scale-free networks with  $\tilde{\gamma} = 2.5$ , we show the cumulative distribution of the ratio for  $\epsilon_1$  and  $\epsilon_2$  in Figs. 2(a) and 2(b), respectively. The cumulative distribution shows the fraction of the networks for which the ratio is larger than the value shown on the horizontal axis. Therefore, if the cumulative distribution is small for a range of the ratio value smaller than 1, then the minimizer of  $\epsilon_1$  or  $\epsilon_2$  is efficient relative to the leading eigenvector for a large proportion of network instances. Figures 2(a) and 2(b) indicate that there are many instances of networks for which the leading eigenvector does not minimize  $\epsilon_1$  or  $\epsilon_2$  and that the minimization of  $\epsilon_1$  or  $\epsilon_2$  reduces the error by a large fraction relative to the case of the leading eigenvector in many cases. The results are similar for  $\tilde{\gamma} = 3.5$  (see Figs. 2(c) and 2(d)). These figures indicate that the minimizer of the error tends to be different from the leading eigenvector and tends to reduce the error by a large amount when the network is large, sparse, or more heterogeneous in terms of the node's degree (i.e.,  $\tilde{\gamma} = 2.5$  as opposed to  $\tilde{\gamma} = 3.5$ ).

Lastly, we investigate the scale-free network model proposed by Holme and Kim, which produces a high clustering coefficient (i.e., many triangles) [29]. We set the number of edges that each new node has, denoted by  $m$ , to  $m = 2$  and  $m = 5$  to produce networks whose average degree is approximately equal to  $\langle k \rangle = 4$  and  $\langle k \rangle = 10$ , respectively. We initialize the network by a star graph having  $m + 1$  nodes. We set the probability of making a triangle for each added edge to 0.5. We show the ratio of the minimized error to the error for the leading eigenvector in Figs. 2(e) and 2(f) for  $\epsilon_1$  and  $\epsilon_2$ , respectively. The results are qualitatively the same as those for the scale-free networks generated by the configuration model. We also find that the leading eigenvector tends to minimize the error for more network instances (i.e., the cumulative distribution is equal to 1 for a wider range of the ratio value on the horizontal axis) in the case of the Holme-Kim model than the configuration model, except for  $(N, \langle k \rangle) = (1000, 4)$ .

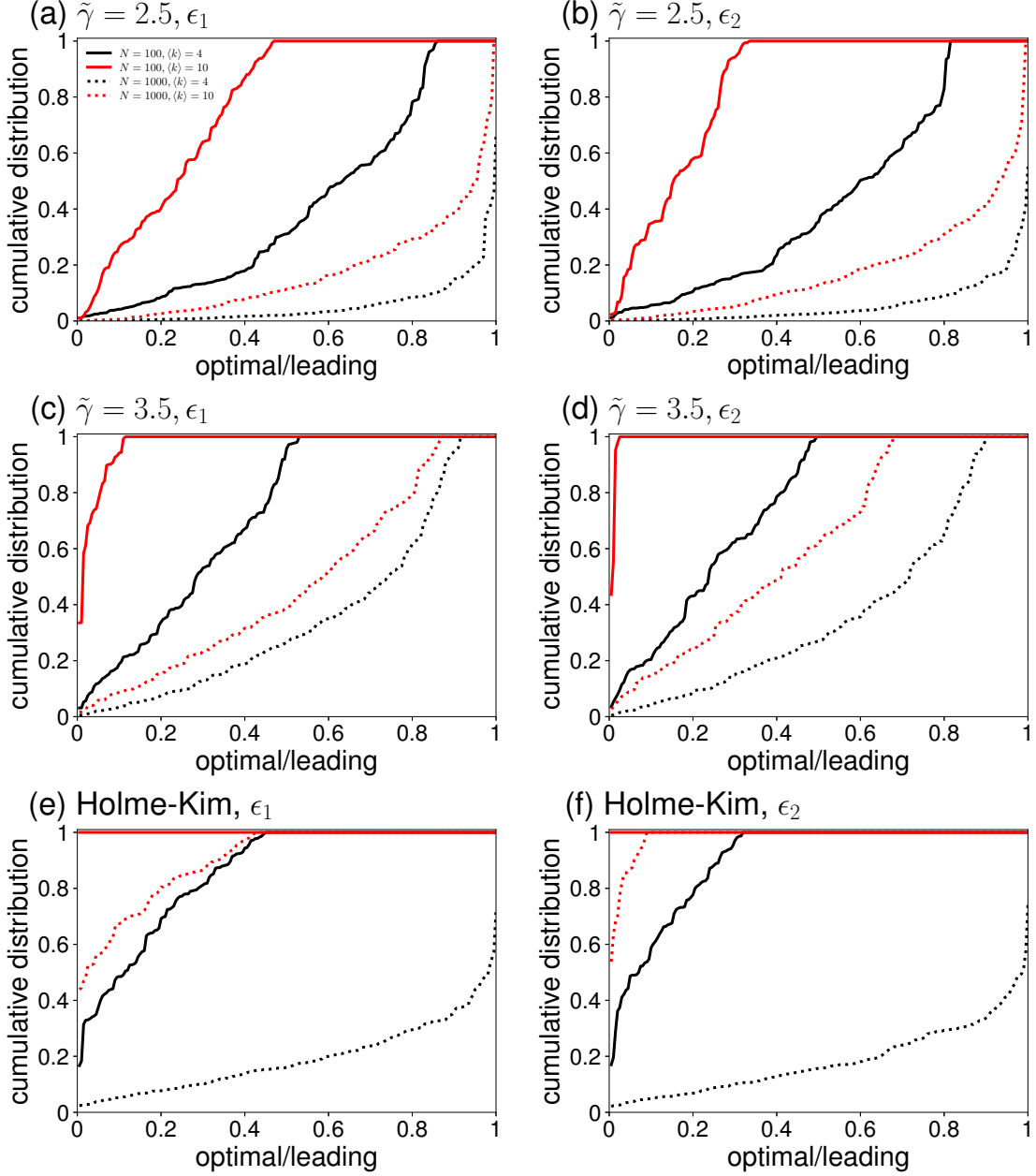


FIG. 2. Reduction in  $\epsilon_1$  and  $\epsilon_2$  by their minimizers relative to the case of the leading eigenvector of the adjacency matrix. We calculate the error (i.e., either  $\epsilon_1$  or  $\epsilon_2$ ) for its minimizer divided by the error for the leading eigenvector. The horizontal axis represents this ratio. The vertical axis represents the the cumulative distribution function of this ratio, i.e., the probability that the ratio is larger than the value specified on the horizontal axis. If the cumulative distribution is small across a wide range on the horizontal axis, then the minimizer of  $\epsilon_1$  or  $\epsilon_2$  tends to be more efficient than the leading eigenvector. (a) Error  $\epsilon_1$  for the scale-free networks with  $\tilde{\gamma} = 2.5$ . (b) Error  $\epsilon_2$  for the scale-free networks with  $\tilde{\gamma} = 2.5$ . (c) Error  $\epsilon_1$  for the scale-free networks with  $\tilde{\gamma} = 3.5$ . (d) Error  $\epsilon_2$  for the scale-free networks with  $\tilde{\gamma} = 3.5$ . (e) Error  $\epsilon_1$  for the Holme-Kim model. (f) Error  $\epsilon_2$  for the Holme-Kim model.

