Physica A 478 (2017) 168–176

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Spectral coarse grained controllability of complex networks

Pei Wang [a,b,*], Shuang Xu [c,a]

[a] School of Mathematics and Statistics, Henan University, Kaifeng 475004, China
[b] Laboratory of Data Analysis Technology, Henan University, Kaifeng 475004, China
[c] School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China

HIGHLIGHTS

- Spectral coarse grained controllability (SCGC) of complex networks are investigated.
- The ER, SF and SW random networks are with distinct SCGC properties.
- The SW networks are very robust, while the SF networks are sensitive to the SCGC.
- The robust SCGC of random ER networks is an emergent property.
- Some social and biological networks are hard to be controlled during the SCGC.

ARTICLE INFO

Article history:
Received 4 January 2017
Received in revised form 3 February 2017
Available online 27 February 2017

Keywords:
Structural controllability
Spectral coarse graining
Complex network
Degree distribution
Degree heterogeneity

ABSTRACT

With the accumulation of interaction data from various systems, a fundamental question in network science is how to reduce the sizes while keeping certain properties of complex networks. Combined the spectral coarse graining theory and the structural controllability of complex networks, we explore the structural controllability of undirected complex networks during coarse graining processes. We evidence that the spectral coarse grained controllability (SCGC) properties for the Erdös–Rényi (ER) random networks, the scale-free (SF) random networks and the small-world (SW) random networks are distinct from each other. The SW networks are very robust, while the SF networks are sensitive during the coarse graining processes. As an emergent properties for the dense ER networks, during the coarse graining processes, there exists a threshold value of the coarse grained sizes, which separates the controllability of the reduced networks into robust and sensitive to coarse graining. Investigations on some real-world complex networks indicate that the SCGC properties are varied among different categories and different kinds of networks, some highly organized social or biological networks are more difficult to be controlled, while many man-made power networks and infrastructure networks can keep the controllability properties during the coarse graining processes. Furthermore, we speculate that the SCGC properties of complex networks may depend on their degree distributions. The associated investigations have potential implications in the control of large-scale complex networks, as well as in the understanding of the organization of complex networks.

1. Introduction

Controllability of complex networks is in the focus of complex network science and control science in recent years [1–10]. Structural controllability is a typical topic. The structural controllability framework [11,12] assumes nodes in the network

* Corresponding author at: School of Mathematics and Statistics, Henan University, Kaifeng 475004, China.
E-mail addresses: wp0307@126.com, wangpei@henu.edu.cn (P. Wang).

http://dx.doi.org/10.1016/j.physa.2017.02.037

follow the canonical linear, time-invariant dynamics

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{1}$$

where $\mathbf{x}(t) = (x_1(t), \ldots, x_N(t))^T$ describes the state of N nodes at time t in the system. $\mathbf{A}$ is a $N \times N$ matrix, which captures the topological structure of the network. The elements of $\mathbf{A}$ are nonnegative, which describe the wiring strengths between components. $\mathbf{B}$ is a $N \times M$ input matrix ($M \le N$), which describes the nodes controlled by outside controllers. $\mathbf{u}(t) = (u_1(t), \ldots, u_M(t))^T$ is the input vector. For the structural controllability of system (1), it is reported that the following statements are equivalent [13,14]:

(i) System (1) is completely controllable.
(ii) Rank condition: The controllability matrix

$$\mathbf{C} = [\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \ldots, \mathbf{A}^{N-1}\mathbf{B}] \tag{2}$$

has full rank, that is $rank(\mathbf{C}) = N$.
(iii) Popov–Belevitch–Hautus (PBH) eigenvector test: The relationship $\mathbf{v}^T\mathbf{A} = \lambda\mathbf{v}^T$ implies $\mathbf{v}^T\mathbf{B} \neq 0^T$, where $\mathbf{v}$ is a nonzero left eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ.

