

Title	Structurally robust control of complex networks
Author(s)	Nacher, Jose C.; Akutsu, Tatsuya
Citation	Physical Review E (2015), 91(1)
Issue Date	2015-01-30
URL	<a href="http://hdl.handle.net/2433/196055">http://hdl.handle.net/2433/196055</a>
Right	©2015 American Physical Society
Type	Journal Article
Textversion	publisher

**Structurally robust control of complex networks**Jose C. Nacher<sup>\*†</sup>*Department of Information Science, Faculty of Science, Toho University, Miyama 2-2-1, Funabashi, Chiba 274-8510, Japan*Tatsuya Akutsu<sup>\*‡</sup>*Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, 611-0011, Japan*

(Received 3 May 2014; revised manuscript received 16 July 2014; published 30 January 2015)

Robust control theory has been successfully applied to numerous real-world problems using a small set of devices called *controllers*. However, the real systems represented by networks contain unreliable components and modern robust control engineering has not addressed the problem of structural changes on complex networks including scale-free topologies. Here, we introduce the concept of structurally robust control of complex networks and provide a concrete example using an algorithmic framework that is widely applied in engineering. The developed analytical tools, computer simulations, and real network analyses lead herein to the discovery that robust control can be achieved in scale-free networks with exactly the same order of controllers required in a standard nonrobust configuration by adjusting only the minimum degree. The presented methodology also addresses the probabilistic failure of links in real systems, such as neural synaptic unreliability in *Caenorhabditis elegans*, and suggests a new direction to pursue in studies of complex networks in which control theory has a role.

DOI: [10.1103/PhysRevE.91.012826](https://doi.org/10.1103/PhysRevE.91.012826)

PACS number(s): 89.75.Hc, 89.70.Eg, 89.75.Da, 89.75.Fb

**I. INTRODUCTION**

Real networks contain unreliable components; in critical infrastructures and technological networks some links may become nonoperational due to disasters or accidents, and in natural networks this might occur due to pathologies. Although the robustness and resilience of networks have been extensively investigated over the past decade [1–4], controllability methods for complex networks that can robustly manage structural changes have not been developed sufficiently. The existing research is limited to recent studies of network controllability under node attack and cascading failures using maximum matching [5–7], the discussion of quantitative measures of network robustness to investigate the effect of edge removal on the number of controllable nodes without a comprehensive theoretical analysis [8] and studies on multiagent systems under simultaneous failure of links and agents [9]. Recently, a relation between controllability robustness and core percolation [10] was investigated as an extension of the results shown in Ref. [5]. Note that the problem of how the number of driver nodes change as function of removal fraction of edges [6–8] and our question of how to design complex networks with structurally robust control feature drastically differ. While the former are heavily relying on percolation and cascading failure techniques well studied over the past decade, our work studies the minimum number of driver nodes to control the entire network against arbitrary single or multiple edge failures.

Robust control theory emerged in the late 1970s and is based on linear, time-invariant transfer functions. The controller is designed to change the system's model dynamics until it reaches a certain degree of uncertainty or disturbance. Thus,

the system is designed to be robust or stable against the presence of bounded modeling errors. To address disturbances, techniques such as single-input, single-output (SISO) feedback and H-infinity loop shaping were developed to avoid dynamic trajectories that deviate when disturbances enter the system [11,12]. Modern robust control engineering has been successfully applied to numerous real-world problems, such as stability in aircrafts and satellites and the efficiency of power, manufacturing, and chemical plants. By using a fault-tolerant control design a system can be robust enough to continue operating as intended even a small system failure occurs. Indeed, robust control for consensus and synchronization of networks have been investigated in a number of papers using a variety of techniques, including complex interconnected neural networks with delay [13]. The problem of fault-tolerant control aims to achieve a certain control objective in the presence of single or multiple edge or node failures. For example, this is a key technological and well-studied problem in networks where agents communicate over wireless connections (e.g., fault-tolerant control of wireless networks) [14,15].

Here we introduce the concept of structurally robust control of complex networks from a different perspective. To provide a concrete example, we adopt the minimum dominating set (MDS) model because it has been widely applied to the control of engineering systems, such as mobile *ad hoc* networks (MANET), transportation routing, computer communication networks [16–21], design of swapped networks for constructing large parallel and distributed systems [22], as well as the investigation of social influence propagation [23–25]. Recently, the relationship between structural controllability and MDS has been established and the size of MDS in a certain type of scale-free networks has been theoretically analyzed [26–28]. Molnár *et al.* further studied the size of MDS by exhaustively comparing several types of artificial scale-free networks using a greedy algorithm [29]. Interestingly, Wuchty demonstrated the applicability of the MDS approach to the controllability of protein interaction networks and showed

<sup>\*</sup>Corresponding authors.<sup>†</sup>nacher@is.sci.toho-u.ac.jp<sup>‡</sup>takutsu@kuicr.kyoto-u.ac.jp

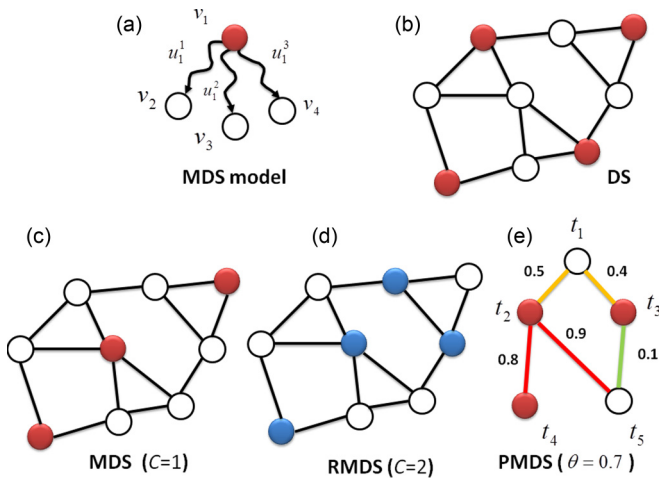


FIG. 1. (Color online) The MDS model and the robust domination configuration. (a) The network is structurally controllable by selecting a minimum dominating set (MDS) because each dominated node has its own control signal. A maximum matching approach needs three driver nodes  $v_1$ ,  $v_2$ , and  $v_4$ , assuming a matching link from  $v_1$  to  $v_3$ . In contrast, the MDS only requires one node,  $v_1$ . The labels  $u_1^1$ ,  $u_1^2$ , and  $u_1^3$  indicate control. (b) Example of a dominating set (DS) in a graph  $G$ . A set of nodes  $S$  (filled) in a graph  $G$  is a dominating set if every node in  $G$  is either an element of  $S$  or adjacent to an element of  $S$ . (c) The graph  $G$  allows for an MDS with cover  $C = 1$  composed of only three nodes. (d) Example of a robust minimum dominating set (RMDS) (i.e., an MDS with cover  $C = 2$ ). In an RMDS, each node in  $G$  is either an element of  $S$  or adjacent to at least two elements of  $S$ . (e) Each edge  $(t_i, t_j)$  has a failure probability of  $P_{t_i, t_j}$  highlighted in the figure. In a probabilistic minimum dominating set (PMDS), the approach shown here, each node must be covered by multiple nodes in the MDS so the probability that at least one edge is active is at least  $\theta = 0.7$ . For instance,  $t_4$  must cover itself because the link  $(t_2, t_4)$  is unreliable with a failure probability of 0.8, which is higher than  $\theta = 0.7$ .

that the MDS of proteins were enriched with essential cancer-related and virus-targeted genes [30]. Moreover, very recently strategies on vulnerability of dominating sets against random and targeted attacks have also been reported [31].

A set of nodes  $S$  in a graph  $G$  is a dominating set (DS) if every node in  $G$  is either an element of  $S$  or adjacent to an element of  $S$ . Then the MDS approach states that a network is made structurally controllable by selecting an MDS (driver set) because each dominated node has its own control signal [26–28] [see Fig. 1(a)]. Whereas each element is controlled by at least one node in  $G$  ( $C = 1$ ) (or is covered by itself) in an MDS, the novel robust MDS (RMDS) approach states that each node must be covered by itself or at least two nodes in  $G$  ( $C = 2$ ) [see Figs. 1(c) and 1(d)]. The analytical results and computer simulations demonstrate that a robust configuration ( $C = 2$ ,  $D = 2$ ) and nonrobust configuration ( $C = 1$ ,  $D = 1$ ) of a scale-free network with minimum degree  $D$  require the same order of driver nodes. The robust configuration guarantees that the system remains controllable even under arbitrary single or multiple link failure. This finding has remarkable implications for designing technical and natural systems that can still operate in the presence of unavailable or damaged links because the implementation of such a robust system in a large

network does not change the order of the required controllers in a conventional system without robustness capability. As a by-product of this research, our results also demonstrate that the minimum degree  $D$  in a network plays an important role in network controllability and significantly affects the size of the MDS. In particular, for  $\gamma < 2$ , the order of the size of an MDS changes if the minimum degree changes, unveiling another tool to reduce the number of driver nodes. These theoretical findings are confirmed by computer simulations and an analysis of real-world undirected, directed, and bipartite networks. In addition, the MDS approach is extended to address probabilistic network domination when we consider the probability of link transmission failure. The derived mathematical tools allow us to identify optimal controllability configurations in real biological systems by mapping the synaptic unreliability distribution experimentally observed in rat brains [32] to the most well-known and recently updated neural network model for *Caenorhabditis elegans* [33].

The connections between the proposed study and developments on graph theory such as eternal domination are worth mentioning. The robust dominating set can be mathematically seen as a  $k$ -dominating set. Further extensions on  $k$ -dominating sets led to the concept of the Roman  $k$ -dominating set and eternal security [34–36]. An eternal dominating or eternally secure set can be defined as a dominating set that allows an eternal defense. The first version of these concepts was mathematically introduced by Burger *et al.* [37,38]. For this definition, we have to understand the nodes as regions, the edges as adjacent regions, and a dominating set  $S$  as military bases of protecting armies located in some selected regions. Then, when a node  $v$  that does not belong to the current dominating set  $S$  is attacked, the army at a node of  $S$  adjacent to  $v$  can be deployed to  $v$  to defend the region from the attack. This process generates a dominating set with the same feature of eternal defense. However, the relations with robust control of complex networks were not investigated in these works.

The concept of structural controllability was first introduced by Lin [39] for single-input systems and it was quickly extended to multi-input systems [40–43]. The maximum matching (MM) algorithm identifies the minimum number of nodes to control the entire network by providing a mapping between structural controllability and network structure [5]. However, there are several striking differences with MDS approach: (1) By using the MM approach, the fraction of driver nodes tends to be minimized in random networks. The MDS does not necessarily give a minimum number of driver nodes in the sense of MM approach. However, MDS gives fewer driver nodes in many cases, including scale-free networks in which hubs are present [1,44]. For example, consider the star graph (all nodes but one node are leaves) with  $n$  leaves. The MM approach then needs  $n-1$  driver nodes, whereas the MDS approach needs only one driver node. (2) The MM approach is based on linear systems, whereas the MDS approach does not even need structural controllability: It is enough to assume that a node is controllable if it is directly connected to a driver node. This represents one of the unique features of the MDS model because it suggests that it can be applied to a certain kind of nonlinear and/or discrete model. However, these striking advantages have a price. (1) The set of driver nodes in MDS is  $O(n)$  in scale-free networks with  $\gamma > 2$ . (2) Each edge has

to be controlled independently. However, even in the case of  $O(n)$ , a relatively small number of drivers is required in most cases. Above all, the proposed concept of structurally robust control is an algorithmic-independent framework. Therefore, engineering applications of structurally robust control may flexibly give preference to one algorithm over another.

## II. THEORETICAL RESULTS FOR ROBUST DOMINATION

### A. Analysis for the case $\gamma < 2$ with minimum degree $D$ and MDS with cover $C = 1$

In the following, we present analytically derived predictions for the minimum number of drivers using the MDS controllability approach by considering specific cases for the degree exponent  $\gamma$  and the minimum degree  $D$ . Then the robust control is also analyzed by considering the number of drivers  $C$  required to cover each node. We first assume that the minimum degree of an undirected graph  $G(V, E)$  with  $n$  nodes and a degree distribution that follows a power law  $P(k) = \alpha k^{-\gamma}$  is  $D$ . We then use a standard mean-field approach that assumes a continuum approximation for the degree  $k$ , so it becomes a continuous real variable [1,44].

We also note that there are some discussions on degree cutoff [45] because our analysis assumes that there exist high-degree nodes. However, we do not introduce such a degree cutoff because we are performing a kind of mean-field analysis. It is to be noted that scale-free networks are a kind of random networks and thus we can have a node with even degree  $n - 1$  with very small probability if  $P(n - 1) > 0$  [1]. In the mean-field analysis, such rare cases are taken into account. However, discussions of degree cutoff are based on average case analysis, and there is no consensus for a cut-off value. Therefore, we do not introduce degree cutoff in our analysis. The results of computer simulation support that our analysis is appropriate. Note also that each node with degree more than 1 must be covered by  $C$  nodes (not by  $C$  edges) in our integer linear programming (ILP) formulation and thus the effect of multiple edges is eliminated in computer simulation. As it has been shown in the field of complex network science, the analysis and classification of networks in terms of their degree distribution is a key feature to understand the complex behavior of complex systems. In particular, the scale-free topology fundamentally changes the system's behavior, with broad implications from spreading processes on networks (like, for example, the spread of infectious diseases) to cascading failures [1,5,44]. It is therefore appropriate to examine the controllability problem in networks governed by power-law degree distributions.

First we assume that the minimum degree  $D$  is 2 in an undirected graph  $G(V, E)$ , where  $V$  is a set of  $n$  nodes and  $E$  is a set of edges connecting nodes in  $V$ . From the following equation:

$$\begin{aligned} \alpha n \int_2^n k^{-\gamma} dk &= \frac{\alpha n}{\gamma - 1} \left( \frac{1}{2^{\gamma-1}} - \frac{1}{n^{\gamma-1}} \right) \\ &\approx \frac{\alpha n}{\gamma - 1} \frac{1}{2^{\gamma-1}} = n \end{aligned}$$

we have  $\alpha = 2^{\gamma-1}(\gamma - 1)$ .