## B. Networks used in dynamical numerical simulations

In the following numerical simulations, we focus on 200 scale-free networks with  $N = 1000$ ,  $\langle k \rangle \in \{4, 10\}$ , and  $\tilde{\gamma} = 3.5$ . We also focus on the minimizer of  $\epsilon_2$ , not that of  $\epsilon_1$ , because we advocate the minimization of  $\epsilon_2$  in the present study (see Section III A). To identify the minimizer of  $\epsilon_2$  for each network, we discarded the eigenvectors associated with an eigenvalue smaller than  $10^{-6}$  including the case of a negative eigenvalue. This exclusion is because, for negative eigenvalues, Eq. (23) behaves qualitatively differently from the case of a positive eigenvalue, and, for positive but tiny eigenvalues, Eq. (23) shows a bifurcation of interest at an extremely large value of the bifurcation parameter. See Section V A for more discussion of this problem. We have found that 162 out of the 200 generated networks with  $\langle k \rangle = 4$  result in the minimal  $e_2$  values that are smaller than 70% of the error attained by the leading eigenvector and satisfy the condition that the eigenvalue is at least  $10^{-6}$ . We have found that 123 networks meet the same criterion when  $\langle k \rangle = 10$ . We only use these networks to calculate the statistics in the following analyses because, in those cases, we expect that the spectral method with a non-leading eigenvector may notably be better than that with the leading eigenvector.

In addition to the scale-free networks with  $\tilde{\gamma} = 3.5$ , we also use a coauthorship network among researchers that published articles on network science by 2006 [30]. The original data set contains 1589 nodes. We only use its largest connected component, which contains  $N = 379$  nodes and 914 edges. We regard this network as an unweighted network.

## C. SIS model

First, we consider the deterministic version of the SIS model, which is also called its individual-based approximation [31, 32], given by

$$\frac{dx_i}{dt} = -\mu x_i + \lambda \sum_{j=1}^N w_{ij}(1 - x_i)x_j, \quad (26)$$

where  $x_i$  represents the probability that the  $i$ th node is infectious at time  $t$ ,  $\lambda$  is the infection rate, and  $\mu$  is the recovery rate. Note that  $w_{ij} \in \{0, 1\}$ . By definition, each infectious node infects its susceptible neighbor independently at rate  $\lambda$ . An infectious node independently recovers at rate  $\mu$ . Because multiplying a common constant to  $\lambda$  and  $\mu$  only changes the timescale of the dynamics, we set  $\mu = 1$  without loss of generality. For each value of  $\lambda$ , we run a simulation with the initial condition  $x_i = 0.01, \forall i \in \{1, \dots, N\}$  until the equilibrium, denoted by  $\mathbf{x}^* = (x_1^*, \dots, x_N^*)$ , is sufficiently closely reached.

We show the bifurcation diagram for observable  $R = \mathbf{a}^\top \mathbf{x}^*$ , where  $\mathbf{a}$  is the leading eigenvector of the adjacency matrix, as a function of the infection rate,  $\lambda$ , for a scale-free network with  $N = 1000$  nodes,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$  in Fig. 3(a). Note that  $R$  is a weighted fraction of infectious nodes in the equilibrium. The solid line represents the results obtained from direct numerical simulations of the model. The dashed line represents the bifurcation diagram for the original spectral method, i.e., with the leading eigenvector as the weight vector  $\mathbf{a}$  and  $\beta = \beta^*$  (see Eq. (15)). We observe that the spectral method is accurate at locating the bifurcation point. However, the spectral method considerably underestimates  $R$  in the endemic phase (i.e., for  $\lambda$  values larger than the epidemic threshold). The dotted line in Fig. 3(a) shows the approximation of the same observable,  $R$ , by the spectral method in which we force  $\beta = 1$ , corresponding to Eq. (23), and continue to use the leading eigenvector as  $\mathbf{a}$ . The use of  $\beta = 1$  is discussed in Ref. [19]. We find that this one-dimensional reduction also locates the bifurcation point of the original dynamical system accurately and that it is better at approximating the  $R$  value in the endemic phase than with  $\beta = \beta^*$ . We compare the numerical results and the spectral method in which  $\mathbf{a}$  is the minimizer of  $\epsilon_2$  in Fig. 3(b). Note that the observable  $R$  is now different from that used in Fig. 3(a) because we have changed  $\mathbf{a}$ . Figure 3(b) indicates that the spectral method with the minimizer of  $\epsilon_2$  is worse than that with the leading eigenvector at accurately locating the bifurcation point. The minimizer of  $\epsilon_2$  is better at approximating  $R$  than the leading eigenvector combined with  $\beta = \beta^*$  but worse than the leading eigenvector combined with  $\beta = 1$ . The reason why the spectral method with the combination of the leading eigenvector as  $\mathbf{a}$  and  $\beta = 1$  does not minimize  $\epsilon_2$  but works better than the minimizer of  $\epsilon_2$  is unclear. It may be because higher-order terms in terms of  $\Delta x_i$  in our theory are nonnegligible or because the range of  $D$  values we have explored is not sufficiently far from the epidemic threshold estimated by the spectral method using the minimizer of  $\epsilon_2$ .

We compare the error among the three methods in Fig. 3(c). At each  $\lambda$  value, we defined the error as the absolute value of the difference between the  $R$  obtained by the direct numerical simulation and that obtained from the one-dimensional reduction, which we divided by the  $R$  value at the largest  $\lambda$  value, i.e.,  $\lambda = 4$ . We normalized the error in this manner because the true value of the observable  $R$  depends on  $\mathbf{a}$ . We use the absolute error instead of the relative error because  $R$  is close to or equal to 0 when  $\lambda$  is small. The error bars represent the mean and standard deviation. The figure confirms that the approximation error is the smallest in the case of the spectral method with the combination of the leading eigenvector and  $\beta = 1$ , the second smallest in the case of the minimizer of  $\epsilon_2$ , and the largest in the case of the combination of the leading eigenvector and  $\beta = \beta^*$ , when  $\lambda$  is sufficiently larger than the epidemic threshold for the

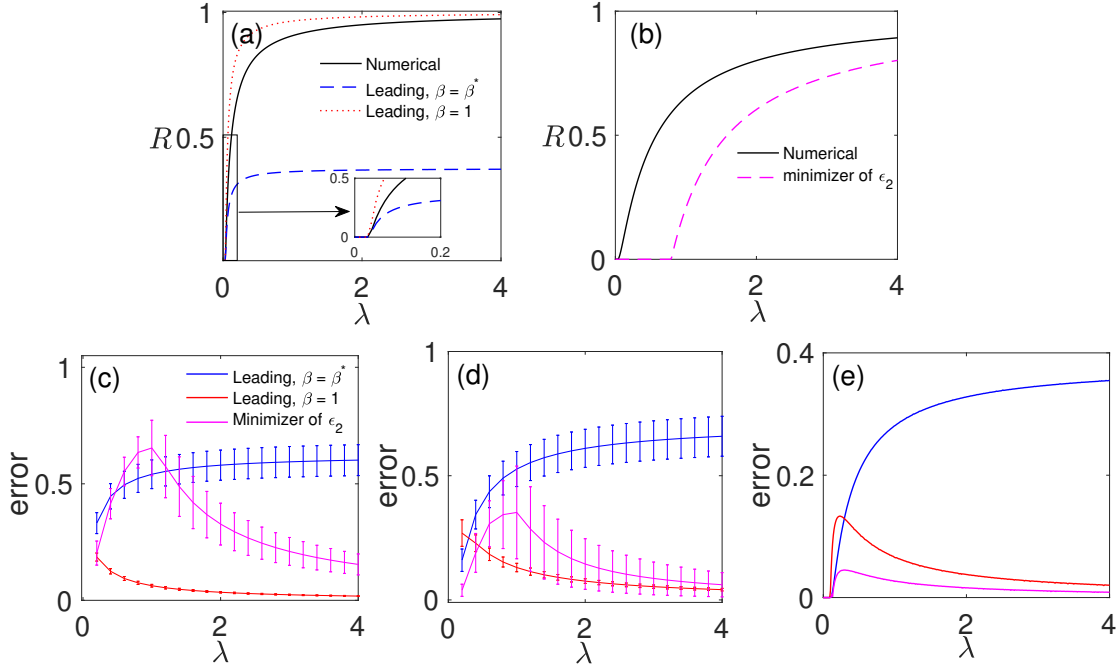


FIG. 3. Results for the SIS model. (a) Bifurcation diagram in terms of the observable,  $R$ , using the spectral method with the leading eigenvector as  $\mathbf{a}$  for a scale-free network with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$ . The inset is a magnification near  $\lambda = 0$  to show the behavior near the epidemic threshold. (b) Bifurcation diagram for the same network when  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . (c) Error for the spectral method with the different  $\mathbf{a}$  and  $\beta$  values for scale-free networks with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$ . We have calculated the error bars, which represent the standard deviations, on the basis of the 123 networks of which the minimized error  $\epsilon_2$  is less than 70% of the error obtained with the leading eigenvector. (d) Error for scale-free networks with  $N = 1000$ ,  $\langle k \rangle = 4$ , and  $\tilde{\gamma} = 3.5$ . We have calculated the error bars on the basis of the 162 networks of which the minimized error  $\epsilon_2$  is less than 70% of the error obtained with the leading eigenvector. (e) Error for the coauthorship network.