For large networks, it is computationally difficult to verify the rank condition (ii) and the PBH eigenvector test (iii). In 2011, based on the rank condition, Liu et al. [1] developed a fast tool to investigate the structural controllability of large-scale directed and weighted complex networks, and they solved the minimum input problem. In 2013, based on the PBH eigenvector test, Yuan et al. [9] developed an exact controllability paradigm to achieve full control of networks with arbitrary structures and link-weight distributions, which solved the structural controllability of undirected complex networks. From Yuan et al. [9], the minimum number of controllers N_D is determined by the maximum geometric multiplicity $\mu(\lambda_i)$ of the eigenvalue λ_i for the adjacency matrix $\mathbf{A}$:

$$N_D = \max_i\{\mu(\lambda_i)\}. \tag{3}$$

For large sparse networks with a small fraction of self-loops, N_D is determined by the rank of $\mathbf{A}$:

$$N_D = \max\{1, N - rank(\mathbf{A})\}. \tag{4}$$

For large dense networks with identical link weights w and a small fraction of self-loops, N_D can be determined by:

$$N_D = \max\{1, N - rank(wI_N + \mathbf{A})\}, \tag{5}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix. It is noted that, many real-world complex networks are sparse, especially those arising from biological systems and other real-world systems [15,16].

One of the biggest challenges in complex network science is the sheer size of large complex networks. Many mathematical algorithms used to investigate complex networks run in times that grow polynomially with the sizes of the networks, therefore, quantitative evaluation of large complex networks is a complicated and often difficult task [17]. The recently proposed spectral coarse graining scheme is a promising way to reduce the complexity of large networks. The coarse graining scheme maps the original network into a smaller one, while the spectral and synchronization properties of the initial system are preserved in the reduced networks [17–19].

A natural question is whether the controllability property can be preserved during the spectral coarse graining processes. In other words, if the initial network is coarse grained into small ones, can the ratio of minimum driver nodes $n_D = N_D/N$ be preserved? In the following, we call the above mentioned question as spectral coarse grained controllability (SCGC). It is interesting to explore the SCGC properties for different types of random complex networks and real-world ones. Motivated by these issues, we give a brief overview on the spectral coarse graining process in Section 2, and investigate the SCGC of three classes of random undirected complex networks in Section 3. In Section 4, the SCGC of some real-world undirected complex networks are investigated. Possible physical mechanisms explanations will be presented in Section 5. Discussions and some concluding remarks are in the last Section 6.

2. Spectral coarse graining of complex networks

The spectral coarse graining strategy is based on the idea of grouping nodes with similar spectral components together. The aim is to obtain a reduced network that preserves some properties of the initial network. The properties of interest can be the main characteristics of random walks [17], synchronization [18,19] and so on.

Suppose $\mathbf{A} = (a_{ij})_{N\times N}$ is the adjacency matrix of a complex network, then stochastic matrix $\mathbf{W} = (w_{ij})_{N\times N}$ with