Let DS be the set of nodes with degree between  $n^\beta$  and  $n$ . Then, the number of nodes in DS (denoted by  $N_{\text{DS}}$ ) is

$$\begin{aligned} N_{\text{DS}} &\approx \alpha n \int_{n^\beta}^n k^{-\gamma} dk = -2^{\gamma-1} n [k^{1-\gamma}]_{n^\beta}^n \\ &= -2^{\gamma-1} n (n^{1-\gamma} - n^{\beta(1-\gamma)}) \approx 2^{\gamma-1} n^{1+\beta(1-\gamma)}. \end{aligned}$$

Let  $E_G$  be the number of edges in  $G(V, E)$ . Then  $E_G$  is given by

$$E_G \approx \frac{\alpha n}{2} \int_2^n k k^{-\gamma} dk \approx \frac{\alpha n}{2(2-\gamma)} n^{2-\gamma},$$

where the factor 2 in  $\frac{\alpha n}{2}$  comes from the fact that each edge is counted by two nodes. The number of edges that are connected to at least one node in DS (i.e., the number of edges covered by DS) is lower bounded by

$$E_{\text{DS}} = \frac{\alpha n}{2} \int_{n^\beta}^n k k^{-\gamma} dk = \frac{\alpha n}{2(2-\gamma)} [n^{2-\gamma} - n^{\beta(2-\gamma)}].$$

It should be noted that  $E_{\text{DS}}$  gives a lower bound and the number of edges covered by DS may be much larger because this estimate considers the case where both end points of these edges are in DS.

The probability that an arbitrary edge is not covered by DS is upper bounded by

$$\frac{E_G - E_{\text{DS}}}{E_G} \approx \frac{n^{\beta(2-\gamma)}}{n^{2-\gamma}} = n^{(\beta-1)(2-\gamma)}.$$

Then the probability that a node with degree  $k$  does not have any edge connected to DS is upper bounded by

$$n^{k(\beta-1)(2-\gamma)},$$

which is also upper bounded by  $n^{2(\beta-1)(2-\gamma)}$  because the minimum degree is assumed to be 2. Therefore, the number of nodes (denoted by  $N_{G-\text{DS}}$ ) not covered by DS is

$$N_{G-\text{DS}} \leq O(n n^{2(\beta-1)(2-\gamma)}) = O(n^{1+2(\beta-1)(2-\gamma)}).$$

Since we can have a dominating set if we merge these nodes with DS, the number of nodes in an MDS is upper bounded by  $N_{\text{DS}} + N_{G-\text{DS}}$ . To minimize the order of  $N_{\text{DS}} + N_{G-\text{DS}}$ , we let

$$1 + \beta(1 - \gamma) = 1 + 2(\beta - 1)(2 - \gamma),$$

which results in

$$\beta = \frac{2(2 - \gamma)}{3 - \gamma}.$$

By using this  $\beta$ , an upper bound of the size of an MDS is estimated as

$$O[n^{1 - \frac{2(\gamma-1)(2-\gamma)}{3-\gamma}}].$$

We can see that this order is smaller than that of our previous result on  $D = 1$  [27]

$$O[n^{1 - (\gamma-1)(2-\gamma)}].$$

In particular, the above takes the minimum order  $O(n^{0.75})$  when  $\gamma^* = 1.5$  for  $D = 1$ , whereas the new bound for  $D = 2$  takes the minimum order  $O(n^{0.657})$  when  $\gamma^* = 3 - \sqrt{2}$  [see Figs. 2(a) and 2(b)]. This difference comes from the fact that a node  $v$  is regarded as not covered by DS if one specific edge

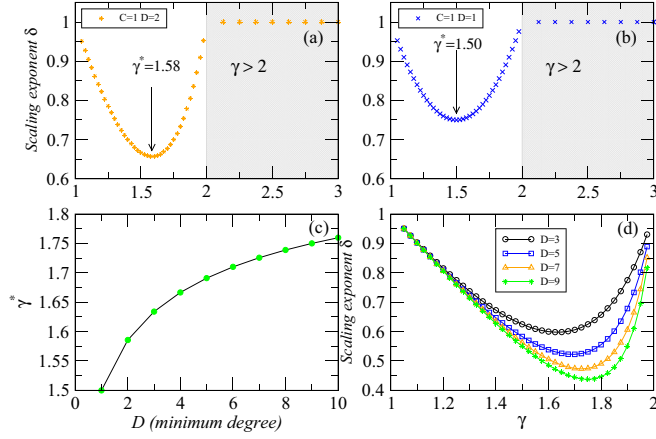


FIG. 2. (Color online) The theoretical predictions for the MDS order. Theoretical analysis illustrates that the MDS size scales according to  $n^\delta$  [see Eq. (3)] in scale-free networks of size  $n$ . The scaling exponent  $\delta$  as a function of the degree exponent  $\gamma$  for the (a)  $C = 1, D = 2$  and (b)  $C = 1, D = 1$  configurations. (c) The degree exponent that minimises the MDS as a function of the minimum degree  $D$ . (d) The dependence of the  $\delta$  exponent on the degree exponent  $\gamma$  calculated for several  $D$  values.

connected to  $v$  is not covered by DS in an existing analysis [27], whereas a node  $v$  is regarded as not covered by DS if no edge connected to  $v$  is not covered by DS in this analysis.

We can extend the above result for the case where the minimum degree is  $D$  by replacing

$$N_{G-DS} \leq O[nn^{2(\beta-1)(2-\gamma)}] = O[n^{1+2(\beta-1)(2-\gamma)}],$$

with

$$N_{G-DS} \leq O[nn^{D(\beta-1)(2-\gamma)}] = O[n^{1+D(\beta-1)(2-\gamma)}].$$

Then we have

$$1 + \beta(1 - \gamma) = 1 + D(\beta - 1)(2 - \gamma), \quad (1)$$

$$\beta = \frac{D(2 - \gamma)}{D(2 - \gamma) + (\gamma - 1)}. \quad (2)$$

By using this  $\beta$ , an upper bound of the size of MDS is estimated as

$$O\left[n^{1 - \frac{D(2-\gamma)(\gamma-1)}{D(2-\gamma)+\gamma-1}}\right]. \quad (3)$$

This order of the MDS size that scales as  $n^\delta$  takes the minimum value when

$$\gamma^* = \frac{(2D - 1) - \sqrt{D}}{D - 1}.$$

It is to be noted that although  $\alpha$  depends on  $D$ , it does not affect the order of the MDS size. The scaling exponent  $\delta$  for the order of the MDS size is shown as a function of the degree exponent  $\gamma$  in Fig. 2(d). This is our first main result and demonstrates that for scale-free networks with  $\gamma < 2$ , the order of the MDS size changes (the exponent  $\delta$  changes in functional form of  $n^\delta$ ) when the minimum degree increases. The dependence of the degree exponent  $\gamma^*$  that minimizes the MDS size on the minimum degree  $D$  is also shown in Fig. 2(c). The results

demonstrate that a higher minimum degree makes it easier to control scale-free networks with  $\gamma < 2$  [Fig. 2(d)].

## B. Analysis on robust domination with minimum degree $D$ and a generic $C$ cover

### 1. Analysis for the case of $\gamma < 2$

Next we show the results for the robust domination (RMDS) [Fig. 1(d)]. For an undirected graph  $G(V, E)$  and a positive integer  $C$ , RDS  $\subseteq V$  is called a  $C$ -robust dominating set if each node  $v \in V$  satisfies the following: either  $v \in$  RDS or  $v$  is connected to  $C$  or more nodes in RDS. Here, we provide an upper bound of the size of the minimum  $C$ -robust dominating set. Note that an RDS is a special case of a generalized dominating set [46,47], which has been studied in the context of the computational complexity. However, it has not been investigated from the perspective of complex networks. We consider the case of  $C$ -robust domination in which the minimum degree of  $G(V, E)$  is  $D$ , where  $C$  and  $D$  are constants such that  $D \geq C$ .

As in Sec. II A, let DS be the set of nodes with degree between  $n^\beta$  and  $n$ . Then the probability that a node is not covered by  $C$  or more nodes in DS is bounded by

$$O\left(\sum_{k=D-C+1}^D \binom{D}{k} n^{(\beta-1)(2-\gamma)k}\right),$$

where we do not include the factor of  $[1 - O(n^{(\beta-1)(2-\gamma)})]^{D-k}$  because we consider an upper bound. This number is further bounded by

$$O[DD^{D-C+1}n^{(\beta-1)(2-\gamma)(D-C+1)}]$$

for sufficiently large  $n$ . Therefore, the number of nodes not covered by DS is

$$O[nn^{(\beta-1)(2-\gamma)E}] = O[n^{1+E(\beta-1)(2-\gamma)}],$$

where we let  $E = D - C + 1$ . It is to be noted that a constant factor is ignored here because we use  $O$  notation.

As before, by balancing the size of DS and the number of noncovered nodes, we have

$$1 + \beta(1 - \gamma) = 1 + E(\beta - 1)(2 - \gamma), \quad (4)$$

$$\beta = \frac{E(2 - \gamma)}{E(2 - \gamma) + (\gamma - 1)}. \quad (5)$$

By using this  $\beta$ , an upper bound of the size of RMDS is estimated as

$$O\left[n^{1 - \frac{E(2-\gamma)(\gamma-1)}{E(2-\gamma)+\gamma-1}}\right]. \quad (6)$$

This is our second and most important finding. This result suggests that the case of an RMDS with minimum cover  $C$  and minimum degree  $D$  corresponds to the case of an MDS with the minimum degree  $D - C + 1$ . For example, the case of an RMDS with  $C = D = 2$  [i.e., the case where each node (with a degree of at least 2) must be covered twice, and the minimum degree  $D$  is 2] corresponds to the case of an MDS with  $D = 1$ .

In all theoretical analyses in Secs. II A and IIB, we assume that multiedges (i.e., multiple edges between the same pair



of nodes) are allowed because it is known that there does not exist a network strictly following a power-law distribution with  $\gamma < 2$  if multiedges are not allowed [48]. However, even if multiedges between the same pairs are replaced by single edges after generating a power-law network with multiedges, the results should hold if the cover parameter  $C$  is 1 because MDS is only concerned with existence of an edge from each node not in MDS to a node in MDS. If  $C > 1$ , we need to consider the possibility that some of  $C$  edges are connected to a node  $v$  are multiedges because such  $v$  may not be dominated by  $C$  nodes. We will show that such a factor can be ignored in many cases if we discuss the order of the size of MDS. Of course, the resulting network does not strictly follow a power-law distribution if multiedges are replaced by single edges. However, because any network with  $\gamma < 2$  cannot strictly follow a power-law distribution, our assumption seems reasonable.

Here we note that Eqs. (4) and (5) are identical to Eqs. (1) and (2) if we replace  $E$  by  $D$ . It suggests that the case of RMDS with the minimum cover  $C$  and the minimum degree  $D$  corresponds to the case of MDS with the minimum degree  $D - C + 1$ . For example, the case of RMDS with  $C = D = 2$  corresponds to the case of MDS with  $D = 1$ .

In the above, we implicitly assumed that all  $C$  edges are connected to different nodes in DS. However, we need to consider the possibility that some of  $C$  edges are connected to the same node in DS because we allow multiedges in theoretical analyses. Suppose that  $C$  (or more) edges from  $v$  are connected to DS. Since the number of edges connected to a node of degree  $k$  in DS is  $O(k)$ , there exist  $O(nk^{-\gamma})$  nodes of degree  $k$  in DS, and there exist  $O(n^{3-\gamma})$  edges connected to DS, the probability that  $C$  edges contain at least one common end point in DS is

$$O\left(C^2 \frac{\int_{k=1}^n k^2 (nk^{-\gamma}) dk}{(n^{3-\gamma})^2}\right) = O\left(C^2 \frac{1}{n^{2-\gamma}}\right).$$

Since there exist  $O(n)$  nodes covered by DS, the number of nodes not covered by  $C$  different nodes would be

$$O(C^2 n^{\gamma-1}).$$

If the exponent  $\gamma - 1$  is smaller than that in Eq. (6), this factor does not affect the order of Eq. (6). For  $\gamma < 1.5$ , it is true for  $E \leq 10$ . However, if  $\gamma \geq 1.7$ , it is true only for  $E = 1$  (i.e.,  $C = D$ ). Therefore, we need to be careful if we consider the case of  $D > C$  and  $\gamma \geq 1.7$ .

## 2. Analysis for the case of $\gamma > 2$

*a. Analysis of lower bound.* First we consider a lower bound. Let  $D$  be the minimum degree. From  $\alpha n \int_D^n k^{-\gamma} dk = n$ , we have  $\alpha = (\gamma - 1)D^{\gamma-1}$ .