one-dimensional reduction with the minimizer of  $\epsilon_2$  (i.e.,  $\lambda \approx 0.80$ ). The results are qualitatively the same for scale-free networks with  $\langle k \rangle = 4$ , as we show in Fig. 3(d). We show the approximation error for the coauthorship network in Fig. 3(e). For this network, the minimizer of  $\epsilon_2$  realizes a smaller error than the leading eigenvector combined with  $\beta = \beta^*$  or  $\beta = 1$  for all values of  $\lambda$ . It should also be noted that the spectral method with the minimizer of  $\epsilon_2$  accurately locates the epidemic threshold for the coauthorship network.

### D. Double-well system

In this section, we consider the coupled double-well system given by

$$\frac{dx_i}{dt} = -(x_i - r_1)(x_i - r_2)(x_i - r_3) + D \sum_{j=1}^N w_{ij}x_j, \quad (27)$$

where  $x_i$  is the state of the node  $i$ ;  $D$  is the coupling strength, which we assume to be common for all edges;  $r_1$ ,  $r_2$ , and  $r_3$  are constants satisfying  $r_1 < r_2 < r_3$  [33–38]. We set  $r_1 = 1$ ,  $r_2 = 2$ , and  $r_3 = 5$ . When the coupling is absent, the dynamics is bistable, with  $x_i = r_1$  and  $x_i = r_3$  being stable equilibria and  $x_i = r_2$  being the unique unstable equilibrium. We use the initial condition  $x_i = 0.01, \forall i \in \{1, \dots, N\}$  for each value of  $D$ . With this initial condition,  $x_i$  will converge to the lower equilibrium, i.e.,  $r_1$ , if there is no coupling.

We plot in Fig. 4(a)  $R = \mathbf{a}^\top \mathbf{x}^*$ , where  $\mathbf{a}$  is the leading eigenvector, against  $D$  for the coupled double-well system on the same scale-free network with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$  as the one used in Figs. 3(a) and 3(b). The theoretical estimate of  $R$  shown in Fig. 4(a) does not depend on the  $\beta$  value because function  $G(x_i, x_j)$  does not depend on  $x_i$  for the double-well system such that Eq. (15) does not depend on  $\beta$ . We show in Fig. 4(b) the same relationships when  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . Similar to the case of the SIS model, the spectral method with the leading eigenvector is substantially better at locating the bifurcation point than the minimizer of  $\epsilon_2$  is. The accuracy at approximating  $R$  is slightly better for the minimizer of  $\epsilon_2$  than the leading eigenvector except near the bifurcation point, which we confirm by the statistical analysis of the absolute error shown in Fig. 4(c). The results are similar for  $\langle k \rangle = 4$  except that the minimizer of  $\epsilon_2$  yields significantly smaller errors than the leading eigenvector when  $D$  is large (see Fig. 4(d)). The results are similar for the coauthorship network (see Fig. 4(e)). We will discuss the relatively low accuracy near the bifurcation point in Section V A.

### E. Generalized Lotka-Volterra dynamics

Third, we consider the generalized Lotka-Volterra (GLV) dynamics is given by

$$\frac{dx_i}{dt} = \bar{\lambda}x_i + D \sum_{j=1}^N w_{ij}x_i x_j, \quad (28)$$

where  $x_i$  represents the abundance of the  $i$ th species,  $\bar{\lambda}$  is the intrinsic growth rate of the species, and  $D$  is the coupling strength as in the case of the double-well system [16, 18]. The nontrivial equilibrium of this dynamical system is given by  $\mathbf{x}^* = -\bar{\lambda}W^{-1}\mathbf{1}$ , where we remind that  $W = (w_{ij})$



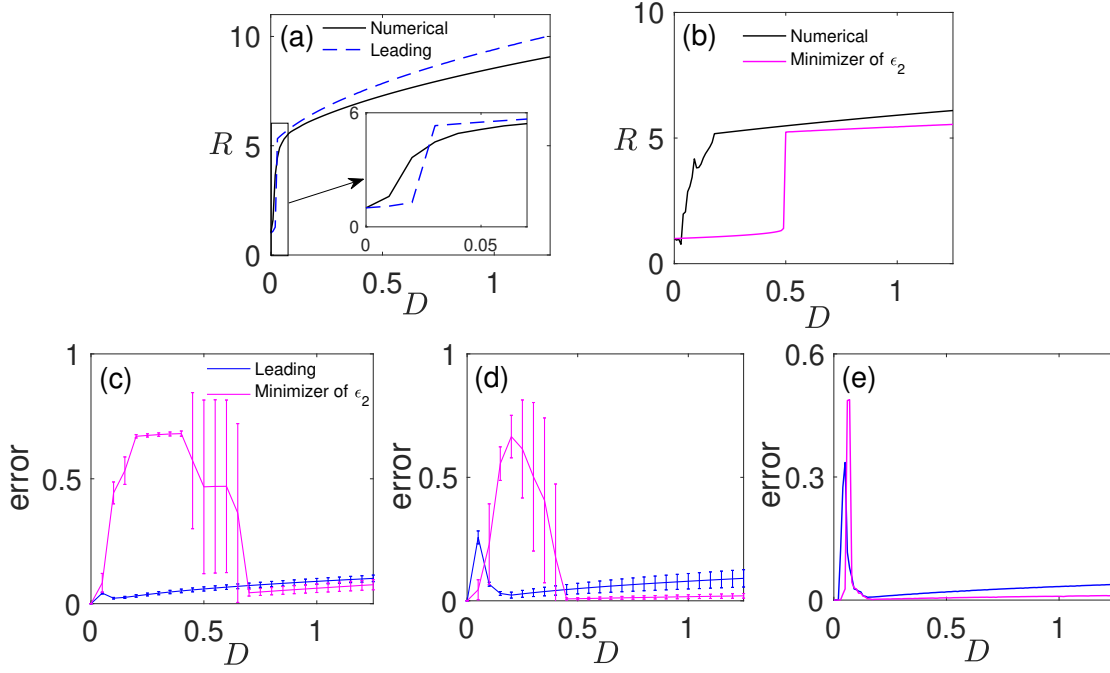


FIG. 4. Results for the double-well system. (a) Bifurcation diagram in terms of  $R = \mathbf{a}^\top \mathbf{x}^*$ , where  $\mathbf{a}$  is the leading eigenvector, for the same scale-free network with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$  as the one used in Fig. 3. (b) Bifurcation diagram for the same network when  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . (c) Error compared between the leading eigenvector and the minimizer of  $\epsilon_2$  as  $\mathbf{a}$  for scale-free networks with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$ . (d) Error for scale-free networks with  $N = 1000$ ,  $\langle k \rangle = 4$ , and  $\tilde{\gamma} = 3.5$ . (e) Error for the coauthorship network.

is the adjacency matrix. This equilibrium is globally asymptotically stable if and only if  $W$  is negative definite [39]. Therefore, only for the GLV dynamics, we change the definition of the adjacency matrix to set  $w_{ii} = -(\alpha_{\max} + 1)/D$ ,  $\forall i \in \{1, \dots, N\}$ , where  $\alpha_{\max}$  is the largest eigenvalue of  $W$ , which makes  $W$  negative definite [17]. With this  $w_{ii}$ , we rewrite Eq. (28) as

$$\frac{dx_i}{dt} = \bar{\lambda}x_i - cx_i^2 + D \sum_{j=1; j \neq i}^N w_{ij}x_ix_j, \quad (29)$$

where  $c = \alpha_{\max} + 1$ . Therefore, we set  $F(x_i) = \bar{\lambda}x_i - cx_i^2$  and  $G(x_i, x_j) = Dx_ix_j$  [17].