$$w_{ij} = \frac{a_{ij}}{\sum_{k=1}^{N} a_{ik}}$$

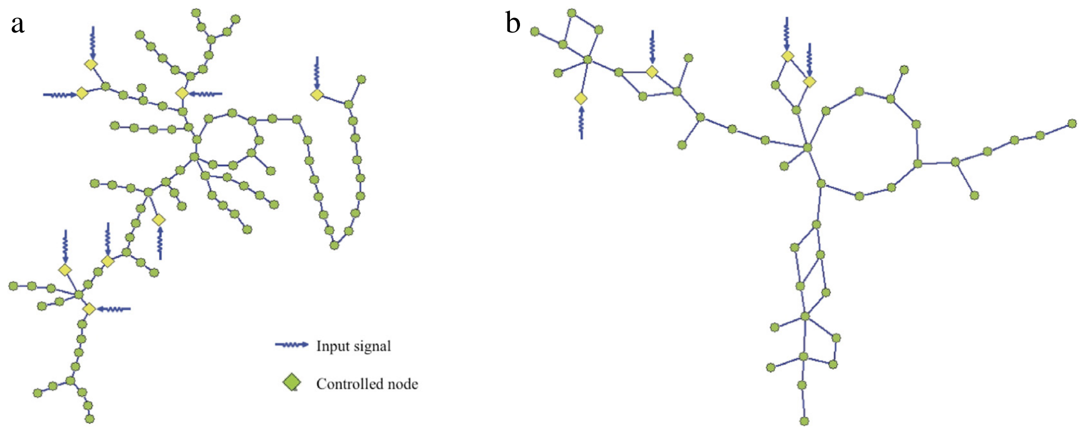

Fig. 1. A toy example of SCGC. (a) Controllability of a toy undirected and unweighted network with 100 nodes. The average degree is 2. One needs control 8 nodes to achieve fully control of the whole network, and $n_D = 8/100 = 8\%$. (b) Controllability of the coarse grained network from (a). The coarse grained network contains 50 nodes, and $\widetilde{n}_D = 4/50 = 8\%$ nodes are needed to be controlled. It is noted that the coarse graining processes are based on the eigenvectors that correspond to the second to the fourth biggest eigenvalues of the stochastic transition matrix **W**. The sets of driver nodes for a network are not unique, we only mark one set of driver nodes in panels (a) and (b) with diamond symbols, respectively.

gives the transition probability from node i to node j. Let the coarse grained network has $\widetilde{N}$ nodes, which correspond to the properly classified $\widetilde{N}$ groups in the initial network, labeled with $L = 1, 2, \ldots, \widetilde{N}$. Based on the spectral coarse graining method, the new stochastic matrix $\widetilde{\mathbf{W}}$ can be written as a product of three matrices,

$$\widetilde{\mathbf{W}} = \mathbf{QWK}, \tag{6}$$

where $\mathbf{Q} \in \mathbb{R}^{\widetilde{N}\times N}$, $\mathbf{K} \in \mathbb{R}^{N\times\widetilde{N}}$, and

$$\mathbf{Q}_{Li} = \delta_{L,L_i}/|C|;\ \mathbf{K}_{iC} = \delta_{C,C_i}. \tag{7}$$

Here, C_i is the index of the group of node i, $|C|$ is the cardinality of group C, δ is the usual Kronecker symbol.

To determine which nodes should be merged, based on Refs. [17,18], consider an eigenvector $\mathbf{v}_k$ of $\mathbf{W}$, merging nodes with the same or similar components in $\mathbf{v}_k$ preserves the eigenvalue λ_k of $\mathbf{W}$. That is, λ_k will be also an eigenvalue of $\widetilde{\mathbf{W}}$. In practice, $\mathbf{v}_k$ can be divided into S equally distributed intervals between the smallest component $\mathbf{v}_{k,\min}$ and the largest component $\mathbf{v}_{k,\max}$, nodes in the same interval can be merged. For simplicity, in the following analysis, we assume the initial networks are undirected and unweighted. The coarse grained networks can be described by either the weighted directed networks $\widetilde{\mathbf{W}} = (\widetilde{w}_{ij})_{\widetilde{N}\times\widetilde{N}}$ or the undirected unweighted networks $\widetilde{\mathbf{A}} = (\widetilde{a}_{ij})_{\widetilde{N}\times\widetilde{N}}$. Here, $\widetilde{\mathbf{A}}$ has the following relationship with $\widetilde{\mathbf{W}}$:

$$\widetilde{a}_{ij} = \begin{cases} 1, & \widetilde{w}_{ij} > 0, \\ 0, & \widetilde{w}_{ij} = 0. \end{cases} \tag{8}$$

$\widetilde{\mathbf{A}}$ is symmetrical, thus the network with the adjacency matrix $\widetilde{\mathbf{A}}$ is undirected and unweighted. When the coarse grained networks are described by $\widetilde{\mathbf{W}}$, it is directed and weighted. In this paper, based on the stochastic matrix **W** of the original network with adjacency matrix **A**, we investigate whether the controllability properties of a network can be preserved while reducing its size. Since the stochastic matrix **W** always has the biggest eigenvalue 1, we consider the eigenvectors of the second, the third and the fourth biggest eigenvalues to perform the spectral coarse graining processes.