For  $S \subseteq V$ ,  $\Gamma(S)$  denotes the set of edges between  $S$  and  $V - S$  [i.e.,  $\Gamma(S) = \{\{u, v\} | u \in S \text{ and } v \in V - S\}$ ]. Here we assume without loss of generality that  $|S| < n/2$  because we are interested only in cases where  $|S|$  is small compared with  $n$ . The following property is trivial:

$$\text{if } |\Gamma(S)| < n/2, S \text{ cannot dominate } V. \quad (7)$$

Let  $S$  be the set of nodes whose degree is greater than or equal to  $K$ . We estimate the size of  $\Gamma(S)$  as follows:

$$\begin{aligned} |\Gamma(S)| &< \alpha n \int_K^n k k^{-\gamma} dk \approx n(\gamma - 1)D^{\gamma-1} \int_K^n k^{-\gamma+1} dk \\ &= nD^{\gamma-1} \left(\frac{\gamma - 1}{\gamma - 2}\right) \left(\frac{1}{K^{\gamma-2}} - \frac{1}{n^{\gamma-2}}\right) \\ &< nD^{\gamma-1} \left(\frac{\gamma - 1}{\gamma - 2}\right) \frac{1}{K^{\gamma-2}}. \end{aligned}$$

If  $S$  is a dominating set, the last term should be no less than  $n/2$ . Therefore, the following inequality should be satisfied:

$$nD^{\gamma-1} \left(\frac{\gamma - 1}{\gamma - 2}\right) \frac{1}{K^{\gamma-2}} > n/2. \quad (8)$$

By solving this inequality, we have

$$\begin{aligned} K &< \left[ D^{\gamma-1} \left(\frac{\gamma - 1}{\gamma - 2}\right) \left(\frac{n}{n/2}\right) \right]^{1/(\gamma-2)} \\ &= \left[ 2D^{\gamma-1} \left(\frac{\gamma - 1}{\gamma - 2}\right) \right]^{1/(\gamma-2)}. \end{aligned}$$

Then the size of  $S$  is estimated as

$$\begin{aligned} |S| &\approx \alpha n \int_K^n k^{-\gamma} dk \approx n \left( \frac{1}{K^{\gamma-1}} - \frac{1}{n^{\gamma-1}} \right) \approx n \frac{1}{K^{\gamma-1}} \\ &> \left[ 2D^{\gamma-1} \left(\frac{\gamma - 1}{\gamma - 2}\right) \right]^{-\frac{\gamma-1}{\gamma-2}} n. \end{aligned}$$

We extend the above analysis to  $C$  domination (i.e., each node must be covered by  $C$  or more edges). In this case, Ineq. (7) should be replaced by

$$\text{if } |\Gamma(S)| < nC/2, S \text{ cannot } C\text{-dominate } V.$$

Then, Ineq. (8) is also replaced by

$$nD^{\gamma-1} \left(\frac{\gamma - 1}{\gamma - 2}\right) \frac{1}{K^{\gamma-2}} > (nC)/2.$$

Finally, we have

$$\begin{aligned} |S| &> \left[ \left( D^{\gamma-1} \frac{2}{C} \right) \left(\frac{\gamma - 1}{\gamma - 2}\right) \right]^{-\frac{\gamma-1}{\gamma-2}} n \\ &= \left( \frac{C}{2D^{\gamma-1}} \right)^{\frac{\gamma-1}{\gamma-2}} \left[ \left(\frac{\gamma - 1}{\gamma - 2}\right) \right]^{-\frac{\gamma-1}{\gamma-2}} n. \quad (9) \end{aligned}$$

For example, consider the case of  $C = 2$  and  $\gamma = 3$  for fixed  $D$ . In this case, 2-domination requires  $2^2 = 4$  times larger MDS.

*b. Analysis of upper bound.* Next we consider an upper bound. As in the above, we have  $\alpha = (\gamma - 1)D^{\gamma-1}$ . Let DS be the set of nodes with degree between  $B$  and  $n$ . Then, the size of DS,  $N_{DS}$ , is estimated as

$$\begin{aligned} N_{DS} &\approx \alpha n \int_B^n k^{-\gamma} dk = -nD^{\gamma-1} [k^{1-\gamma}]_B^n \\ &= -nD^{\gamma-1} (n^{1-\gamma} - B^{1-\gamma}) \approx n \frac{D^{\gamma-1}}{B^{\gamma-1}}. \end{aligned}$$

As in Sec. II B 2, let  $E_G$  and  $E_{DS}$  be the number of edges in  $G(V, E)$  and the number of edges connected to DS,

respectively. In addition, let  $D$  denote the minimum degree. Then we have

$$E_G \approx \alpha n \int_D^n k k^{-\gamma} dk \approx \frac{\gamma - 1}{\gamma - 2} n \frac{D^{\gamma-1}}{D^{\gamma-2}},$$

$$E_{DS} \approx \alpha n \int_B^n k k^{-\gamma} dk \approx \frac{\gamma - 1}{\gamma - 2} n \frac{D^{\gamma-1}}{B^{\gamma-2}}.$$

The probability that an arbitrary edge is covered by DS is

$$\frac{E_{DS}}{E_G} \approx \left(\frac{D}{B}\right)^{\gamma-2}.$$

Thus, a lower bound of the probability that an arbitrary node is not covered by  $C$  or more edges is estimated as

$$1 - \sum_{k=C}^D \binom{D}{k} \left[\left(\frac{D}{B}\right)^{\gamma-2}\right]^k \left[1 - \left(\frac{D}{B}\right)^{\gamma-2}\right]^{D-k}$$

$$= \sum_{k=0}^{C-1} \binom{D}{k} \left[\left(\frac{D}{B}\right)^{\gamma-2}\right]^k \left[1 - \left(\frac{D}{B}\right)^{\gamma-2}\right]^{D-k}.$$

Since it is very difficult to consider a general pair  $(C, D)$ , we consider the case of  $C = D$ . Then this probability is simplified into

$$1 - \left(\frac{D}{B}\right)^{C(\gamma-2)}.$$

Therefore, an upper bound  $f(B)$  of the size of MDS is estimated as

$$f(B) = n \left[ \left(\frac{D}{B}\right)^{\gamma-1} + 1 - \left(\frac{D}{B}\right)^{C(\gamma-2)} \right].$$

It is to be noted that this number does not give a meaningful bound for many  $(C, \gamma)$ . For example, if  $C = 1$  and  $\gamma = 3$ ,  $f(B) = n$  holds.

By solving  $f'(B) = 0$ , we see that  $f(B)$  takes the minimum value

$$n \left\{ 1 + \left[ \frac{C(\gamma-2)}{\gamma-1} \right]^{\frac{\gamma-1}{C(2-\gamma)+(\gamma-1)}} - \left[ \frac{C(\gamma-2)}{\gamma-1} \right]^{\frac{C(\gamma-2)}{C(2-\gamma)+(\gamma-1)}} \right\} \quad (10)$$

at  $B = D \left[ \frac{\gamma-1}{C(\gamma-2)} \right]^{1/(C(2-\gamma)+(\gamma-1))}$ . It is interesting to note that this minimum value does not depend on the minimum degree  $D$ . If  $C = D = 1$  (i.e., original MDS), this minimum value is simplified into

$$n \left[ 1 + \left(\frac{\gamma-2}{\gamma-1}\right)^{\gamma-1} - \left(\frac{\gamma-2}{\gamma-1}\right)^{\gamma-2} \right].$$

It is also interesting to consider the case of  $C = 1$  and  $D = 2$ . In this case,  $f(B)$  is given by

$$f(B) = n \left\{ \left(\frac{2}{B}\right)^{\gamma-1} + \left[ 1 - \left(\frac{2}{B}\right)^{(\gamma-2)-2} \right] \right\}.$$

Although it is difficult to analytically derive its minimum, we can estimate it by numerical computation. Figure 3 compares upper bounds for  $(C, D) = (1, 1)$  and  $(C, D) = (1, 2)$  and lower bounds for  $(C, D) = (1, 1)$ . This figure shows that the upper bound becomes smaller as  $D$  increases in the case of  $C = 1$ .

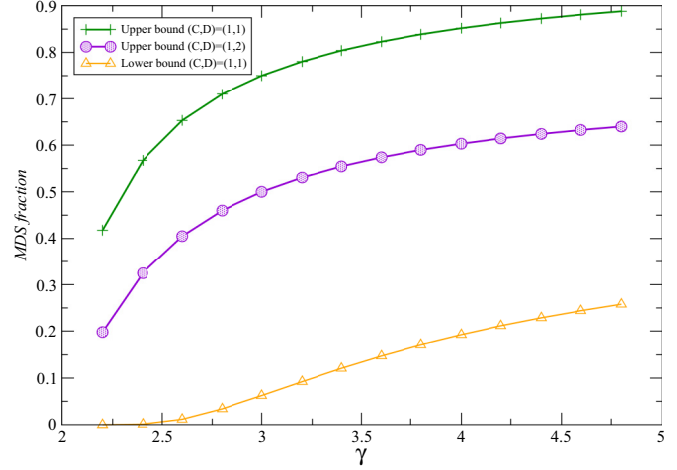


FIG. 3. (Color online) Analytical results for several configurations. Comparison of upper bounds for  $(C, D) = (1, 1), (1, 2)$  and lower bounds for  $(C, D) = (1, 1)$ .

Although, as shown in Fig. 4, the gap between the derived lower and upper bounds is large (especially for larger  $C$  and  $D$ ), these are not trivial. For example, suppose that there exist two nodes with degree no less than  $n - 1$ . Then the size of RMDS for  $C = 2$  is 2, which is  $O(1)$  [much less than  $O(n)$ ]. The  $\Omega(n)$  lower bound suggests that such a case seldom occurs in random scale-free networks.

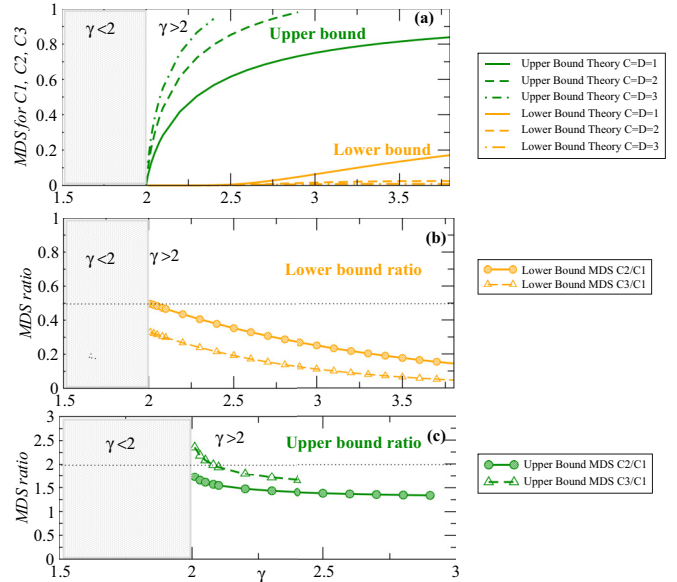


FIG. 4. (Color online) The theoretical predictions for MDS ( $C = 1$ ) and robust domination RMDS ( $C = 2, 3$ ) with  $D = C$ . (a) Lower and upper bound predictions for the fraction of nodes required to control the entire network for covers  $C = 1, 2, 3$ . (b) Ratio of lower bound MDS sizes for  $C = 2$  and  $C = 1$  (circles) and  $C = 3$  and  $C = 1$  (triangles). (c) Ratio of upper bound MDS sizes for  $C = 2$  and  $C = 1$  (circles) and  $C = 3$  and  $C = 1$  (triangles).

### III. COMPUTATION OF ROBUST DOMINATION

#### A. ILP formulation for MDS in unipartite networks

Let  $G(V, E)$  be an undirected graph, where  $V$  and  $E$  are sets of nodes and edges, respectively. We begin with the ILP formulation for computation of an MDS [26]. From  $G(V, E)$ , we construct the following ILP instance:

$$\begin{aligned} & \text{minimize } \sum_{v \in V} x_v, \\ & \text{subject to } x_u + \sum_{\{u,v\} \in E} x_v \geq 1 \text{ for all } u \in V, \\ & x_v \in \{0, 1\} \text{ for all } v \in V. \end{aligned}$$

---


$$\begin{aligned} & \text{minimize } \sum_{v \in V} x_v, \\ & \text{subject to } 2x_u + \sum_{\{v,u\} \in E} x_v \geq 2 \text{ for all } u \in V \text{ such that } \deg(u) > 1, \\ & x_u + \sum_{\{v,u\} \in E} x_v \geq 1 \text{ for all } u \in V \text{ such that } \deg(u) = 1, \\ & x_v \in \{0, 1\} \text{ for all } v \in V. \end{aligned}$$

However, the GNU Linear Programming Solver (glpsol) executable could not solve this problem in reasonable CPU time. So we strengthen the condition so each node with degree greater than 2 is covered by at least two nodes in MDS even if the node belongs to MDS. Then the resulting IP becomes as follows:

$$\begin{aligned} & \text{minimize } \sum_{v \in V} x_v, \\ & \text{subject to } x_u + \sum_{\{v,u\} \in E} x_v \geq 2 \text{ for all } u \in V \text{ such that } \deg(u) > 1, \\ & x_u + \sum_{\{v,u\} \in E} x_v \geq 1 \text{ for all } u \in V \text{ such that } \deg(u) = 1, \\ & x_v \in \{0, 1\} \text{ for all } v \in V. \end{aligned}$$

It is to be noted that the solution obtained by the above ILP also satisfies the conditions of the original formulation. Therefore, the solution obtained by use of this ILP also gives a robust dominating set although it is not necessarily minimum. We can also consider a variant of MDS in which weight  $w(u, v)$  is assigned for each edge and each node  $u$  must be covered by edges with total weight  $W_u$ . Then this variant can be formulated as

$$\begin{aligned} & \text{minimize } \sum_{v \in V} x_v, \\ & \text{subject to } w(u, u)x_u + \sum_{\{v,u\} \in E} (w(u, v)x_v) \geq W_u \text{ for all } u \in V \text{ such that } \deg(u) > 0, \\ & x_v \in \{0, 1\} \text{ for all } v \in V. \end{aligned}$$

#### C. Implementation of the ILP problems

For the MDS ( $C = 1$ ) and RMDS ( $C = 2$ ) configurations computed in real-world and simulated networks, the optimal solution was calculated using the glpsol solver [49]. The GNU Linear Programming Kit (GLPK) supplies a software package intended for solving large-scale linear programming, mixed integer programming, and other related problems. In

Then the set  $\{v|x_v = 1\}$  clearly gives an MDS. It is known that, in contrast to the bipartite matching [5], the MDS problem is NP-hard. Therefore, it is reasonable to use ILP.

#### B. ILP formulation for robust domination

Suppose that each node must be covered twice except the degree 1 and 0 nodes. Then we can formulate that this *robust dominating set* problem for  $C = 2$  [i.e., each node (with degree greater than 2) is either in MDS or is covered by at least two nodes in MDS, where each node with degree 1 is either in MDS or is covered by at least one node in MDS] as follows. Because it is impossible to cover each degree 1 node by two edges, we have introduced this exceptional handling of degree 1 nodes. However, if the minimum degree is 2 or more, we need not consider this exceptional case,

---

our problem, after translating the mathematical problem into an ILP problem, the input model is solved using the glpsol executable.

For the probabilistic MDS (PMDS), to be shown later, the optimal solution for the ILP formulation was calculated using the IBM ILOG CPLEX OPTIMIZER STUDIO, ver. 12.02. As the GLPK, it is a software package that allows us to



solve large-scale mathematical optimization problems. The computation of the PMDS is more intensive than that of MDS and RMDS, therefore we used CPLEX because it performed faster than GLPK to find the optimal solution.

#### D. Generation of unipartite scale-free networks

We employ the following algorithm to construct unipartite scale-free networks of size  $n$ , in which the degree distribution of  $V$  (a set of nodes) follows  $P(k) \propto k^{-\gamma}$  under the constraint that the minimum and maximum degrees are  $D$  and  $n$ , respectively.