The equilibrium of Eq. (29) is given by  $\mathbf{x}^* = -\bar{\lambda}(DW - cI)^{-1}\mathbf{1}$ , where  $W$  is the weighted adjacency matrix of the original network with the diagonal entries being equal to 0, and  $I$  is the

$N \times N$  identity matrix. Therefore, we obtain

$$\begin{aligned}
R &= \mathbf{a}^\top \mathbf{x}^* \\
&= -\bar{\lambda} \mathbf{a}^\top (DW - cI)^{-1} \mathbf{1} \\
&= -\bar{\lambda} \mathbf{a}^\top (D\alpha - c)^{-1} \mathbf{1} \\
&= \frac{\bar{\lambda}}{\alpha_{\max} + 1 - D\alpha}.
\end{aligned} \tag{30}$$

On the other hand, the  $R$  value in the equilibrium for the one-dimensional reduction, Eq. (15), including the case of our modified spectral method (i.e., Eq. (23)) with the replacement of  $\beta^*$  by 1, is given by

$$R = \frac{\bar{\lambda}}{\alpha_{\max} + 1 - \beta^* D\alpha}. \tag{31}$$

Therefore, the spectral method is exact if and only if  $\beta^* = 1$  and regardless of which eigenvector of  $W^\top$  we use as  $\mathbf{a}$ .

With  $\bar{\lambda} = 0.5$ , we show in Fig. 5 the  $R$  values and the error as a function of  $D$  for the GLV model on the different networks. For Figs. 5(a) and 5(b), we use the same scale-free network with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$  as the one used in Figs. 3(a) and 3(b). With the leading eigenvector as  $\mathbf{a}$  and  $\beta = \beta^*$ , the spectral method is not accurate at approximating  $R$  (see the dashed line in Fig. 5(a)). In contrast, as the theory predicts, the theoretically obtained  $R$  perfectly matches the numerically obtained  $R$  when one combines the leading eigenvector and  $\beta = 1$  (see the dotted line in Fig. 5(a), which completely overlaps the solid line) or uses the minimizer of  $\epsilon_2$  (see Fig. 5(b), in which the dashed line completely overlaps the solid line). Note that  $R$  as a function of  $D$  looks qualitatively different between Figs. 5(a) and 5(b). This is because Eq. (30) diverges to infinity at  $D = D_c \equiv (\alpha_{\max} + 1)/\alpha$  as one increases  $D$  from a small value, and  $D \approx 1$ , which is the largest value of  $D$  in these figures, is much closer to  $D_c$  when  $\alpha = \alpha_{\max}$  than when  $\alpha$  is the eigenvalue associated with the minimizer of  $\epsilon_2$ . We statistically confirm the results shown in Figs. 5(a) and 5(b) for various scale-free networks in Figs. 5(c) and 5(d) and the coauthorship network in Fig. 5(e).

## V. ADVANTAGES OF THE LEADING EIGENVECTOR TO THE MINIMIZER OF $\epsilon_2$

The theoretical and numerical results shown in the previous sections suggest that the spectral method using the minimizer of  $\epsilon_2$  as  $\mathbf{a}$  outperforms that using the leading eigenvector as  $\mathbf{a}$  in many

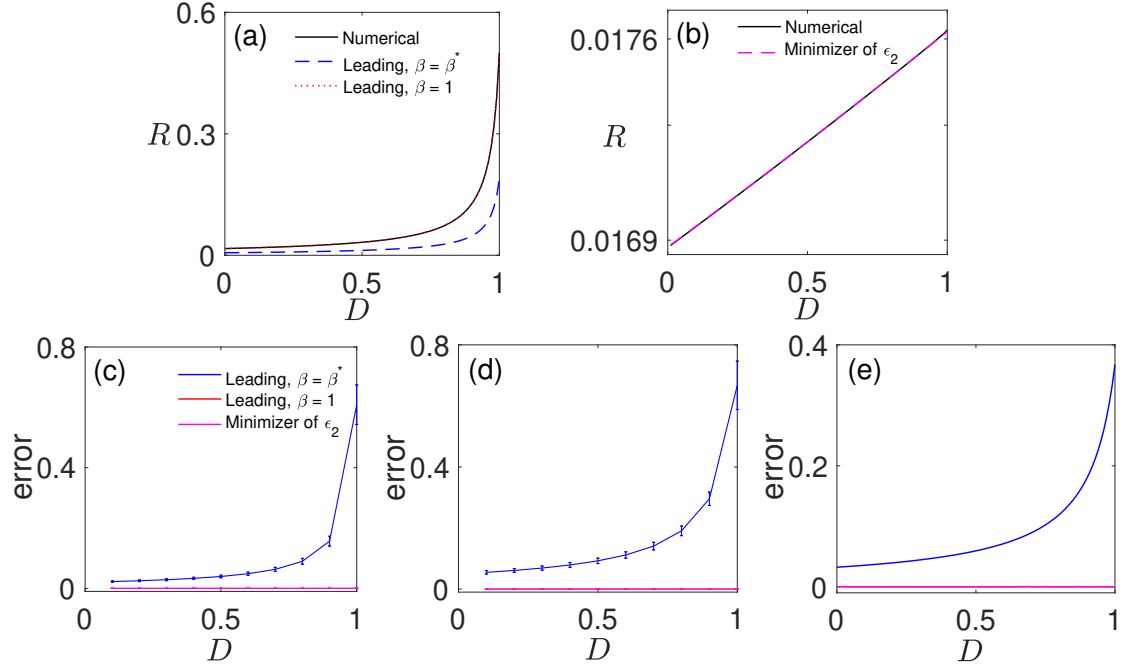


FIG. 5. Results for the GLV model. (a) Relationship between  $R$  and  $D$  for the spectral method with the leading eigenvector as  $\mathbf{a}$  for the same scale-free network with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$  as the one used in Fig. 3. The theoretical estimate for  $\beta = 1$  shown by the dotted line completely overlaps the numerical results shown by the solid line. (b) Same when we use the minimizer of  $\epsilon_2$  as  $\mathbf{a}$ . The theoretical estimate by the dashed line completely overlaps the numerical results shown by the solid line. (c) Error compared among the three cases for scale-free networks with  $N = 1000$ ,  $\langle k \rangle = 10$ , and  $\tilde{\gamma} = 3.5$ . (d) Error for scale-free networks with  $N = 1000$ ,  $\langle k \rangle = 4$ , and  $\tilde{\gamma} = 3.5$ . (e) Error for the coauthorship network. In (c), (d), and (e), the error in the case of the leading eigenvector combined with  $\beta = 1$  and the minimizer of  $\epsilon_2$  is equal to 0 regardless of the  $D$  value.

cases. However, in this section, we discuss two reasons why we still prefer the leading eigenvector as  $\mathbf{a}$ , but combined with  $\beta = 1$ , in the spectral method.

#### A. The spectral method with the leading eigenvector better predicts the bifurcation point

For the SIS model and double-well system, we have found that the spectral method using the minimizer of  $\epsilon_2$  as  $\mathbf{a}$  is much less accurate at locating the bifurcation point than the spectral method using the leading eigenvector as  $\mathbf{a}$ , as shown in Figs. 3(a), 3(b), 4(a), and 4(b). The reason for this phenomenon is as follows.