Fig. 1 shows a toy example, where the original network is with 100 nodes and average degree equals 2. After performing the coarse graining processes, the size of the network is reduced to 50 nodes. The second to the fourth biggest eigenvalues of the stochastic matrix of the original network are 0.9978, 0.9962 and 0.9898, while the values for the coarse grained network in Fig. 1(b) are 0.9921, 0.9824, 0.9624. The errors for the three pairs of eigenvalues are no more than 0.03, which indicates the coarse grained network can preserve some eigenvalues. It is noted that, there are many sets of minimum driver nodes, for simplicity, we only mark one set of such nodes with diamond symbol in Fig. 1. In the original network in Fig. 1(a), $n_D = 8\%$ nodes are needed to be controlled to achieve fully control of the whole network, and interestingly, for the coarse grained network in Fig. 1(b), the ratio of the controlled nodes $\widetilde{n}_D$ is the same as that for the original network. The toy example indicates some networks can keep the controllability property unchanged even when their sizes are remarkably reduced.

3. Spectral coarse grained controllability of random complex networks

In the following, we investigate the structural controllability of spectral coarse grained random complex networks. For a network with N nodes, we assume nodes follow the dynamics as shown in Eq. (1). We consider the controllability measure

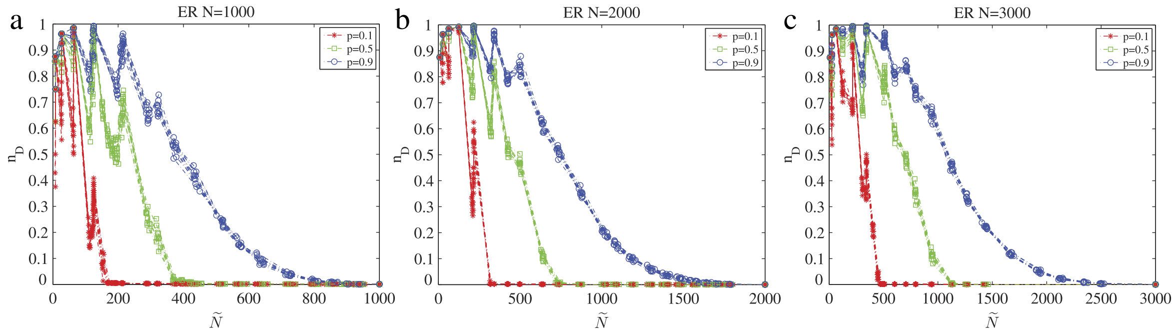

Fig. 2. SCGC of ER networks. The three panels show the cases for $N = 1000, 2000, 3000$, respectively. In each panel, we consider $p = 0.1, 0.5, 0.9$, respectively. Under each set of parameters, we perform 10 simulation runs.

n_D of a large number of artificial networks, which include the ER random networks [20] with the connecting probability p, the SW [21] networks with rewiring probability p and SF networks [22,23] with power-law exponent γ and average degree $\langle k \rangle$.

3.1. Erdös–Rényi random networks

For the ER random networks, every possible edge occurs independently with probability $0 < p < 1$. We consider networks with different sizes N and connecting probabilities p. For simplicity, we consider the cases $N = 1000, 2000, 3000$ and $p = 0.1, 0.5, 0.9$. The curves of the coarse grained network sizes $\widetilde{N}$ versus n_D are shown in Fig. 2, where 10 simulation runs are performed for each case.

From Fig. 2, on one hand, with the increasing of the connecting probabilities p, one needs to control higher fraction of nodes. In fact, with the increasing of p, the average degree of the ER networks increase, and the results indicate that dense ER networks need more nodes to be controlled to achieve fully control. On the other hand, with the decreasing of network sizes according to the coarse graining processes, one needs to control a higher fraction of nodes to fully control the network. Furthermore, for the networks with $N = 1000, 2000, 3000$, if one shrinks the sizes of the three sets of networks to the same size $\widetilde{N}$, then, one can derive the conclusion that relatively more fractions of nodes should be controlled with the increasing of N. For the ER networks, n_D tends to be constant for $\widetilde{N} \geq \textit{threshold}$. For $p = 0.1$, the thresholds are around 320, 340 and 500 for $N = 1000, 2000, 3000$, respectively. For $p = 0.5$, such thresholds are around 600, 860 and 1500, respectively. Whereas, for $p = 0.9$, the threshold values are around 940, 1785 and >1500. This observations indicate that the ER networks can keep certain controllability properties during the spectral coarse graining processes, especially for these sparse and small size networks. However, after the sizes of the networks are reduced to certain thresholds, the controllability of the coarse grained networks become abruptly sensitive to the coarse graining processes, which is an emergent property of the ER network.