For given  $n, \gamma, D$  we generate a random unipartite network in the following way:

(1) For each node  $v \in V$ , generate half edges  $e_i = (v, u_i)$  ( $u_i$  is a virtual node) according to the degree distribution  $\alpha_1 k^{-\gamma}$  under the constraint of the minimum degree  $D$  and the maximum degree  $n$ , where  $\alpha$  is selected so the number of nodes in  $V$  is almost  $n$ .

(2) Repeat the following until there are almost no remaining half edges: randomly select nonconnected  $e_i = (v, u_i)$  and  $e_j = (v', u_j)$  such that  $v \neq v'$  and then connect  $v$  and  $v'$ .

The probabilistic MDS, to be introduced later, was computed using generated samples of synthetic scale-free networks with a variety of scaling exponent  $\gamma$  and average degree  $\langle k \rangle$  values using the Havel-Hakimi algorithm with random (Monte Carlo) edge swaps (HMC) [50].

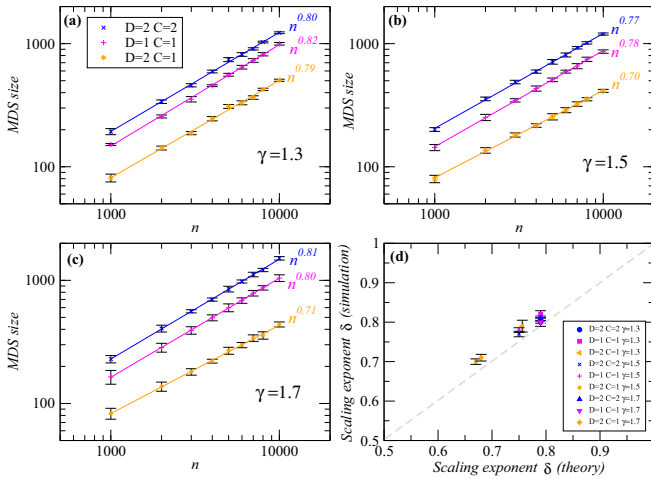


FIG. 5. (Color online) Computer simulations for MDS size for  $\gamma < 2$ . The MDS size calculated in computer-generated scale-free networks for  $\gamma = 1.3, 1.5$ , and  $1.7$  are shown in (a), (b), and (c), respectively. The configurations for minimum degree  $D$  and cover  $C$  are shown in the figure legend. The lines display a scaling law of  $n^\delta$ . The precise values for  $\delta$  are, from top to bottom, (a)  $0.804 \pm 0.011, 0.8234 \pm 0.009, 0.791 \pm 0.015$ ; (b)  $0.771 \pm 0.007, 0.781 \pm 0.006, 0.704 \pm 0.007$ ; and (c)  $0.810 \pm 0.006, 0.804 \pm 0.003, 0.714 \pm 0.008$ . As predicted by the theory, configurations  $D = 1, C = 1$  (blue) and  $D = 2, C = 2$  (red) exhibit very similar scaling exponents. Note that  $D = 2$  significantly decreases the MDS size. (d) The scaling exponent  $\delta$  predicted by theory compared with the scaling exponent observed in computer simulations for each  $D$  and  $C$  configurations. The results were averaged over 10 realizations. The error bars (s.e.m.) are shown in the figure. The correlation coefficient  $r$  is above 0.999 in all cases.

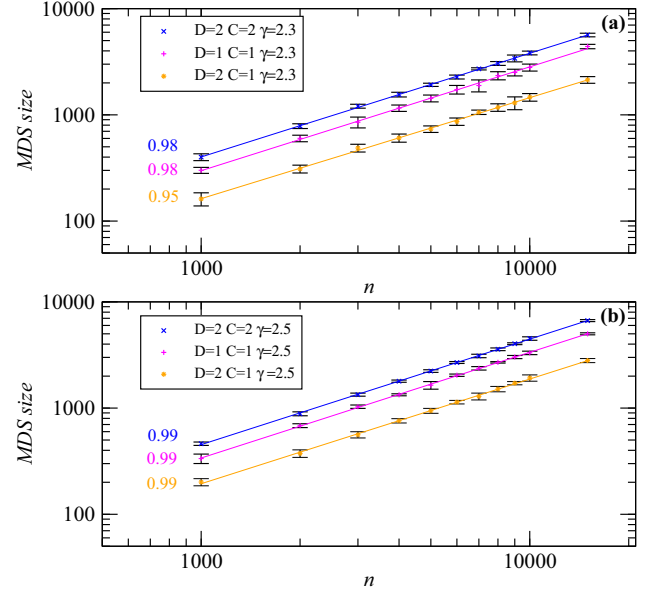


FIG. 6. (Color online) The MDS size calculated in computer generated scale-free networks for  $\gamma = 2.3, 2.5$ . Configurations for minimum degree  $D$  and cover  $C$  are shown in figure legend. The lines show a scaling law as  $n^\delta$ . The precise values for  $\delta$  are from up to down (a)  $0.976 \pm 0.005, 0.978 \pm 0.011, 0.951 \pm 0.010$  and (b)  $0.993 \pm 0.005, 0.995 \pm 0.005, 0.986 \pm 0.010$ . All three configurations show very similar scaling exponents close to 1, as predicted by theory. Note that  $D = 2$  significantly decreases the MDS size. The error bars (s.e.m.) are shown in the figure. The correlation coefficient  $r$  is above 0.999 in all cases.

#### E. Computer simulations for RMDS

To confirm the theoretical predictions shown above, we constructed artificial scale-free networks with a variety of degree exponents  $\gamma$  and minimum degree  $D = 1$  and  $D = 2$ . An ensemble of scale-free networks was constructed for each network size up to 10 000 nodes, and the mean value together with standard error of the mean (s.e.m.) for MDS size with  $C = 1$  and  $C = 2$  were computed. For  $\gamma < 2$ , the theoretical results predict the same order of MDS size (the same exponent  $\delta$  in the scaling function  $n^\delta$ ) for configurations  $(D = 2, C = 2)$  and  $(D = 1, C = 1)$  [see Eq. (6)]. In contrast, the results predict a different scaling for the configuration  $(D = 2, C = 1)$ , as shown by Eq. (3). Figure 5 presents the simulation results for  $\gamma < 2$ , which agree with the analytical predictions.

For  $\gamma > 2$ , the analytical computations predict the same scaling functional form  $n^\delta$  with  $\delta = 1$  for all the three configurations. The simulation results agree with this prediction with high accuracy (see Figs. 6 and 7).

#### F. Robust control of real-world networks

We used the concepts and mathematical tools presented above to investigate the robust control of several real networks. The experimental data analysis includes undirected, directed, and bipartite networks from biological and sociotechnical systems (see Tables I, II, and III). We first present the results for undirected networks and show that the MDS density for

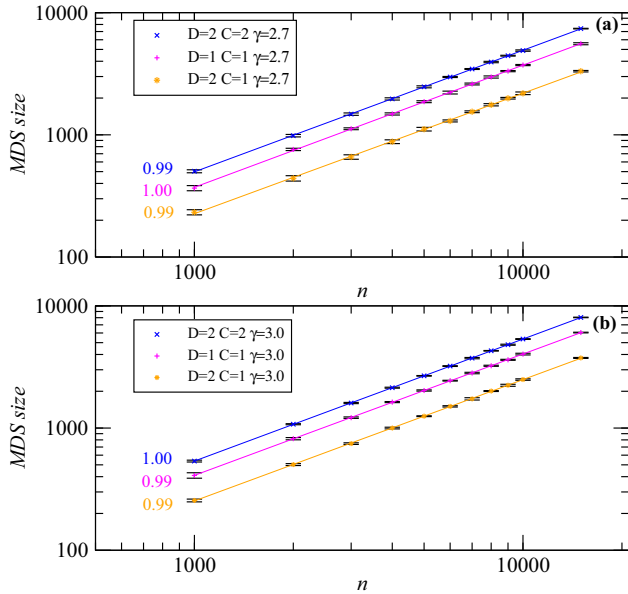


FIG. 7. (Color online) The MDS size calculated in computer generated scale-free networks for  $\gamma = 2.7, 3.0$ . Configurations for minimum degree  $D$  and cover  $C$  are shown in figure legend. The lines show a scaling law as  $n^\delta$ . The precise values for  $\delta$  are from up to down (a)  $0.994 \pm 0.002$ ,  $0.999 \pm 0.004$ ,  $0.986 \pm 0.006$  and (b)  $0.998 \pm 0.008$ ,  $0.992 \pm 0.011$ ,  $0.993 \pm 0.022$ . All three configurations show very similar scaling exponents close to 1, as predicted by theory. Note that  $D = 2$  significantly decreases the MDS size. The error bars (s.e.m.) are shown in the figure. The correlation coefficient  $r$  is above 0.999 in all cases.

$C = 1$  increases with increasing  $\gamma$ . The computation of the robust MDS density ( $C = 2$ ) exhibits a similar dependency, as predicted by Eq. (9) [Figs. 8(a) and 8(b)]. Interestingly, the MDS ratio for  $C = 2$  and  $C = 1$  differs, on average, by a

factor of 2 or less [Fig. 9(d)], in agreement with the theoretical predictions shown in Fig. 4(c). When overlapping the real data and the predictions from Eqs. (9) and (10) for the lower and upper bounds, respectively, for networks with  $\gamma > 2$ , we see that the real data are always within the theoretical boundaries (Fig. 10).

The MDS size for both  $C = 1$  and  $C = 2$  scales linearly with  $n$  [see Fig. 9(a)], which is in agreement with the theoretical predictions shown in Eqs. (9) and (10) for  $\gamma > 2$  and the computer simulations (see Figs. 6 and 7). Note that Fig. 8 displays the MDS fraction, and Fig. 9(a) represents the MDS size.

To investigate the influence of the frequency of nodes with degree 1 and 2 [ $P(1) = n(1)/n$  and  $P(2) = n(2)/n$ ] on robust control, we computed MDS1 and MDS2 versus  $P(1)$  and  $P(2)$ . We then calculated the ratio of MDS2:MDS1 versus  $P(1)$  and  $P(2)$ . The results indicate that a small  $P(1)$  and large  $P(2)$  tend to be associated with a small MDS density (see Fig. 11). The ratio of MDS2:MDS1 is less than 2 in most cases.

The analysis for directed networks included an Internet peer-to-peer (P2P) network, the transcriptional regulatory network for *Escherichia coli* from the Regulon database, a set of food webs from different ecosystems, U.S. political blogs, and the chemical synapse network for *C. elegans*. The results demonstrate that MDS1 and MDS2 densities increase with increasing  $\gamma_{\text{out}}$  [Figs. 12(a) and 12(b)] and  $\gamma_{\text{in}}$  [Figs. 13(a) and 13(b)], which is in agreement with the dependence found for undirected networks. In addition, the MDS sizes for  $C = 1$  and  $C = 2$  scale linearly with  $n$  [Figs. 9(b) and 9(c)]. Moreover, as in the undirected case, the MDS ratio between  $C = 1$  and  $C = 2$  is almost always less than 2, with only one exception [see Figs. 9(d)–9(f)]. Moreover, less than 50% of nodes are needed to control the network in both the typical ( $C = 1$ ) and robust ( $C = 2$ ) control configurations [see Figs. 12(c), 12(d), 13(c), and 13(d)].

TABLE I. The real undirected networks analysed in this work. Type, name, and description of each undirected network. We used the discrete maximum-likelihood fitting method to estimate the degree exponent  $\gamma$  from the cumulative degree distribution of each network [56,57]. The standard error of  $\gamma$  is derived from the width of the maximum likelihood. The same method was used to estimate the degree exponent in the directed and bipartite networks shown in Tables II and III.

Type	Name	Description
Protein	PPI network DIPS (6 org.) [58]	Protein networks for 6 organisms from DIPS.
	PPI Human HPRD [59]	Protein network of <i>H. sapiens</i> from HPRD.
	PPI Yeast BioGrid [60]	Protein network of <i>S. cerevisiae</i> from BioGrid.
Transportation	U.S. airports [61]	The largest U.S. airports connected by flights.
Collaboration	Hep-Th [62]	The High Energy Physics-Theory collaboration.
	Gr-QC [62]	The Quantum Cosmology research collaboration.
Communication	Email [63]	Email network in a university.
Languages	Japanese [64]	The connectivity of words in Japanese.
	Spanish [64]	The connectivity of words in Spanish.
Neuronal	Neuronal junction [33]	The electric junction network of <i>C. elegans</i> .
Intra-org.	Sawmill [65]	A communication network within a small enterprise.
Information	Wiki [66]	Linked information.
Recommendation	U.S. politics books [67]	U.S. politics books copurchased by the same buyers.

TABLE II. The real directed networks analyzed in this work. Type, name, and description of each directed network. The networks whose degree distribution follows a power law for in-degree or out-degree are indicated by I or O, respectively.

Type	Name	Description
Internet	Internet P2P [68]	Gnutella peer to peer network from August 5, 2002.
	Internet P2P [68]	Gnutella peer to peer network from August 6, 2002.
	Internet P2P [68]	Gnutella peer to peer network from August 8, 2002.
	Internet P2P [68]	Gnutella peer to peer network from August 9, 2002.
Gene regulation	Transcriptional network (O) [69]	Transcription regulatory network of <i>E. coli</i> .
Food web	Cheslower (I) [70]	Lower Chesapeake Bay in Summer food web.
	Chesapeake (I) [70]	Chesapeake Bay Mesohaline food web.
	Everglade [70]	Everglades Graminoid Marshes food web.
	Florida (O) [70]	Florida Bay Trophic food web.
	Michigan (I) [70]	Lake Michigan food web.
	St. Marks [70]	St. Marks River (Florida) flow network.
Political	Mondego (O) [70]	Mondego Estuary - Zostrea site.
	Political blogs [71]	Blog network related to politics.
Neuronal	Chemical Synapse [33]	The chemical synapse network of <i>C. elegans</i> .