The modified spectral method (i.e., Eq. (23)) as well as the original spectral method (i.e., Eq. (15)) with  $\beta^*$  being replaced by 1 (see Figs. 3, 4, and 5) uses the eigenvalue of the adjacency

matrix,  $\alpha$ , as the effective coupling strength. We can interpret the earlier one-dimensional reduction developed by Gao et al. [15] as a mean-field theory to approximate the leading eigenvalue  $\alpha$  by a function of the nodes' degrees. Suppose that  $G(x_i, x_j) = D\tilde{G}(x_i, x_j)$ , where  $D$  is the coupling strength, which we regard as the bifurcation parameter. Note that, in the SIS model, the infection rate,  $\lambda$ , plays the role of  $D$ . Assume that the one-dimensional dynamical system given by Eq. (23) and  $G(x_i, x_j) = D\tilde{G}(x_i, x_j)$  with  $\alpha$  being the leading eigenvalue (i.e.,  $\alpha_{\max}$ ) undergoes a bifurcation at  $D = D_{c,\max}$ . We denote by  $D_{c,\text{org}}$  the value of  $D$  at which the bifurcation occurs in the original  $N$ -dimensional dynamical system. Empirically, the spectral method using the leading eigenvector as  $\mathbf{a}$  is not necessarily accurate at anticipating the bifurcation point, and one tends to obtain

$$\Delta D_{c,\max} \equiv D_{c,\max} - D_{c,\text{org}} > 0 \quad (32)$$

such that the spectral method tends to overestimate the bifurcation point for some models of population dynamics [17].

If we use a non-leading eigenvector as  $\mathbf{a}$ , we only replace  $\alpha$  in Eq. (23) by the eigenvalue associated with  $\mathbf{a}$ . Therefore, the bifurcation in the one-dimensional reduced dynamical system occurs at  $D = D_{c,\text{NL}}$ , where

$$\alpha_{\text{NL}} D_{c,\text{NL}} = \alpha_{\max} D_{c,\max}, \quad (33)$$

and  $\alpha_{\text{NL}}$  is the non-leading eigenvalue associated with the eigenvector  $\mathbf{a}$  under consideration. Using Eqs. (32) and (33), we evaluate the error in locating the bifurcation point by the spectral method with a non-leading eigenvector as follows:

$$\Delta D_{c,\text{NL}} \equiv D_{c,\text{NL}} - D_{c,\text{org}} = \frac{\alpha_{\max}}{\alpha_{\text{NL}}} D_{c,\max} - D_{c,\text{org}}. \quad (34)$$

We find that  $\Delta D_{c,\text{NL}} > \Delta D_{c,\max} > 0$  and that the difference between  $\Delta D_{c,\text{NL}}$  and  $\Delta D_{c,\max}$  is large if the eigenvalue ratio,  $\alpha_{\max}/\alpha_{\text{NL}}$ , is large. Therefore, if the minimizer of  $\epsilon_2$  is associated with a small eigenvalue of the adjacency matrix, the bifurcation point for the one-dimensional dynamical system,  $D = D_{c,\text{NL}}$ , tends to be much larger than  $D_{c,\text{org}}$ . Then, for  $D \in (D_{c,\text{org}}, D_{c,\text{NL}})$ , the one-dimensional reduction is qualitatively wrong at describing the original  $N$ -dimensional dynamical system such that our method is not expected to be accurate at approximating  $R$  for the original dynamical system. This is why we excluded in the beginning of Section IV C the eigenvectors  $\mathbf{a}$  that are associated with positive eigenvalues with tiny magnitudes and negative eigenvalues from the candidates of the minimizer of the error.

To further demonstrate the relevance of this reasoning, we show in Fig. 6 the largest eigenvalue and the eigenvalue whose associated eigenvector minimizes  $\epsilon_2$ . We use scale-free networks with

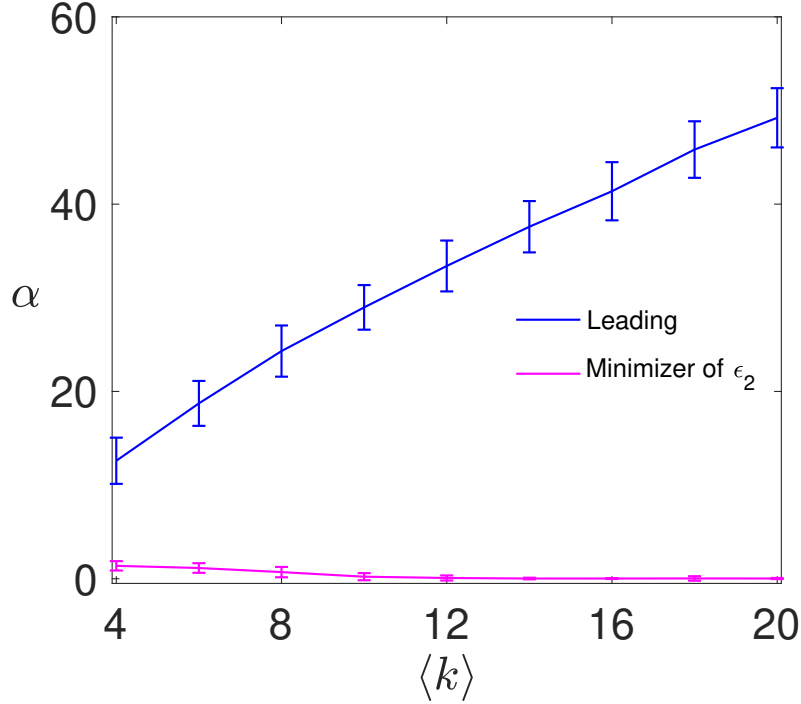


FIG. 6. Eigenvalues of the adjacency matrix of scale-free networks with  $N = 1000$  nodes and  $\tilde{\gamma} = 3.5$ . We show the leading eigenvalue and the eigenvalue associated with the minimizer of  $\epsilon_2$ . For each  $\langle k \rangle$  value, the error bar represents the average and standard deviation calculated on the basis of the subset of 200 networks for which the eigenvalue associated with the eigenvector minimizing  $\epsilon_2$  is at least  $10^{-6}$ .

$N = 1000$  and  $\tilde{\gamma} = 3.5$ , and vary  $\langle k \rangle$ . As we have done in the numerical simulations whose results are shown in Figs. 3, 4, and 5, we exclude the eigenvalues smaller than  $10^{-6}$ . Figure 6 indicates that the leading eigenvalue increases as  $\langle k \rangle$  increases, as theory predicts [27], and that the eigenvalue associated with the minimizer of  $\epsilon_2$  decreases as  $\langle k \rangle$  increases. Therefore, Eq. (34) implies that  $D_{c,NL}$  tends to be much larger than  $D_{c,org}$  as  $\langle k \rangle$  increases.