3.2. Scale-free random networks

The first SF model [22] is proposed by Barabási and Albert in 1999, and called as the BA SF networks. The degree distributions of SF networks follow $p(k) \propto k^{\gamma}$, where γ is called the power-law exponent. The BA SF networks are with $\gamma = -3$ [22]. Then, Bollobás and Riordan developed the configuration model [23], which can generate random SF networks with arbitrary power-law exponent and arbitrary network sizes. In this paper, we used the configuration model [23] to investigate the SCGC of SF networks. It is reported that many real-world complex networks are SF, and their power-law exponents mostly range among -2 and -3 [22,24]. In this subsection, we consider the SF networks with $N = 1000, 2000, 3000$, for different sizes of networks, we consider the power-law exponents $\gamma \in [-2, -4.5]$.

Fig. 3 shows the average curves of n_D versus $\widetilde{N}$ for different networks, where each curve is averaged over 10 independent simulation runs. For most of the cases, with the decreasing of the network sizes, n_D firstly increases and then decreases. For most of the coarse grained SF networks, more than 20% nodes need to be controlled to achieve fully control of the networks. Another finding is that with the increasing of the absolute values of power-law exponents $|\gamma|$, relatively higher fractions of nodes needed to be controlled.

3.3. Small-world random networks

Small-world networks are with short average path lengths and high clustering coefficients [21,25,26]. Plenty of random SW network generation models have been proposed during the last decades [21,25,24]. For example, the WS model [21], the Newman and Watts (NW) model [25]. In this paper, we consider the WS algorithms. The WS model is based on the random edge rewiring of regular networks.

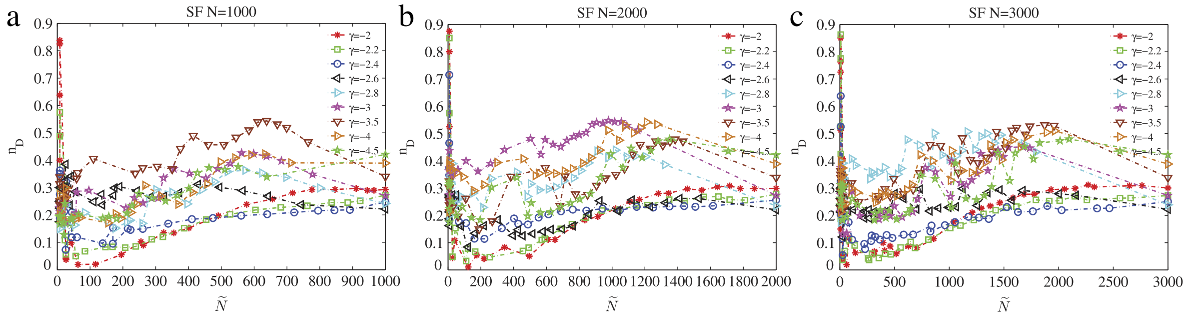

Fig. 3. SCGC of SF networks. The three panels show the cases for $N = 1000, 2000, 3000$, respectively. Each panel shows the average curves under different power-law exponents and each curve is averaged over 10 simulation runs.

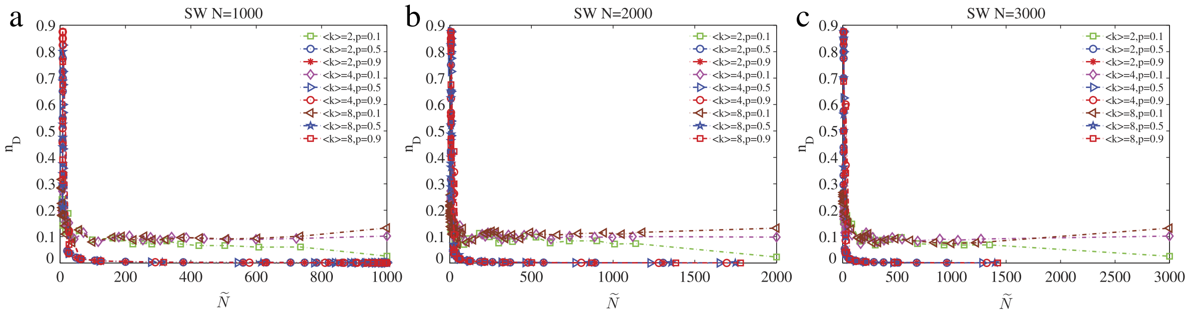

Fig. 4. SCGC of SW networks with $N = 1000, 2000, 3000$. The three figures shows the cases with different average degrees and rewiring probabilities. Under each set of parameter, we perform 10 simulation runs.