#### IV. ANALYSIS ON ROBUST DOMINATION IN BIPARTITE NETWORKS

##### A. Computation of MDS in bipartite networks

We define a bipartite graph as  $G(V_{\top}, V_{\perp}; E)$ , where  $V_{\top}$  is a set of top nodes,  $V_{\perp}$  is a set of bottom nodes, and  $E$  is a set of edges ( $E \subseteq V_{\top} \times V_{\perp}$ ). In our analysis, the directions of the edges are considered from  $V_{\top}$  to  $V_{\perp}$ . Therefore, the set of driver nodes will be a subset of  $V_{\top}$ , where nodes in  $V_{\top}$  need not be covered.

The computation of an MDS of a bipartite network is equivalent to the computation of a minimum set cover. Although it is an NP-hard problem, we have verified that the optimal solution is obtained in networks with power-law distributions of up to approximately 110 000 nodes within a few seconds. The computation was formalised as the following ILP problem:

$$\begin{aligned}
 & \text{minimize} && \sum_{v \in V_{\top}} x_v, \\
 & \text{subject to} && \sum_{\{v,u\} \in E} x_v \geq 1 \quad \text{for all } u \in V_{\perp}, \\
 & && x_v \in \{0,1\} \quad \text{for all } v \in V_{\top}. \quad (11)
 \end{aligned}$$

##### B. Computation of RMDS in bipartite networks

The above-mentioned ILP can be extended for computation of an RMDS in bipartite networks. It is formalized as

$$\begin{aligned}
 & \text{minimize} && \sum_{v \in V_{\top}} x_v, \\
 & \text{subject to} && \sum_{\{v,u\} \in E} x_v \geq 1 \quad \text{for all } u \in V_{\perp} \text{ such that } \deg(u) = 1, \\
 & && \sum_{\{v,u\} \in E} x_v \geq 2 \quad \text{for all } u \in V_{\perp} \text{ such that } \deg(u) > 1,
 \end{aligned}$$

where  $\deg(u)$  indicates the degree of node  $u$ . It should be noted that for any node  $u \in V_{\perp}$  with degree 1, it is not possible to cover  $u$  twice and thus we must relax the condition for these nodes.

##### C. Generation of bipartite scale-free networks

We employ the following algorithm to construct bipartite scale-free networks, in which the degree distributions of  $V_{\top}$  and  $V_{\perp}$  follow  $P_{\top}(k) \propto k^{-\gamma_1}$  and  $P_{\perp}(k) \propto k^{-\gamma_2}$ , respectively. Here we consider  $n_1 = |V_{\top}|$  and  $n_2 = |V_{\perp}|$ . The maximum degree for the nodes in  $V_{\top}$  corresponds to  $n_1$ .

TABLE III. The real bipartite networks analyzed in this work. Type, name, and description of each bipartite network.

Type	Name	Description
Social	Firms-World Cities [72]	Services of firms across cities.
	Facebook Forum UCA [73]	Facebook users linked to topics.
	Davis's Southern Women Club [74]	Attendance at social events by women.
	Cond-Mat Sci. Coll. [75]	Collaboration of scientists and papers.
	Graph Book Bibliography [76]	Author-by-paper network.
Biological	The Dutch Elite [77]	Individuals connected to administrative bodies.
	Drugs-Targets [78]	Drugs binding to protein targets.
	Transcriptional network (Yeast) [79]	Transcription regulatory network of <i>S. cerevisiae</i> .
	ncRNA-protein network (human) [80]	Interactions between ncRNAs and proteins in <i>H. sapiens</i> .
	ncRNA-protein (6 organisms) [80]	All ncRNA-protein interactions of six organisms.

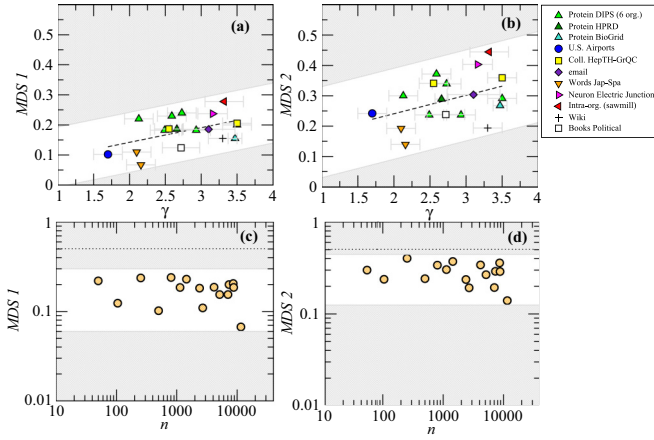


FIG. 8. (Color online) The MDS fraction in real undirected networks. The MDS fraction as a function of the degree exponent  $\gamma$  for real undirected networks  $C = 1$  in (a) and  $C = 2$  in (b). The MDS fraction as a function of the network size for the same real undirected networks for  $C = 1$  in (c) and  $C = 2$  in (d). Classification of nodes into network classes is only displayed for (a) and (b) (see figure legend). Highlighted regions are visual guidance of the observed tendency.

Then, for given  $n_1, \gamma_1, \gamma_2$ , we generate a random bipartite network in the following way:

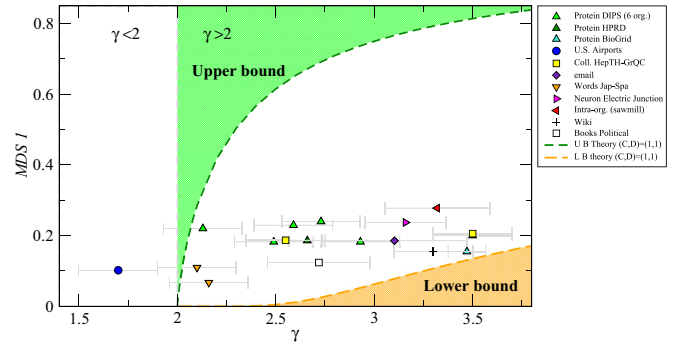


FIG. 10. (Color online) Comparison between theoretical predicted bounds and MDS in real undirected networks. The overlap between the MDS fraction with  $C = 1$  (MDS1) for undirected networks and the predictions of Eqs. (6) and (7) for the lower and upper bounds of networks with  $\gamma > 2$ . The real data are always within the theoretical boundaries.

(1) For each node  $v \in V_T$ , generate half edges  $e_i = (v, u_i)$  ( $u_i$  is a virtual node) according to the degree distribution  $\alpha_1 k^{-\gamma_1}$ , where  $\alpha_1$  is selected so the number of nodes in  $V_T$  is almost  $n_1$ .

(2) For each node  $w \in V_\perp$ , generate half edges  $e'_j = (u'_j, w)$  ( $u'_j$  is a virtual node) according to the degree distribution  $\alpha_2 k^{-\gamma_2}$ , where  $\alpha_2$  is selected so the number of  $e'_j$ 's is equal to the number of  $e_i$ 's.

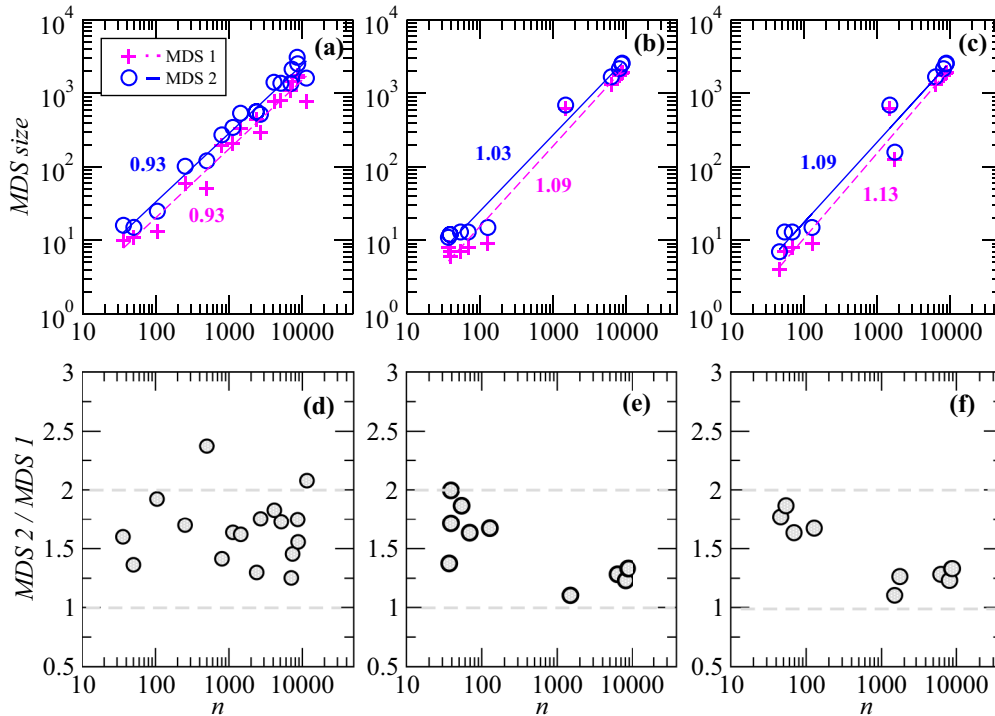


FIG. 9. (Color online) The MDS size scaling in real undirected and directed networks. The MDS size as a function of the network size  $n$  for real (a) undirected and directed networks with (b) incoming and (c) outgoing links. The results display a scaling function of  $n^\delta$  with  $\delta$  close to 1, as predicted by the theoretical analysis. The precise values of the scaling exponent  $\delta$  for  $C = 1$  (plus) and  $C = 2$  (circle) are (a)  $0.934 \pm 0.046$  ( $r = 0.981$ ),  $0.931 \pm 0.036$  ( $r = 0.989$ ); (b)  $1.094 \pm 0.053$  ( $r = 0.990$ ),  $1.034 \pm 0.044$  ( $r = 0.992$ ); and (c)  $1.134 \pm 0.071$  ( $r = 0.984$ ),  $1.092 \pm 0.067$  ( $r = 0.985$ ), respectively. The correlation coefficient  $r$  is indicated between parentheses. The ratio between the MDS sizes (MDS2:MDS1) computed with covers  $C = 2$  and  $C = 1$  is shown in parts (d)–(f) for the same real networks. The results show that the ratio is almost always lower than 2. See Tables I and II for real data details.

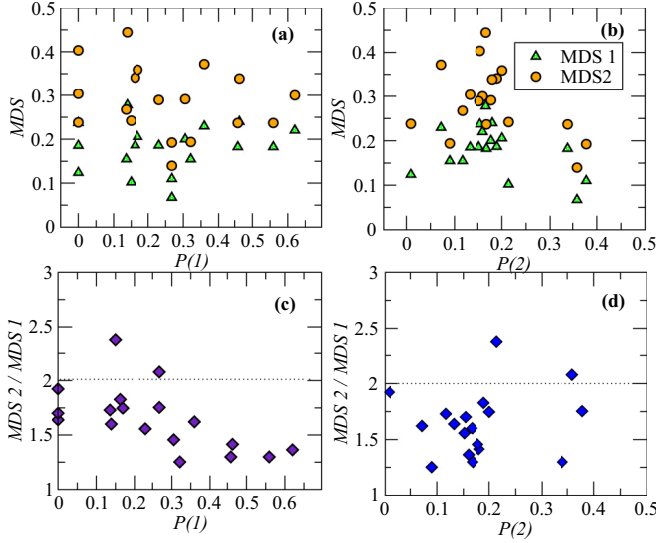


FIG. 11. (Color online) The dependence of the MDS fraction on the degree distribution for real undirected networks. The MDS fraction with  $C = 1$  (MDS1) (a) and  $C = 2$  (MDS2) (b) as a function of the degree probability with degree 1 [ $P(1) = n(1)/n$ ] and 2 [ $P(2) = n(2)/n$ ]. The ratio of MDS2:MDS1 versus (c)  $P(1)$  and (d)  $P(2)$ , respectively. The ratio is below 2 in most cases. The results demonstrate that small  $P(1)$  and large  $P(2)$  tend to be associated with a small MDS fraction.

(3) Randomly connect  $e_i$ 's and  $e_j$ 's in a one-to-one manner. It is to be noted that  $n_2$  (the number of nodes of  $V_\perp$ ) is determined automatically in step (2) to satisfy the condition on edge numbers.

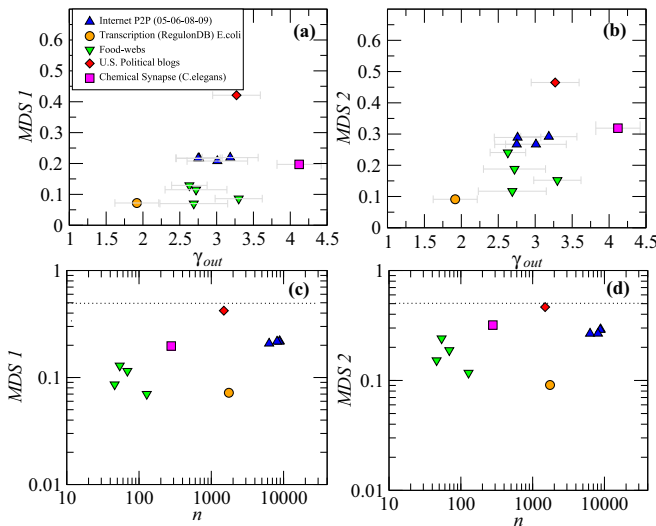


FIG. 12. (Color online) The MDS fraction as a function of the degree exponent  $\gamma_{out}$  for real directed networks  $C = 1$  in (a) and  $C = 2$  in (b). The MDS fraction as a function of the network size for the same real directed networks for  $C = 1$  in (c) and  $C = 2$  in (d). In all cases the MDS fraction is lower than 0.5 (dotted line). Note that four of seven food webs show power-law behavior for outgoing degrees. Each color in circles corresponds to a network type as shown in the legend.