### B. The spectral method with the leading eigenvector is more robust against noise

We expect that the spectral method using the leading eigenvector as  $\mathbf{a}$  is more robust against noise than that using the minimizer of  $\epsilon_2$  for the following reason. Consider dynamical noise added to our dynamical system on networks, i.e., Eq. (1). Assume that each  $x_i$  fluctuates around the equilibrium in the case without noise,  $x_i^*$ , with mean 0 and standard deviation  $\sigma_i$ . For simplicity, we also assume that  $x_i$  and  $x_j$ , where  $i \neq j$ , are uncorrelated in the equilibrium; note that this assumption is just for the sake of discussion here and does not hold in general due to the interaction

between different nodes via edges [40, 41]. Then, the expectation of  $R$  in the equilibrium is given by  $R = \sum_{i=1}^N a_i x_i^*$ . The standard deviation of  $R$ , denoted by  $\sigma_R$ , is given by  $\sigma_R = \sqrt{\sum_{i=1}^N |a_i| \sigma_i^2}$ . We remind that  $\mathbf{a}$  is normalized such that  $\sum_{i=1}^N a_i = 1$ . We combine the following two observations to argue that  $R$  using the leading eigenvector as  $\mathbf{a}$  is more robust against noise than  $R$  using the minimizer of  $\epsilon_2$ . First, the expectation of  $R$  depends on  $\mathbf{a}$  unless  $x_1 = \dots = x_N$ . In our numerical results shown in Figs. 3(a), 3(b), 4(a), 4(b), 5(a), and 5(b),  $R$  tends to be smaller when  $\mathbf{a}$  is the minimizer of  $\epsilon_2$  than when  $\mathbf{a}$  is the leading eigenvector. Second, and more importantly,  $\sigma_R$  would be larger with the minimizer of  $\epsilon_2$  than with the leading eigenvector. When  $\mathbf{a}$  is the leading eigenvector, the Perron-Frobenius theorem guarantees that  $a_i > 0, \forall i \in \{1, \dots, N\}$ . In contrast, when  $\mathbf{a}$  is a nonleading eigenvector, including the case of the minimizer of  $\epsilon_2$ , the orthogonality of the eigenvectors associated with the different eigenvalues implies that some of  $a_i$  are positive and others are negative. Then, because  $\sum_{i=1}^N a_i = 1$ , the  $|a_i|$  value tends to be larger for a nonleading eigenvector than the leading eigenvector. For example, suppose that  $N = 3$ , that the normalized leading eigenvector is  $\mathbf{a} = (1/2, 1/4, 1/4)^\top$ , and a normalized nonleading eigenvector is  $\mathbf{a} = (-1, 1, 1)^\top$ . Then, if  $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$ , one obtains  $\sigma_R = \sqrt{\sum_{i=1}^N |a_i|} \sigma = \sigma$  for the leading eigenvector and  $\sigma_R = \sqrt{3} \sigma$  for the nonleading eigenvector. Therefore, the signal-to-noise ratio for  $R$  in the case of the leading eigenvector, quantified by the expectation divided by the standard deviation of  $R$ , should be smaller than the same ratio in the case of a non-leading eigenvector.

To examine the relevance of this argument, we run numerical simulations of the three dynamical systems with noise. We add an independent white noise with the intensity  $\sqrt{\mathcal{D}}$  to the right-hand side of each of the  $N$  differential equations constituting the dynamical system on the network. In other words, we add a value sampled from the Gaussian distribution with mean 0 and standard deviation  $\sqrt{\mathcal{D}dt}$  to the right-hand side of the differential equation in each integration time step of size  $dt$ . We set  $\sqrt{\mathcal{D}} = 0.2$  for the SIS and GLV models and  $\sqrt{\mathcal{D}} = 10$  for the double-well system. For the SIS and GLV models, once  $x_i$  becomes negative due to the added noise, it tends to diverge to  $-\infty$ . To prevent this phenomenon, once any  $x_i$  becomes negative in any simulation time step, we reset  $x_i$  to  $10^{-6}$ .

We show in Figs. 7(a) and 7(b) the numerical results for the SIS model run on the same scale-free network as the one used in Figs. 3(a) and 3(b). We verify that the fluctuation in  $R$  using the leading eigenvector as  $\mathbf{a}$ , shown by the solid line in Fig. 7(a), carries less noise than  $R$  using the minimizer of  $\epsilon_2$  as  $\mathbf{a}$ , shown by the solid line in Fig. 7(b). We stress that we have obtained Figs. 3(a) and 3(b) from the same simulation; we have only changed the observable,  $R$ . The results are qualitatively similar for the coauthorship network, while the difference between the two cases is

now smaller (see Figs. 7(c) and 7(d)). The results for the double-well system, shown in Figs. 7(e)–(h), and those for the GLV model, shown in Figs. 7(i)–(l), reinforce our claim that  $R$  is more robust against noise when one uses the leading eigenvector rather than the minimizer of  $\epsilon_2$  as  $\mathbf{a}$ .

The theoretical estimate based on the spectral method does not depend on the noise (see the dashed and dotted lines in Fig. 7). However, the target observable to be approximated,  $R$ , is noisier when  $\mathbf{a}$  is a non-leading eigenvector than the leading eigenvector. Therefore, we conclude that the one-dimensional reduction using the leading eigenvector as  $\mathbf{a}$  better describes the one-dimensional projection of the original network dynamics in the presence of noise unless the magnitude of the noise is small.

## VI. DISCUSSION

In the present study, we have explored two ideas to try to improve the spectral method to reduce dynamical systems on networks into a one-dimensional dynamics. The first idea is to constrain the use of the Taylor expansion of the dynamical variables,  $\{x_1, \dots, x_N\}$ , around one reference point,  $R$ . The original spectral method uses the Taylor expansion around  $R$  and  $\beta^* R$ , where  $\beta^* \neq 1$  in general. Our Taylor expansion has led to a one-dimensional reduction that does not contain  $\beta^*$  (i.e., Eq. (23)). Our second idea is to use the non-leading eigenvector of the adjacency matrix that minimizes the error. We have obtained explicit expressions of the errors, with which one can systematically search the minimizer of the error, especially that of  $\epsilon_2$  (see Eq. (22)). Note that our method requires the calculations of all the eigenvalues and eigenvectors, which costs  $O(N^3)$  time. This is an important limitation of the present method when we apply the method to large networks. In contrast, the previous studies used the leading eigenvector [19, 20]. We have found that, for networks that are relatively homogeneous in the degree, such as the Erdős-Rényi random graph and the Watts-Strogatz small-world network model, the leading eigenvector almost always minimizes the errors. In contrast, when the degree is heterogeneously distributed, the optimal eigenvector tended to be a non-leading one, which is particularly the case for larger and sparser networks. For these networks, our modified spectral method is expected to enjoy reduced approximation errors.

We have assessed the performance of approximating the one-dimensional observable,  $R$ , for three dynamical systems on scale-free networks for which the leading eigenvector and the minimizer of  $\epsilon_2$  do not tend to coincide. We have shown that the spectral method using the minimizer of  $\epsilon_2$  as  $\mathbf{a}$  tends to surpass the original spectral method across the different dynamical systems and networks (see Figs. 3, 4, and 5). However, we have also found that our spectral method using the minimizer