The evolution curves of controllability index n_D for the coarse grained SW networks with initial sizes 1000, 2000 and 3000 are shown in Fig. 4. Here, for each cases, networks with average degrees $\langle k \rangle = 2, 4, 8$ and rewiring probabilities $p = 0.1, 0.5, 0.9$ are considered. Different panels show similar phenomena, that is, the SW networks can robustly keep the controllability property during the coarse graining processes. Only when the sizes of the coarse grained networks are very small, the controllability property suffers abruptly changes. For the SW networks with the same size, the fractions n_D keep roughly the same even if the network sizes were reduced to 200 nodes. That is to say, the SCGC of the SW networks can be truly representative of the initial ones even the size of the network is reduced to its 20%. Another finding is that, for sparse SW networks and $\widetilde{N} > 200$, one needs to control relatively more nodes, while for dense networks ($\langle k \rangle \geq 4$), no more than 1% nodes are needed to be controlled to achieve fully control of the whole networks, which are just going in the opposite direction as that for the ER networks. For the effect of network sizes, we observe that they have no much effect on the controllability during the coarse graining processes.

4. Spectral coarse grained controllability of real-world networks

Hereinafter, we consider the controllability of some real-world networks during the coarse graining processes. Some collaboration, communication, computer, infrastructure, social, biological, lexical as well as power networks [27–31] will be investigated.

The sizes of these networks range from hundreds to tens of thousands nodes, and with thousands or tens of thousands of edges, as shown in Table 1. For the SCGC of the considered real-world networks, we summarize the percent of nodes needed to be controlled $\widetilde{n}_D$ when the sizes of the original networks are reduced to around 10%N, 20%N, 40%N, 60%N, 70%N. As well, we show the n_D for the original networks for comparison. For the networks from different categories, we select some representative networks, and show their SCGC curves in Fig. 5.

For the controllability of the original networks, their properties are varied among different categories of networks and different kinds of networks. On one hand, the social networks, some of the communication networks, the route views computer network and many biological networks are difficult to be fully controlled. For example, the Facebook network is very difficult to be controlled, one needs control almost all nodes to achieve fully control of the network. For the human PPI networks and the yeast PPI networks, more than 25% nodes are needed to controlled to achieve fully control. However, the C. Elegans metabolic network is not difficult to be fully controlled, one should only control no more than 6% nodes. On the other hand, the infrastructure networks and the power networks are relatively easier to be controlled. The proportion of driver nodes are all no more than 15.5%. One possible explanation of the observations is that some networks, such as the biological networks and social networks, are highly organized, they evolved from long time evolution, and are very robust to

Table 1
Statistics of some real-world networks and their SCGC. For each network, we show its size, number of edges, number of driver nodes, as well as the fractions of driver nodes while the size of the original network is reduced to around its 10%, 20%, 40%, 60%, 70%.