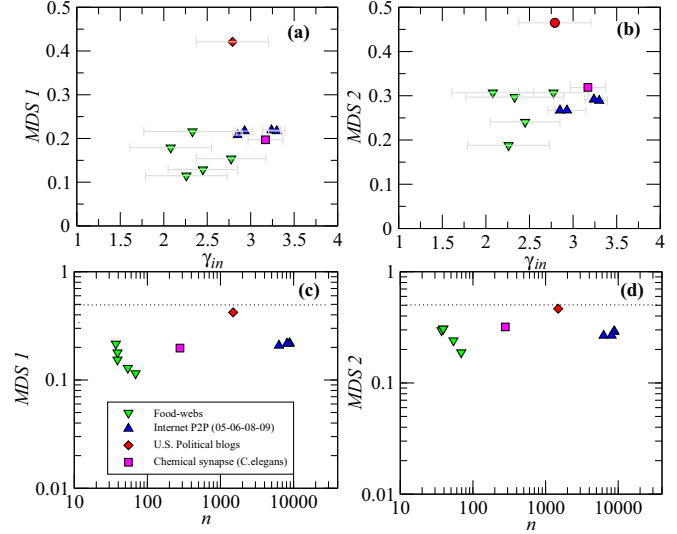


FIG. 13. (Color online) The MDS fraction as a function of the degree exponent  $\gamma_{in}$  for real directed networks  $C = 1$  in (a) and  $C = 2$  in (b). The MDS fraction as a function of the network size for the same real directed networks for  $C = 1$  in (c) and  $C = 2$  in (d). In all cases the MDS fraction is lower than 0.5 (dotted line). Note that five of seven food webs show power-law behavior for incoming degrees. The data for the transcriptional regulatory network is absent because it does not follow a power law for incoming degrees. Each color in circles corresponds to a network type as shown in the legend.

**D. Data analysis of real-world bipartite networks**

We collected a set of 10 real-world bipartite networks corresponding to sociotechnical [Fig. 14(a)] and biological systems [Fig. 14(b)]. We then formalized and computed the MDS for the  $C = 1$  and  $C = 2$  configurations. Although the MDS with  $C = 2$  is always larger than the MDS with  $C = 1$ , the difference is proportionally very small in most cases. Figures 14(a) and 14(b) also illustrates that biological systems tend to require a larger MDS size than sociotechnical systems. The computer simulation of ensembles of bipartite scale-free networks with a variety of degree exponents also demonstrates that  $C = 1, D = 1$  and  $C = 2, D = 2$  can control the network with a similar fraction of nodes. Therefore, robustly controlling a bipartite network requires a similar fraction of nodes as the typical, nonrobust system [see Figs. 14(c) and 14(d)].

**V. THEORETICAL ANALYSIS FOR THE PROBABILISTIC DOMINATION (PMDS)**

In some real networks, each link has a probability of failing, which leads to the probabilistic concept of robust control (PMDS). For example, experimental analyses on neural networks have confirmed the unreliability of central synaptic transmission in rat brains [32]. The mean transmission failure probability was found to be  $p = 0.71$ , with a range of 0.3 to 0.95 ( $w = 0.24$ ). In this work, we used the well-studied *C. elegans* neural network to investigate probabilistic robust control. To investigate this type of systems from a theoretical perspective using the robust MDS approach, we assume that each edge  $(v,u)$  has the probability of failure



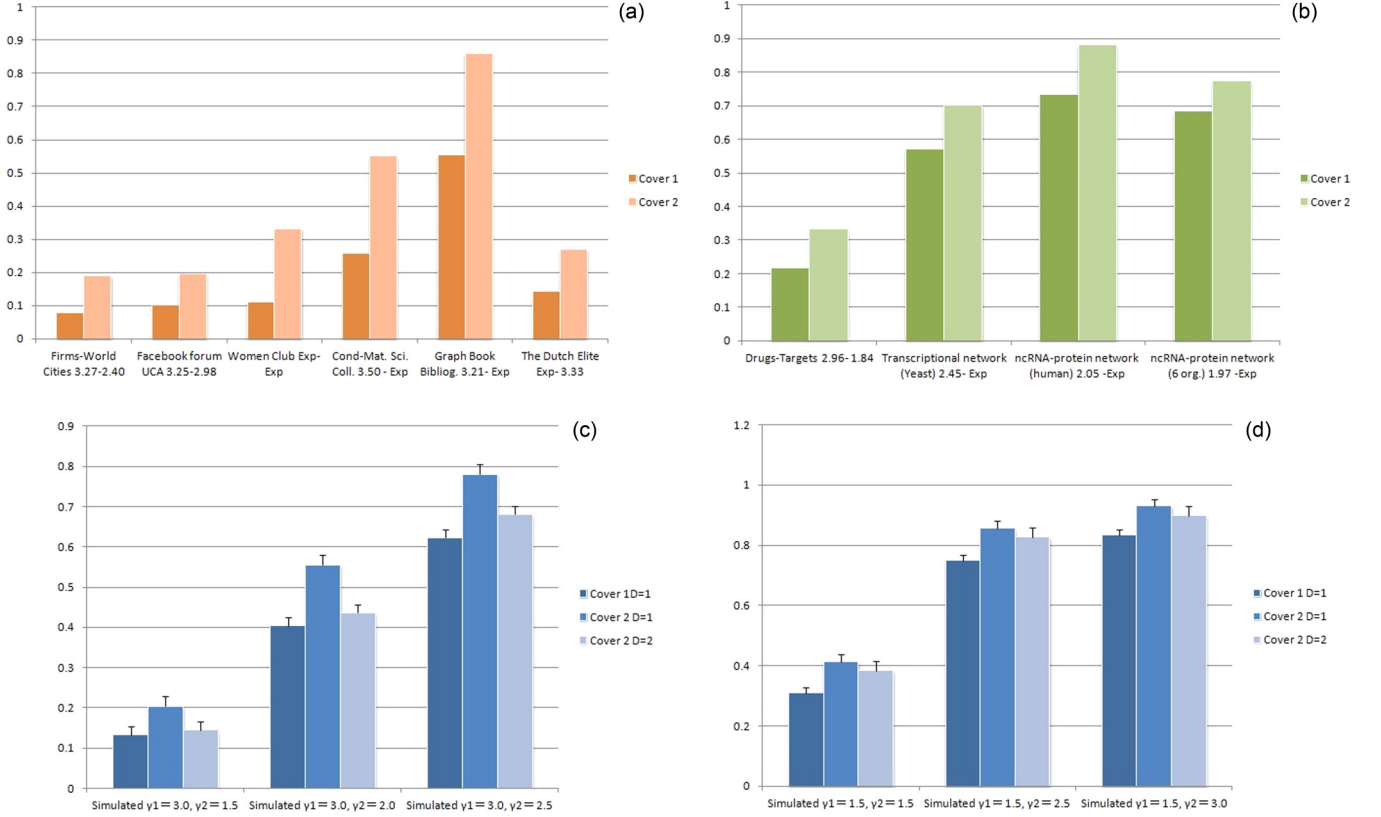


FIG. 14. (Color online) The MDS size computed on real-world bipartite networks. (a) Social networks and (b) biological networks computed for  $C = 1$  and  $C = 2$  as shown in the legend. Note that social networks tend to be controlled with a smaller fraction of nodes compared to biological networks. [(c)–(d)] Computation of the MDS in synthetic bipartite networks generated with the model shown in Sec. IV C with several values of degree exponents  $\gamma_1$  and  $\gamma_2$  and three configurations ( $C = 1, D = 1$ ), ( $C = 2, D = 1$ ), and ( $C = 2, D = 2$ ). Note that the minimum degree  $D = 2$  almost completely compensates the increasing of the MDS size from the robust control  $C = 2$ .

$P_{v,u}$  [see Fig. 1(e)]. We require that each node is covered by multiple nodes in an MDS so the probability that at least one edge is active is at least  $\theta$ . Let  $S$  be a DS. Then  $S$  must satisfy

$$(\forall u)(1 - \prod_{v \in S} P_{v,u}) \geq \theta. \quad (12)$$

As we will show later, this problem can be also formalized and solved using ILP.

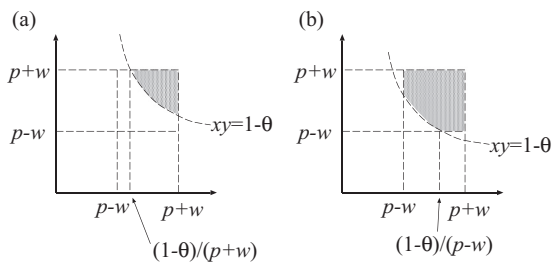


FIG. 15. Graphical representation of the case  $D = 2$  and  $w > 0$ . In order to estimate the fraction of degree-2 nodes to be added to a DS, it is enough to consider the regions (A) and (B) for the cases of  $(p + w)(p - w) < 1 - \theta < (p + w)(p + w)$  and  $(p - w)(p - w) < 1 - \theta < (p + w)(p - w)$ , respectively.

#### A. The case of $D = 1$

First, we consider the case of  $D = 1$  (i.e., the minimum degree is 1) and  $w = 0$ . Let DS be an MDS for  $G(V, E)$  for the nonprobabilistic version. Let  $U$  be a set of degree 1 nodes, each of which does not belong to DS but is dominated by a node in DS. Let  $\{u, v\}$  be the only edge connecting to  $v \in U$ . We can observe:

if  $1 - P_{u,v} < \theta$ ,  $v$  must be covered by itself.

Therefore, all nodes in  $U$  should be added to DS (in a probabilistic version) when  $\theta > 1 - P_{u,v}$ . Therefore, it is expected that the MDS size increases approximately from  $|\text{DS}|$  to  $|\text{DS}| + |U|$  at around  $\theta = 1 - p$ . For example, consider the case of  $p = 0.71$ . Then there should be great increase of the MDS size at  $\theta = 1 - 0.71 = 0.29$ . It shows good agreement with the simulation result (see Figs. 16 and 17).

#### B. The case of $D = 2$

We can extend the above analysis to the case of  $D = 2$  (i.e., the minimum degree is 2) and  $w = 0$ . Let DS be an MDS for  $G(V, E)$  for the nonprobabilistic version. In this case, we consider two types of nodes of degree 2:

- (a)  $v$  has one edge connecting to a node in DS,
- (b)  $v$  has two edges connecting to nodes in DS,

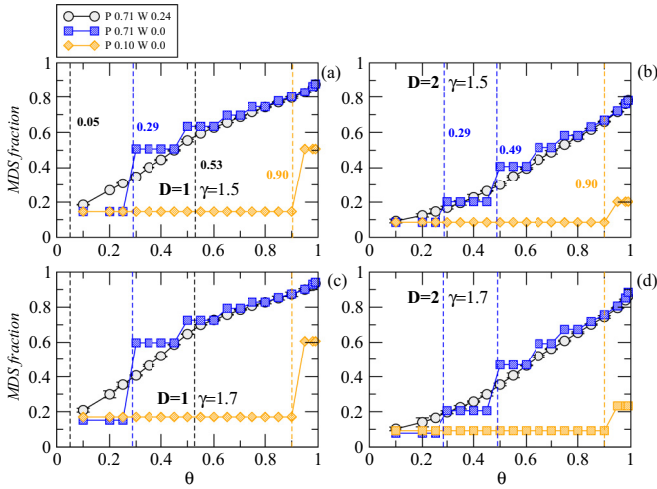


FIG. 16. (Color online) Probabilistic MDS simulation results. The probabilistic MDS fraction in simulated scale-free networks with (a)  $D = 1$  and (b)  $D = 2$  for  $\gamma = 1.5$  and  $D = 2$  (c)  $D = 1$  and (d)  $D = 2$  for  $\gamma = 1.7$ . The predicted theoretical thresholds (dashed lines) that significantly changes the MDS size are in fair agreement with observed results in computer simulations. The configurations for the probability of link failure  $P$  and the variability change  $w$  of the failure probability [ $p - w$ ,  $p + w$ ] are shown in the figure legend.

where each node does not belong to DS but is dominated by a node in DS. Let  $U_1$  and  $U_2$  be the sets of type (a) and type (b) nodes, respectively. Then nodes in  $U_1$  should be added to DS if  $1 - p < \theta$ . On the other hand, nodes in  $U_2$  should be added to DS if  $1 - p^2 < \theta$ . Therefore, it is expected that the MDS size increases approximately from  $|\text{DS}|$  to  $|\text{DS}| + |U_1|$  at around  $\theta = 1 - p$  and from  $|\text{DS}| + |U_1|$  to  $|\text{DS}| + |U_1| + |U_2|$

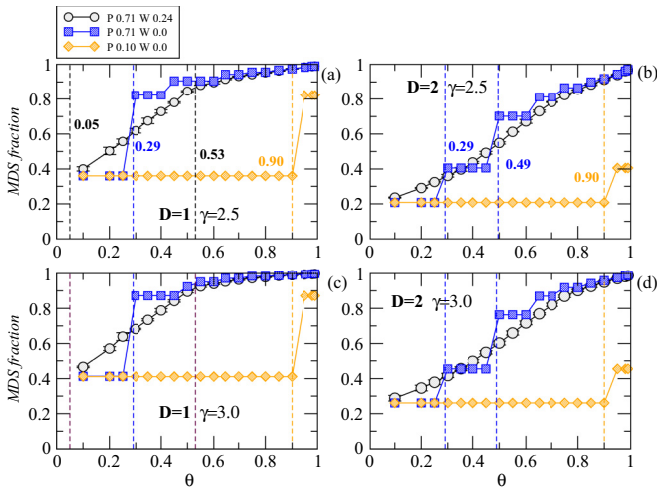


FIG. 17. (Color online) The probabilistic robust domination on synthetic scale-free networks. The probabilistic MDS fraction in simulated scale-free networks with (a)  $D = 1$  and (b)  $D = 2$  for  $\gamma = 2.5$  and (c)  $D = 1$  and (d)  $D = 2$  for  $\gamma = 3.0$ . The predicted theoretical thresholds (dashed lines) that significantly change the MDS size are in fair agreement with the observed results from the computer simulations. The configurations for the probability of link failure  $p$  and the variability change  $w$  of the failure probability [ $p - w$ ,  $p + w$ ] are shown in the figure legend.

at around  $\theta = 1 - p^2$ . In the case of  $p = 0.71$ , these two threshold values are 0.29 and 0.49, in good agreement with the simulation result.

### C. The case of $D = 1$ and $w > 0$

Next we consider the case of  $D = 1$  and  $w > 0$ . As in the above, let DS be an MDS for  $G(V, E)$  for the nonprobabilistic version, and let  $U$  be a set of degree 1 nodes each of which does not belong to DS but is dominated by a node in DS. Let  $e = \{u, v\}$  be the only edge connecting to  $v \in U$ . Let  $p + \Delta_v$  be the failure probability of this edge  $e$ , where  $-w \leq \Delta \leq w$ . We can observe:

if  $1 - (p + \Delta_v) < \theta$ ,  $v$  must be covered by itself.