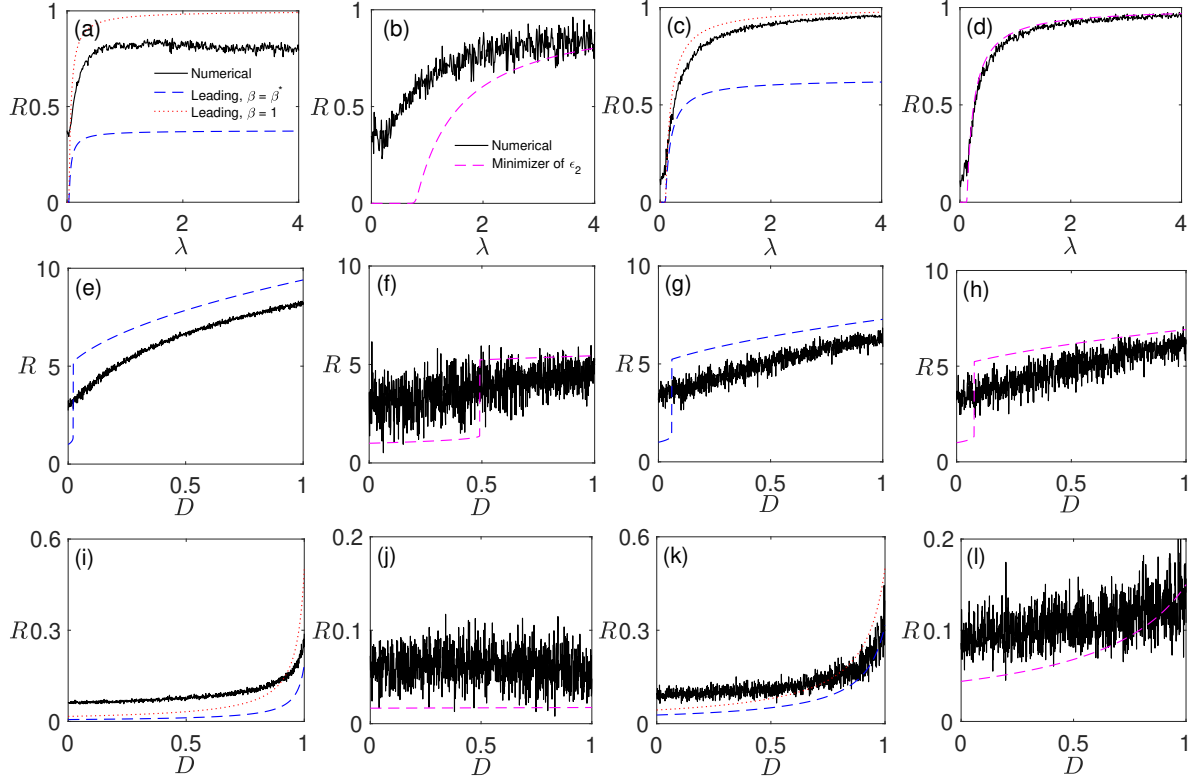


FIG. 7.  $R$  in the presence of noise. (a) SIS model on the scale-free network;  $\mathbf{a}$  is the leading eigenvector. (b) SIS model on the scale-free network;  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . (c) SIS model on the coauthorship network;  $\mathbf{a}$  is the leading eigenvector. (d) SIS model on the coauthorship network;  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . (e) Double-well system on the scale-free network;  $\mathbf{a}$  is the leading eigenvector. (f) Double-well system on the scale-free network;  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . (g) Double-well system on the coauthorship network;  $\mathbf{a}$  is the leading eigenvector. (h) Double-well system on the coauthorship network;  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . (i) GLV model on the scale-free network;  $\mathbf{a}$  is the leading eigenvector. (j) GLV model on the scale-free network;  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . (g) GLV model on the coauthorship network;  $\mathbf{a}$  is the leading eigenvector. (l) GLV model on the coauthorship network;  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . The scale-free network used in (a), (b), (e), (f), (i), and (j) is the same as the one used in Figs. 3(a) and 3(b). The theoretical estimates, i.e., the dashed and dotted lines, for the scale-free network shown in (a), (b), (e), (f), (i), and (j) are the same as those shown in Figs. 3(a), 3(b), 4(a), 4(b), 5(a), and 5(b), respectively.

of  $\epsilon_2$  as  $\mathbf{a}$  has two essential limitations. First, it is not good at estimating the bifurcation point. In particular, our method locates the bifurcation point extremely far from the correct value if the minimizer of  $\epsilon_2$  is associated with an eigenvalue that is much smaller than the largest eigenvalue (see Section V A). In this case, our one-dimensional reduction is qualitatively wrong over a wide range of the bifurcation parameter. Therefore, one cannot accurately approximate  $R$  of the original



high-dimensional dynamical system in a range of the bifurcation parameter of interest. Second, the spectral method using the minimizer of  $\epsilon_2$  as  $\mathbf{a}$  is not robust against noise. This is because, in the presence of dynamical noise, if one uses a non-leading eigenvector as  $\mathbf{a}$ , the variance of  $R$  is larger than in the case of the leading eigenvector used as  $\mathbf{a}$ . We have also examined the modified spectral method (i.e.,  $\beta = 1$ ) with  $\mathbf{a}$  being replaced by the leading eigenvector. Up to our numerical efforts, this combination is the best performer in that (i) it performs better than the original spectral method (i.e.,  $\beta = \beta^*$  and the leading eigenvector as  $\mathbf{a}$ ) and comparably with our method using the minimizer of  $\epsilon_2$  as  $\mathbf{a}$  in the absence of noise and that (ii) the target  $R$  fluctuates less than when  $\mathbf{a}$  is the minimizer of  $\epsilon_2$ . The spectral method using  $\beta = 1$  and the leading eigenvector is also better at locating the bifurcation point than that using the minimizer of  $\epsilon_2$ . Therefore, we recommend the combination of  $\beta = 1$  and the leading eigenvector, i.e., Eq. (22) provided with the leading eigenvector and eigenvalue. We may be able to find pairs of eigenvalue and eigenvector,  $(\alpha, \mathbf{a})$ , that suppress  $\epsilon_2$  much better than the leading eigenvector and the associate eigenvalue is not much smaller than the leading eigenvalue. Using such a pair  $(\alpha, \mathbf{a})$  may improve overall performances of the spectral method. Systematically investigating this issue warrants future work.

Both the original and modified spectral methods are based on the Taylor expansion of the differential equations in terms of  $x_1, \dots, x_N$  around one or multiple common values (e.g.,  $R$ ). Therefore, the methods are expected to work better when  $\{x_1, \dots, x_N\}$  is relatively homogeneous. We in fact reached the same conclusion for the GBB reduction in our previous work [17]. Our previous numerical simulations suggested that  $\{x_1, \dots, x_N\}$  was not wildly heterogeneous for various dynamical systems including those used in this paper, except near the bifurcation points [17]. However, we do not have a systematic understanding when  $\{x_1, \dots, x_N\}$  is more homogeneous than in other cases. Such an understanding is expected to help us to assess the applicability of the GBB reduction and spectral methods.

The DART is a systematic method to map high-dimensional dynamics on networks into a low-dimensional dynamical system, extending the spectral method [20]. The DART is applicable to a diversity of dynamics including synchronization dynamics on networks. Like the original spectral method [19], the DART is based on the Taylor expansion at multiple reference points and the leading eigenvectors of the adjacency matrix. It is interesting to apply the present approach to the DART, in particular to the cases where the reduced dynamics have the dimension larger than one. When the dimension of the reduced dynamics is larger than one, at least one eigenvector with which to take the weighted average of  $\{x_1, \dots, x_N\}$  (e.g., the eigenvector associated with the second largest eigenvalue of the adjacency matrix) will necessarily contain both positive and

negative elements due to the Perron-Frobenius theorem and the orthogonality of different eigenvectors. Then, the signal-to-noise ratio for the corresponding observable may be compromised in the presence of dynamical noise (see Section V B). With this possibility being included, it is worth further examining dimension reduction and resilience of noisy dynamical systems on networks.

## ACKNOWLEDGMENTS

N.M. acknowledges support from AFOSR European Office (under Grant No. FA9550-19-1-7024), the Sumitomo Foundation, the Japan Science and Technology Agency (JST) Moonshot R&D (under Grant No. JPMJMS2021), and the National Science Foundation (under Grant No. 2052720).

### Appendix A: Minimizing the error when the weight vector is an eigenvector of $K$

Let us require Eq. (12) to hold exactly instead of Eq. (13) in the spectral method. Then, the weight vector  $\mathbf{a}$  needs to be an eigenvector of  $K$ . Because  $K$  is a diagonal matrix,  $\mathbf{a}$  is a standard unit vector, i.e.,  $\mathbf{a} = \mathbf{e}_i \equiv (0, \dots, 0, \underbrace{1}_{i \text{ th entry}}, 0, \dots, 0)$  and  $\alpha = k_i^{\text{in}}$ , where  $i \in \{1, \dots, N\}$ . This implies that we use a single variable  $x_i$  to represent the entire system. Then, it is straightforward to derive  $\beta = 1$  and minimize the error for Eq. (13) in terms of  $\gamma$  as follows:

$$\epsilon_3 \equiv \min_{\gamma} \left\| W^{\top} \mathbf{e}_i - k_i^{\text{in}} \gamma \mathbf{e}_i \right\|^2 = \min_{\gamma} \left[ \sum_{\ell=1}^N (w_{i\ell})^2 + (k_i^{\text{in}})^2 \gamma^2 \right] = \sum_{\ell=1}^N (w_{i\ell})^2, \quad (\text{A1})$$

which is realized by  $\gamma = 0$ . We find that  $\epsilon_3$  is minimized with respect to  $i$  when the  $i$ th node has the smallest in-degree in the network. The corresponding one-dimensional reduction is given by

$$\frac{dx_{\bar{i}}}{dt} = F(x_{\bar{i}}) + k_{\bar{i}}^{\text{in}} G(R, 0), \quad (\text{A2})$$

where  $\bar{i}$  is the index of the node with the smallest in-degree.