Category	Network	N	Edges	Refs.	$\widetilde{n}_D(10\%)$	$\widetilde{n}_D(20\%)$	$\widetilde{n}_D(40\%)$	$\widetilde{n}_D(60\%)$	$\widetilde{n}_D(70\%)$	$n_D(100\%)$
Collaborat.	Arxiv general relat.	5 242	28 980	[27]	0.0086	0.0050	0.0195	–	0.1328	0.2224
	Comput. geometry	7 343	23 796	[27]	0.0391	0.0361	0.0740	0.1530	–	0.2517
	High-energy phys.	8 361	31 502	[27]	0.0093	0.0161	0.0398	0.0701	–	0.1563
	Netscience	1 589	5 484	[27]	0.1037	0.0985	0.1720	0.2694	0.3339	0.1089
Communicat.	U.Rovirai Virgili	1 133	5 451	[28]	0.0083	0.0043	0.0024	0.0089	0.0079	0.0371
	as-735	7 716	26 467	[27]	0.0458	0.1235	–	–	–	0.6274
	Pretty good privacy	10 680	24 316	[28]	0.0079	0.0087	–	–	–	0.2436
Computer	Route views	6 474	13 895	[28]	0.0150	0.0751	0.2216	–	–	0.5559
Infrastruct.	Minnesota	2 642	6 606	[27]	0.0100	0.0109	0.0156	0.0145	0.0154	0.0167
	Euroroad	1 174	1 417	[28]	0.0085	0.0468	0.0166	0.0390	0.0429	0.0417
	US power grid	4 941	6 594	[28]	0.0195	0.0417	–	–	–	0.1200
Social	Facebook	2 888	2 981	[28]	0.9327	–	–	–	–	0.9931
	Hamster	2 426	16 631	[28]	0.0117	0.0104	0.0379	0.0825	0.1185	0.2960
Biological	C. elegans	453	4 596	[28]	0.0217	0.0110	0.0052	0.0035	0.0030	0.0596
	Human PPI	1 870	2 277	[28]	0.0149	0.0166	0.1196	0.2327	0.2525	0.3246
	Human PPI	3 133	6 726	[28]	0.0224	0.0725	0.1455	0.1840	0.1982	0.2515
	Yeast PPI	2 361	7 182	[29]	0.0046	0.0021	0.0544	0.1225	0.1704	0.3037
	Yu08	2 018	2 930	[31]	0.0054	0.0256	0.1106	0.1843	0.2158	0.2592
Lexical	King James	1 773	9 131	[28]	0.0046	0.0085	0.0074	0.0426	0.0602	0.1647
Power netw.	662-bus	662	2 474	[27]	0.0143	0.0145	0.0269	–	0.0206	0.0725
	685-bus	685	3 249	[27]	0.0141	0.0070	0.0172	–	0.0161	0.0161
	1138-bus	1 138	4 054	[27]	0.0087	0.0340	0.0603	0.0763	–	0.1520
	bcspwr06	1 454	5 300	[27]	0.0135	0.0071	0.0205	–	0.0452	0.0956
	bcspwr07	1 612	5 824	[27]	0.0238	0.0145	0.0298	0.0428	–	0.1297
	bcspwr08	1 624	6 050	[27]	0.0184	0.0314	0.0350	0.0409	–	0.1225
	bcspwr09	1 723	6 511	[27]	0.0238	0.0149	0.0277	0.0443	–	0.1172
	bcspwr10	5 300	21 842	[27]	0.0084	–	–	–	–	0.0343

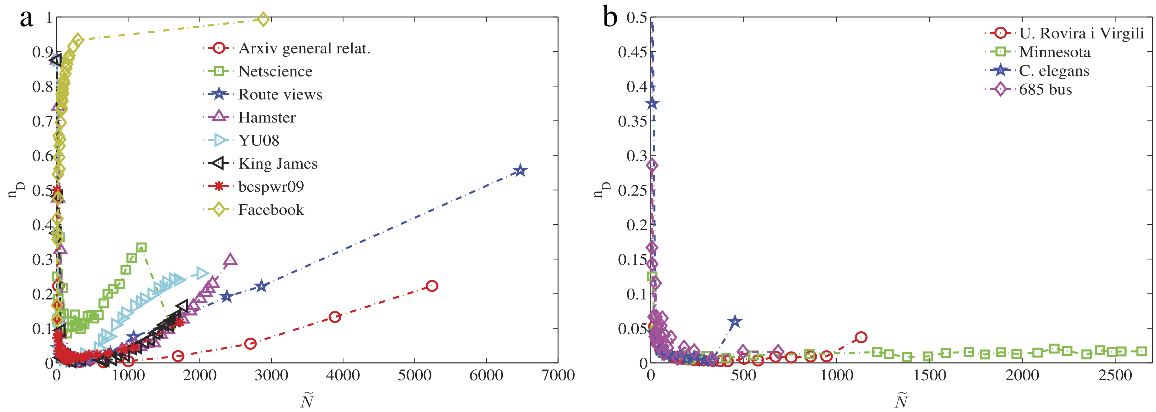

Fig. 5. The SCGC of some real-world networks. The associated networks are selected from Table 1.

environmental stimulus, thus they are very difficult to be controlled. Whereas, some man-made networks, such as the power networks and the infrastructure networks, are influenced by human will, therefore, are relatively more fragile and easier to be controlled. Without doubt, this explanation is not all-or-none, such as the C. Elegans metabolic network. However, the observation for the C. elegans may be influenced by its small network size (only 453 nodes are considered).