We define  $U_\Delta$  by

$$U_\Delta = \{v | v \in U, p + \Delta_v > p + \Delta\}.$$

Therefore, all nodes in  $U_\Delta$  should be added to a DS (in a probabilistic version) where  $\Delta = 1 - p - \theta$ . Here the size of  $U_\Delta$  is estimated as

$$|U_\Delta| \approx \left( \frac{\Delta + w}{2w} \right) |U|,$$

where  $-w \leq \Delta \leq w$ . By replacing  $\Delta$  with  $1 - p - \theta$ , we have

$$|U_\Delta| \approx \left( \frac{1 - p - \theta + w}{2w} \right) |U|.$$

Therefore, it is expected that the MDS size is approximately given by  $|\text{DS}| + \left( \frac{1 - p - \theta + w}{2w} \right) |U|$ . It should be noted that  $\frac{1 - p - \theta + w}{2w}$  becomes 0 and 1 at  $\theta = 1 - p - w$  and  $\theta = 1 - p + w$ , respectively. In the case of  $p = 0.71$  and  $w = 0.29$ , these two threshold values are 0.05 and 0.53, in good agreement with the simulation results (see Figs. 16 and 17). This discussion can be generalized for the cases in which  $\Delta_v$  does not follow the uniform distribution.

### D. The case of $D = 2$ and $w > 0$

Finally, we consider the case of  $D = 2$  and  $w > 0$ . Let  $v$  be a degree-2 node and  $u_1$  and  $u_2$  be the neighboring nodes to  $v$ . Then we estimate the fraction of degree-2 nodes that does not satisfy

$$1 - P_{u_1, v} P_{u_2, v} \geq \theta,$$

where such a node  $v$  should be added to a DS. For that purpose, it is enough to calculate the area shown in Fig. 15.

For  $\theta$  satisfying  $(p + w)(p - w) < 1 - \theta < (p + w)(p + w)$ , we consider the region (A) whose area is given by

$$\begin{aligned} P_A &= (p + w) \left[ (p + w) - \left( \frac{1 - \theta}{p + w} \right) \right] - \int_{\frac{1 - \theta}{p + w}}^{p + w} \frac{1 - \theta}{x} dx \\ &= (p + w) \left[ (p + w) - \left( \frac{1 - \theta}{p + w} \right) \right] \\ &\quad - (1 - \theta) \left[ \ln(p + w) - \ln \left( \frac{1 - \theta}{p + w} \right) \right]. \end{aligned}$$

Therefore, in this case, the fraction is given by  $P_A/((2w)^2)$  because  $(P_{u_1,v}, P_{u_2,v})$  is uniformly distributed in the region of  $[p-w, p+w] \times [p-w, p+w]$ .

For  $\theta$  satisfying  $(p-w)(p-w) < 1-\theta < (p+w)(p-w)$ , we consider the region (B) whose area is given by

$$\begin{aligned} P_B &= (2w)^2 - \int_{p-w}^{\frac{1-\theta}{p-w}} \frac{1-\theta}{x} dx + (p-w) \\ &\quad \times \left[ \left( \frac{1-\theta}{p-w} \right) - (p-w) \right] \\ &= (2w)^2 - (1-\theta) \left[ \ln \left( \frac{1-\theta}{p-w} \right) - \ln(p-w) \right] \\ &\quad + (p-w) \left[ \left( \frac{1-\theta}{p-w} \right) - (p-w) \right]. \end{aligned}$$

Again, the fraction is given by  $P_B/((2w)^2)$ . The simulation results with 1000 nodes for  $D=1$ ,  $D=2$  with  $w > 0$  configurations are shown in Figs. 16 and 17. We also compared the theoretical results for the case  $D=2$  with those from the simulations performed on scale-free networks with  $D=2$  and  $\gamma=3$ . The plots show a similar overall tendency, although the inflection point is more noticeable in the theoretical curve (see Fig. 18). It is worth noticing that the theoretical values  $P_A$  and  $P_B$  are scaled so these take almost the same values as the simulated ones at the beginning and ending points (i.e., so the values take between 0.3 and 0.85 instead of between 0.0 and 1.0) since it is assumed in theoretical analysis that all nodes are of degree 2 and the effects of the other nodes are ignored (note also that degree 2 nodes occupy a major portion of nodes in the case of  $D=2$ ). This comparison result suggests that theoretical analysis captures some tendency even if nodes with degree more than 2 are ignored.

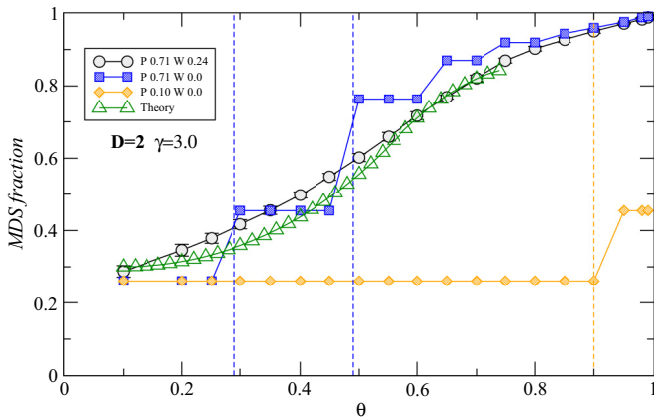


FIG. 18. (Color online) Theoretical predictions and simulations of PMDS. Comparison of the probabilistic MDS size computed in simulated scale-free networks with  $D=2$  and  $\gamma=3$  and the case of  $D=2$  predicted by theory. Theoretical values are scaled so these take almost the same values as the simulated ones at the beginning and ending points because, in the theoretical analysis, it is assumed that all nodes are of degree 2 and the effects of the other nodes are ignored.

### E. ILP formulation for probabilistic robust domination (PMDS)

We assume that each edge  $(v,u)$  has the probability of failure  $P_{v,u}$ . We want each node be covered by multiple nodes in MDS so the probability that at least one edge is active is at least  $\theta$ . Let  $S$  be a dominating set. Then we require  $S$  to satisfy

$$(\forall u)(1 - \prod_{v \in S} P_{v,u}) \geq \theta.$$

Then we have

$$1 - \prod_{v \in S} P_{v,u} \geq \theta,$$

$$\prod_{v \in S} P_{v,u} \leq 1 - \theta,$$

$$\sum_{v \in S} \ln(P_{v,u}) \leq \ln(1 - \theta),$$

$$\sum_{v \in S} -\ln(P_{v,u}) \geq -\ln(1 - \theta).$$

Then we have

$$\text{minimize } \sum_{v \in V} x_v,$$

$$\text{subject to } x_u \geq 1, \quad \text{for all } u \in V \text{ such that } \deg(u) = 0,$$

$$-\ln(1 - \theta)x_u + \sum_{\{v,u\} \in E} ((-\ln(P_{v,u}))x_v)$$

$$\geq -\ln(1 - \theta),$$

$$\text{for all } u \in V \text{ such that } \deg(u) > 0,$$

$$x_v \in \{0, 1\}, \quad \text{for all } v \in V,$$

where  $\deg(u)$  indicates the degree of node  $u$ .

### F. Probabilistic robust domination applied to the *C. elegans* neuronal network

Recent reconstructions of the *C. elegans* neural network have significantly updated the wiring diagram of the somatic nervous system. The new reconstruction includes original data from White *et al.* [51] and Hall and Russel [52] and adds new information. In particular, 3000 synaptic contacts, including gap junctions, chemical synapses, and neuromuscular junctions, were updated or added to the latest network version [33]. As a result, the large-scale structure of the network has significantly changed with respect to that of White *et al.* Here we focus on the connectivity of gap junction and chemical synapse networks of *C. elegans* neurons. The channels that provide electrical coupling between neurons are called gap junctions. In contrast, chemical synapses use neurotransmitters to link neurons. Because these network biologically differ, they are treated independently, as done in Ref. [33]. Although it might be possible that gap junctions could conduct current in only one direction, this feature has not been observed or confirmed yet in *C. elegans* [33]. Therefore, this network was considered as undirected network. The chemical synapses, in contrast, contains directionality capability, a feature that has been confirmed using micrographs [33]. The analyzed gap junction network consisted of 279 neurons and 514 gap junction connections. The giant connected component is

composed of 248 neurons and two smaller components of 2 and 3 neurons. After removing the 26 isolated neurons, we performed our analysis using 253 neurons and 514 connections. The statistical analysis revealed a power-law distribution for the degree distribution with a characteristic degree exponent of  $\gamma = 3.14$  [33]. The chemical synapse network consisted of 279 neurons and 2194 directed connections. The statistical analysis showed that the in-degree (out-degree) distribution followed a power law with degree exponent  $\gamma_{in} = 3.17$  ( $\gamma_{out} = 4.22$ ), respectively. These results contrast with analyses done using the data set from White *et al.* [51], which reported an exponential decay for the degree distribution [53].

Experimental analyses on neural networks have confirmed the unreliability of central synaptic transmission in rat brains [32]. The mean transmission failure probability was found to be  $p = 0.71$ , with a range from 0.3 to 0.95 ( $w = 0.24$ ). In this work, we used the most well-studied neural network corresponding to the *C. elegans* (chemical synapse and gap junction) to investigate probabilistic robust control. A visual representation of experimental neural gap junction (undirected) for *C. elegans* is shown in Fig. 19. A transmission failure probability distribution similar to that observed in rat brains was mapped on the links of these networks, making a fraction of them unreliable. The results of the analyses are described in Figs. 16 and 17 for computer simulations and Figs. 20–22 for real neural gap junction and chemical synapse networks and suggest that the presence of variance of the

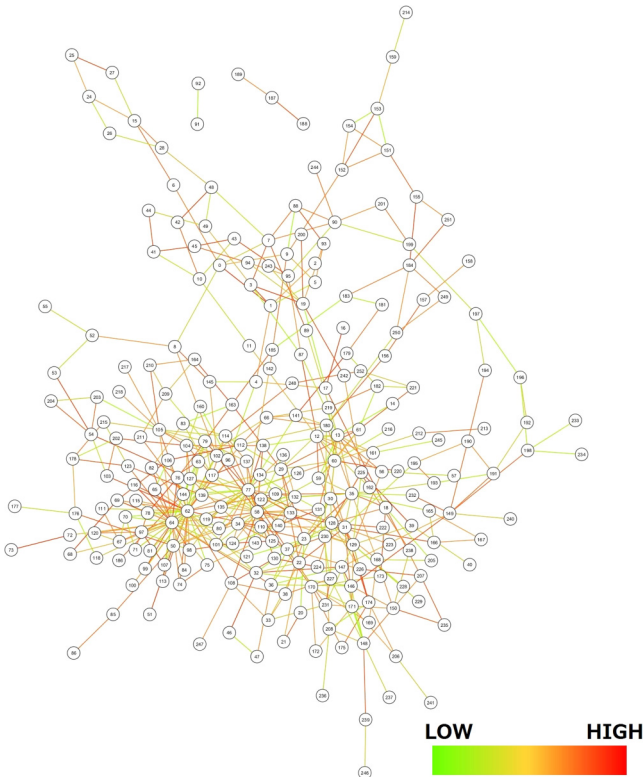


FIG. 19. (Color online) The neural gap junction network. Visualization of the experimental gap junction undirected network for *C. elegans*. A probability distribution of synapse transmission failure with a peak at  $p = 0.71$  and width of  $w = 0.24$  (0.47–0.95) is mapped onto the links.

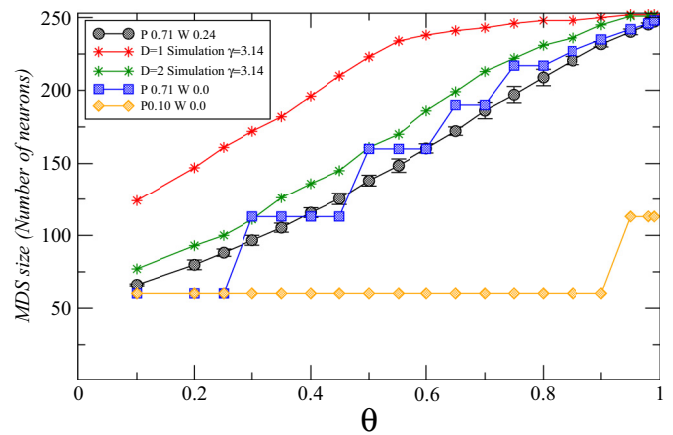


FIG. 20. (Color online) The probabilistic MDS size computed on the neural gap junction undirected network. The computation of the probabilistic MDS size using the real neural gap junction network for *C. elegans* organism where a distribution of link failure was mapped onto the network as shown in legend for three values of  $w$  (circle, square, and diamond symbols). The results of synthetic scale-free networks constructed using the model shown in Sec. III D and calculated with the same degree exponent and the number of nodes observed in the real *C. elegans* network (star symbols). When simulated networks have a minimum degree  $D = 2$ , the MDS size decreases as predicted by theory. Because in this simulation the synthetic network have different average degree, the results tend to be higher than those from the real network.

failure probability  $w$  does not significantly affect the fraction of driver nodes. In contrast, it is strongly affected by both the minimum degree  $D$  and the average failure probability  $p$ . This biological example of unreliable links suggests that theoretical results and simulations on probabilistic robust control analysis may have an impact on understanding and controlling at will real-world systems with unreliable components.

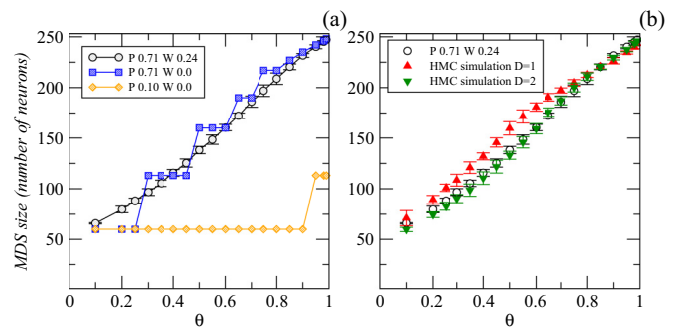


FIG. 21. (Color online) The probabilistic robust domination on a real neural gap junction network. The computation of the probabilistic MDS size using (a) the real neural gap junction network for *C. elegans* where the distribution of link failures was mapped to the network as shown in the legend. (b) The results of synthetic scale-free networks constructed using the HMC model with the same number of nodes and average degree display a similar tendency. The MDS size decreases as predicted by theory when simulated networks have a minimum degree  $D = 2$ .