In the case of our modified spectral method, we require that Eq. (20) exactly holds instead of Eq. (21). Then, we again obtain  $\mathbf{a} = \mathbf{e}_i$  and  $\alpha = k_i^{\text{in}}$ , where  $i \in \{1, \dots, N\}$ . Then, the error for Eq. (21) is given by

$$\epsilon_4 = \left\| W^{\top} \mathbf{a} - \alpha \mathbf{a} \right\|^2 = \sum_{\ell=1}^N (w_{i\ell})^2 + (k_i^{\text{in}})^2. \quad (\text{A3})$$

If the network is unweighted, we use  $(w_{i\ell})^2 = w_{i\ell} \in \{0, 1\}$  to obtain  $\epsilon_4 = k_i^{\text{in}} + (k_i^{\text{in}})^2$ , which is minimized when we select  $i$  with the smallest in-degree. In this case, the one-dimensional reduction

is given by

$$\frac{dx_{\bar{i}}}{dt} = F(x_{\bar{i}}) + k_{\bar{i}}^{\text{in}} G(R, R). \quad (\text{A4})$$

- 
- [1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Complex networks: Structure and dynamics, *Phys. Rep.* **424**, 175 (2006).
  - [2] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).
  - [3] M. A. Porter and J. P. Gleeson, *Dynamical Systems on Networks* (Springer, Cham, Switzerland, 2016).
  - [4] M. E. J. Newman, *Networks*, 2nd ed. (Oxford University Press, Oxford, UK, 2018).
  - [5] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Synchronization in complex networks, *Phys. Rep.* **469**, 93 (2008).
  - [6] A. E. Motter, S. A. Myers, M. Anghel, and T. Nishikawa, Spontaneous synchrony in power-grid networks, *Nat. Phys.* **9**, 191 (2013).
  - [7] E. Thébault and C. Fontaine, Stability of ecological communities and the architecture of mutualistic and trophic networks, *Science* **329**, 853 (2010).
  - [8] S. Allesina and S. Tang, Stability criteria for complex ecosystems, *Nature* **483**, 205 (2012).
  - [9] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman Hall/CRC, Boca Raton, FL, 2007).
  - [10] I. Hanski, Metapopulation dynamics, *Nature* **396**, 41 (1998).
  - [11] V. Colizza, R. Pastor-Satorras, and A. Vespignani, Reaction-diffusion processes and metapopulation models in heterogeneous networks, *Nat. Phys.* **3**, 276 (2007).
  - [12] G. E. Hinton and R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science* **313**, 504 (2006).
  - [13] Y. Bengio, A. Courville, and P. Vincent, Representation learning: A review and new perspectives, *IEEE Trans. Pat. Anal. Machine Intel.* **35**, 1798 (2013).
  - [14] J. P. Cunningham and Z. Ghahramani, Linear dimensionality reduction: Survey, insights, and generalizations, *J. Machine Learn. Research* **16**, 2859 (2015).
  - [15] J. Gao, B. Barzel, and A. L. Barabási, Universal resilience patterns in complex networks, *Nature* **530**, 307 (2016).
  - [16] 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).
  - [17] P. Kundu, H. Kori, and N. Masuda, Accuracy of a one-dimensional reduction of dynamical systems on networks, *Phys. Rev. E* **105**, 024305 (2022).
  - [18] C. Tu, P. D’Odorico, and S. Suweis, Dimensionality reduction of complex dynamical systems, *iScience* **24**, 101912 (2021).

- [19] E. Laurence, N. Doyon, L. J. Dubé, and P. Desrosiers, Spectral dimension reduction of complex dynamical networks, *Phys. Rev. X* **9**, 011042 (2019).
- [20] V. Thibeault, G. St-Onge, L. J. Dubé, and P. Desrosiers, Threefold way to the dimension reduction of dynamics on networks: An application to synchronization, *Phys. Rev. Research* **2**, 043215 (2020).
- [21] A. V. Goltsev, S. N. Dorogovtsev, J. G. Oliveira, and J. F. F. Mendes, Localization and spreading of diseases in complex networks, *Phys. Rev. Lett.* **109**, 128702 (2012).
- [22] T. Martin, X. Zhang, and M. E. J. Newman, Localization and centrality in networks, *Phys. Rev. E* **90**, 052808 (2014).
- [23] R. Pastor-Satorras and C. Castellano, Distinct types of eigenvector localization in networks, *Sci. Rep.* **6**, 18847 (2016).
- [24] R. Pastor-Satorras and C. Castellano, Eigenvector localization in real networks and its implications for epidemic spreading, *J. Stat. Phys.* **173**, 1110 (2018).
- [25] B. Barzel and A.-L. Barabási, Universality in network dynamics, *Nat. Phys.* **9**, 673 (2013).
- [26] N. Biggs, *Algebraic Graph Theory*, 2nd ed. (Cambridge University Press, Cambridge, UK, 1993).
- [27] F. Chung, L. Y. Lu, and V. Vu, Spectra of random graphs with given expected degrees, *Proc. Natl. Acad. Sci. USA* **100**, 6313 (2003).
- [28] D. J. Watts and S. H. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature* **393**, 440 (1998).
- [29] P. Holme and B. J. Kim, Growing scale-free networks with tunable clustering, *Phys. Rev. E* **65**, 026107 (2002).
- [30] M. E. J. Newman, Finding community structure in networks using the eigenvectors of matrices, *Phys. Rev. E* **74**, 036104 (2006).
- [31] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, *Rev. Mod. Phys.* **87**, 925 (2015).
- [32] I. Z. Kiss, J. C. Miller, and P. L. Simon, *Mathematics of Epidemics on Networks* (Springer, Cham, Switzerland, 2017).
- [33] N. Wunderling, B. Stumpf, J. Krönke, A. Staal, O. A. Tuinenburg, R. Winkelmann, and J. F. Donges, How motifs condition critical thresholds for tipping cascades in complex networks: Linking micro-to macro-scales., *Chaos* **30**, 043129 (2020).
- [34] C. D. Brummitt, G. Barnett, and R. M. D’Souza, Coupled catastrophes: Sudden shifts cascade and hop among interdependent systems, *J. R. Soc. Interface* **12**, 20150712 (2015).
- [35] J. Krönke, N. Wunderling, R. Winkelmann, A. Staal, B. Stumpf, O. A. Tuinenburg, and J. F. Donges, Dynamics of tipping cascades on complex networks, *Phys. Rev. E* **101**, 042311 (2020).
- [36] N. Wunderling, M. Gelbrecht, R. Winkelmann, J. Kurths, and J. F. Donges, Basin stability and limit cycles in a conceptual model for climate tipping cascades, *New J. Phys.* **22**, 123031 (2020).
- [37] A. K. Klose, V. Karle, R. Winkelmann, and J. F. Donges, Emergence of cascading dynamics in interacting tipping elements of ecology and climate, *R. Soc. Open Sci.* **7**, 200599 (2020).
- [38] N. Wunderling, J. F. Donges, J. Kurths, and R. Winkelmann, Interacting tipping elements increase

- risk of climate domino effects under global warming, *Earth Syst. Dyn.* **12**, 601 (2021).
- [39] 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).
- [40] H. Risken, *The Fokker-Planck Equation*, 2nd ed. (Springer, Berlin, 1989).
- [41] S. Chen, E. B. O’Dea, J. M. Drake, and B. I. Epureanu, Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems, *Sci. Rep.* **9**, 2572 (2019).