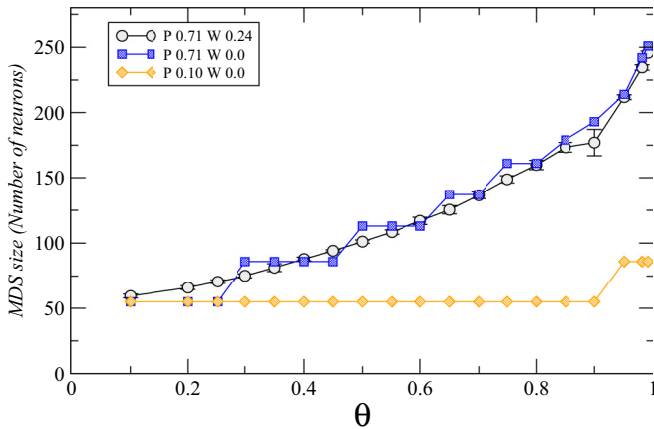


FIG. 22. (Color online) The probabilistic MDS size computed on the chemical synapse directed network. The computation of the probabilistic MDS size using the real chemical synapse directed network for *C. elegans* where a distribution of link failure was mapped onto the network as shown in legend for three values of  $w$ . Note that the distribution follows a curve function rather than a straight line as shown in the undirected case.

## VI. CONCLUSION

We have introduced the concept of structurally robust control of complex networks and have used the MDS model, which is widely applied in engineering problems, to illustrate an example of robust complex network controllability. Counterintuitively, the developed analytical tools, computer simulations and real-world network analyses demonstrate that robust control in a large network does not change the order of required driver nodes compared to a conventional system without such robust capability. When using an MDS with  $C = 1$ ,  $D = 1$ , the system can easily become uncontrollable if only one power or communication line fails during major natural disasters. In contrast, in the RMDS framework ( $C = 2$ ,  $D = 2$ ) the system remains controllable even under arbitrary single or multiple link failure. Therefore, both configurations require exactly the same order of controllers. Engineering and biological systems could benefit from these findings.

In addition, the order of the MDS changes for  $\gamma < 2$  by changing the minimum degree  $D$  (e.g., constructing real networks with degree  $D > 1$ ), unveiling another tool to decrease the number of driver nodes. Because some real networks have unreliable links, we have extended our framework to PMDS and have successfully applied the developed analytical tools to real neural networks of *C. elegans* with unreliable synaptic transmission. With the forthcoming comprehensive map of neural connections in the human brain [54,55], the presented method could offer new avenues to examine the brain's large-scale structure to address synaptic reliability and

to stimulate large fractions of the brain by interacting only with relatively few components.

It is worth remarking that our analysis on robust network control emphasizes the importance of the network topology and focuses on deriving the relative change of the required number of controllers in both robust and nonrobust configurations. Studies done in the context of fault-tolerant control often include more elaborate node dynamics to simulate synchronization of networks and agents communicating over wireless connections. Instead, our MDS model-based mathematical analysis relies on mean-field theory concepts. To our best knowledge, there is no other theoretical results for robust property of MDS size on scale-free networks in both control theory and network science fields.

The proposed concept of structurally robust control of complex networks could also be investigated using a different algorithmic framework. As discussed above, we selected the MDS model because it has already found applications in real engineering systems. However, the concept could also be mathematically formalized and implemented using, for example, the maximum matching model [5]. In this case, additional computations would be needed to investigate the order of drivers in an optimal robust control configuration; therefore, this analysis is left for future work.

In addition, the presented method can also address the simultaneous failure of multiple links. The aim of the RMDS ( $C = 2$ ) framework is to construct a system that remains controllable even if an arbitrary link is damaged. However, the developed analytical tools also allow us to design a system with a more robust configuration ( $C = 3$ ) or ( $C = 4$ ) so the network is still controllable even in case of arbitrary failure of pair or triplet of links, respectively.

The emerging picture for probabilistic failure or malfunction of transportation and transmission lines in real-world complex infrastructures, sociotechnical networks, and biological networks emphasizes the importance and role of the presented robust DS approach for controllability. The proposed framework and tools offer a new direction for understanding the linkage between controllability and robustness in complex networks, with implications from engineering to biological systems.

## ACKNOWLEDGMENTS

J.C.N. was partially supported by MEXT, Japan (Grant-in-Aid No. 25330351), and T.A. was partially supported by MEXT, Japan (Grant-in-Aid No. 26540125). This work was also partially supported by research collaboration projects by Institute for Chemical Research, Kyoto University.

- [1] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).  
 [2] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin, *Phys. Rev. Lett.* **85**, 4626 (2000).

- [3] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Phys. Rev. Lett.* **86**, 3682 (2001).  
 [4] R. Albert, H. Jeong, and A.-L. Barabási, *Nature* **406**, 378 (2000).



- [5] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, *Nature* **473**, 167 (2011).
- [6] C.-L. Pu, W.-J. Pei, and A. Michaelson, *Physica A* **391**, 4420 (2012).
- [7] S. Nie, X. Wang, H. Zhang, Q. Li, and B. Wang, *PLoS ONE* **9**, e89066 (2014).
- [8] J. Ruths and D. Ruths, *Complex Networks IV* (Springer, Berlin, 2013), p. 185.
- [9] M. A. Rahimian and A. G. Aghdam, *Automatica* **49**, 3149 (2013).
- [10] T. Jia and M. Posfai, *Sci. Rep.* **4**, 5379 (2014).
- [11] U. Mackenroth, *Robust Control Systems: Theory and Case Studies* (Springer-Verlag, Berlin, 2010).
- [12] M. Green and D. J. N. Limebeer, *Linear Robust Control* (Dover, Mineola, NY, 2002).
- [13] Z. Wang, L. Tieshan, and H. Zhang, *Int. J. Adapt. Control Signal Process.* **28**, 859 (2014).
- [14] X.-Y. Li, P.-J. Wan, Y. Wang, and C.-W. Yi, *Wireless Commun. Mobile Comput.* **4**, 109 (2004).
- [15] M. Cardei, Y. Shuhui, and J. Wi, in *IEEE Proceedings of the Int. Conf. on Mobile Adhoc and Sensor Systems (MASS)* (IEEE, Pisa, 2007).
- [16] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater, *Fundamentals of Domination in Graphs*, Pure Applied Mathematics (Chapman and Hall/CRC, New York, 1998).
- [17] I. Stojmenovic, M. Seddigh, and J. Zunic, *IEEE Trans. Parallel Distrib. Syst.* **13**, 14 (2002).
- [18] K. M. Alzoubi, P.-J. Wan, and O. Frieder, in *IEEE Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS-35)* (IEEE, Hawaii, 2002).
- [19] R. Pushpalakshmi and A. V. Kumar, in *IEEE Proceedings of Computing Communication and Networking Technologies (ICCCNT)* (IEEE, Karur, 2010).
- [20] I. Khalil and E. R. Weippl, *Innovations in Mobile Multimedia Communications and Applications: New Technologies* (Information Science Reference, Hershey, 2011).
- [21] A. H. Karbasi and R. E. Atani, *Int. J. Security Appl.* **7**, 185 (2013).
- [22] W. Chena, Z. Lub, and W. Wub, *Inform. Sci.* **269**, 286 (2014).
- [23] L. Kelleher and M. Cozzens, *Graphs. Math. Soc. Sci.* **16**, 267 (1988).
- [24] F. Wang *et al.*, *Theor. Comput. Sci.* **412**, 265 (2011).
- [25] S. Eubank, V. S. Anil Kumar, M. V. Marathe, A. Srinivasan, and N. Wang, in *SODA '04 Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (2004), pp. 718–727.
- [26] J. C. Nacher and T. Akutsu, *New J. Phys.* **14**, 073005 (2012).
- [27] J. C. Nacher and T. Akutsu, *J. Phys. Conf. Ser.* **410**, 012104 (2013).
- [28] J. C. Nacher and T. Akutsu, *Sci. Rep.* **3**, 1647 (2013).
- [29] F. Molnár, S. Sreenivasan, B. K. Szymanski, and G. Korniss, *Sci. Rep.* **3**, 1736 (2013).
- [30] S. Wuchty, *Proc. Natl. Acad. Sci. USA* **111**, 7156 (2014).
- [31] F. Molnár, N. Derzsy, B. K. Szymanski, and G. Korniss, [arXiv:1409.7831](https://arxiv.org/abs/1409.7831) [physics.soc-ph] (2014).
- [32] C. Allen and C. F. Stevens, *Proc. Natl. Acad. Sci. USA* **91**, 10380 (1994).
- [33] L. R. Varshney *et al.*, *PLoS Comput. Biol.* **7**, e1001066 (2011).
- [34] J. F. Fink and M. S. Jacobson, *Graph Theory with Applications to Algorithms and Computer Science* (John Wiley & Sons, New York, 1985).
- [35] M. A. Henning and S. T. Hedetniemi, *Discrete Math.* **266**, 239 (2003).
- [36] W. Goddard, S. M. Hedetniemi and S. T. Hedetniemi, *J. Combin. Math. Combin. Comput.* **52**, 169 (2005).
- [37] A. P. Burger, E. J. Cockayne, W. R. Grundlingh, C. M. Mynhardt, J. H. Van Vuuren, and W. Winterbach, *J. Combin. Math. Combin. Comput.* **49**, 159 (2004).
- [38] A. P. Burger, E. J. Cockayne, W. R. Grundlingh, C. M. Mynhardt, J. H. Van Vuuren, and W. Winterbach, *J. Combin. Math. Combin. Comput.* **50**, 179 (2004).
- [39] C.-T. Lin, *IEEE Trans. Automat. Control* **19**, 201 (1974).
- [40] R. W. Shields and J. B. Pearson, *IEEE Trans. Automat. Control* **21**, 203 (1976).
- [41] S. Hosoe and K. Matsumoto, *IEEE Trans. Automat. Control* **24**, 963 (1979).
- [42] H. Maeda, *IEEE Trans. Automat. Control* **26**, 795 (1981).
- [43] K. Murota, *Systems Analysis by Graphs and Matroids: Structural Solvability and Controllability* (Springer-Verlag, Berlin, 1987).
- [44] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, Oxford, 2007).
- [45] M. Boguna, R. Pastor-Satorras, and A. Vespignani, *Eur. Phys. J. B* **38**, 205 (2004).
- [46] J. A. Telle, *Nordic J. Comput.* **1**, 157 (1994).
- [47] P. A. Golovach, J. Kratochvíl, and O. Suchý, *Discrete Appl. Math.* **160**, 780 (2012).
- [48] C. I. Del Genio, T. Gross, and K. E. Bassler, *Phys. Rev. Lett.* **107**, 178701 (2011).
- [49] <http://www.gnu.org/software/glpk>.
- [50] F. Viger and M. Latapy, in *Proceedings of the 11th Intl. Computing and Combinatorics Conf. (COCOON'05)* (Springer LNCS, Kunming, 2005), pp. 440–449.
- [51] J. G. White *et al.*, *Phil. Trans. R. Soc. Lond. B* **314**, 1 (1986).
- [52] D. H. Hall and R. L. Russell, *J. Neurosci.* **11**, 1 (1991).
- [53] L. A. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, *Proc. Natl. Acad. Sci. USA* **97**, 11149 (2000).
- [54] E. Bullmore and O. Sporns, *Nat. Rev. Neurosci.* **10**, 186 (2009).
- [55] R. C. Craddock *et al.*, *Nat. Methods* **10**, 524 (2013).
- [56] A. Clauset, C. R. Shalizi, and M. E. J. Newman, *SIAM Rev.* **51**, 661 (2009).
- [57] <http://tuvalu.santafe.edu/~aaronc/powerlaws>
- [58] The Database of Interacting Proteins. [<http://dip.doe-mbi.ucla.edu/dip>]
- [59] T. S. K. Prasad *et al.*, *Nucl. Acids Res.* **37**, D767 (2009).
- [60] C. Stark *et al.*, *Nucl. Acids Res.* **39**, D698 (2011).
- [61] V. Colizza, R. Pastor-Satorras, and A. Vespignani, *Nat. Phys.* **3**, 276 (2007).
- [62] J. Leskovec, J. Kleinberg and C. Faloutsos, *ACM Trans. Knowl. Discov. Data (ACM TKDD)* **1**, 2 (2007).
- [63] J.-P. Eckmann, E. Moses, and D. Sergi, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14333 (2004).
- [64] R. Milo *et al.*, *Science* **303**, 1538 (2004).
- [65] J. H. Michael and J. G. Massey, *Forest Prod. J.* **47**, 25 (1997).
- [66] J. Leskovec, D. Huttenlocher, and J. Kleinberg, in *Proceedings of the 28th ACM Conf. of Human Factors and Computing Systems (CHI)* (ACM, Atlanta, GA, 2010).
- [67] V. Krebs (unpublished) [<http://www.orgnet.com/>]

- [68] M. Ripeanu, I. Foster, and A. Iamnitchi, *IEEE Internet Computing Journal* **6**, 50 (2002).
- [69] H. Salgado *et al.*, *Nucl. Acids Res.* **41**, D203 (2012).
- [70] C. J. Melian and J. Bascompte, *Ecology* **85**, 352 (2004).
- [71] L. A. Adamic and N. Glance, in *Proceedings of the WWW-2005 Workshop on the Weblogging Ecosystem* (ACM, Chiba, 2005).
- [72] P. J. Taylor, *World City Network: A Global Urban Analysis* (London, Routledge, 2000).
- [73] T. Opsahl, *Soc. Networks* **35**, 159 (2013).
- [74] S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, UK, 1994).
- [75] M. E. J. Newman, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 404 (2001).
- [76] W. Imrich and S. Klavzar, *Products Graphs: Structure and Recognition* (Wiley, New York, 1998).
- [77] <http://vlado.fmf.uni-lj.si/pub/networks/data>.
- [78] M. A. Yildirim *et al.*, *Nat. Biotech.* **25**, 1119 (2007).
- [79] M. C. Costanzo *et al.*, *Nucl. Acids Res.* **29**, 75 (2001).
- [80] T. Wu *et al.*, *Nucl. Acids Res.* **34**, D150 (2006).