The SCGC property are also varied among different categories of networks and different kinds of real-world networks. The properties of the citation network for Arxiv general relativity, the route views, the Hamster, the Yeast PPI network from Ref. [31], are similar, which are in accordance with the SF random networks. With the decreasing of network sizes, $\widetilde{n}_D$ quickly decreases to around zero, and then increases after the sizes of the network become very small. This helps us to further confirm the SF property of these real-world networks. The performance of the Minnesota infrastructure network, the C. elegans PPI network and the 685 bus power network are very similar to the performance of the small-world networks, it is very robust to the coarse graining processes. The Facebook network is very difficult to be controlled, even when its sizes are drastically coarse grained. For example, when the size of the network is reduced to around to ten percents of its original size, more than 90% nodes are still needed to be controlled. There are some differences between the netscience collaboration network and the other networks, one only needs control about 11% nodes to achieve fully control the original

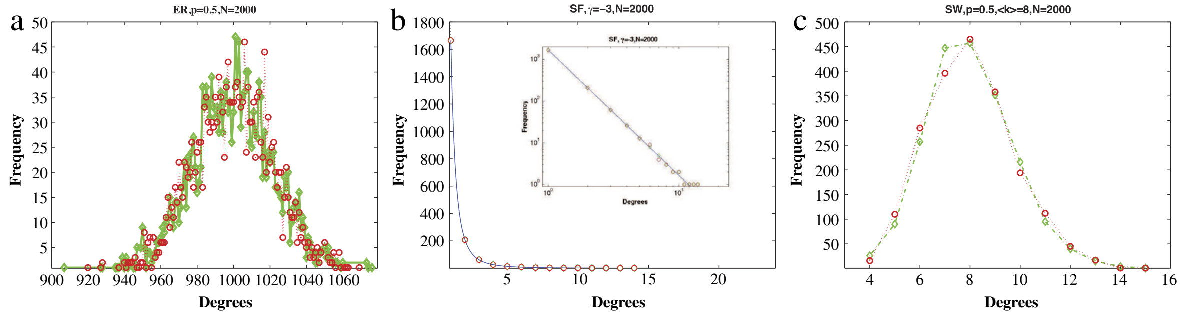

Fig. 6. Degree distributions for the three types of networks. Networks with $N = 2000$ are considered, the ER networks are with $p = 0.5$, the SF networks are with $\gamma = -3$, while the SW networks are with $p = 0.5$, $\langle k \rangle = 8$.

network, however, for the coarse grained networks, when its size is reduced to around 1200, more than 30% driver nodes are needed, after that, with the decreasing of the network sizes, the $\widetilde{n}_D$ quickly decreases and then increases after the sizes of the coarse grained network become very small.

In summary, the SCGC of real-world networks are varied among different categories and different kinds of networks, and they may also be influenced by their sampling sizes. For small networks, such as the C. elegans PPI network and the power network with 685 bus, their SCGC properties are different from larger networks of the same categories. Moreover, the observations also indicate us the social networks and the biological networks, which are highly evolved and organized, are more difficult to be controlled, even these networks are reduced to very small sizes.

5. Physical mechanisms

In this subsection, we briefly discuss the mechanisms behind the observations in the above sections. In Ref. [1], it is reported that the number of driver nodes is determined mainly by the networks' degree distributions. We declare that the SCGC properties of complex networks also correlate to its degree distributions. Based on the degree distributions of different types of networks, we explore why certain kinds of networks can preserve SCGC, and why the dense ER networks are more difficult to be controlled than sparse ones, while dense SW networks are easier to be controlled than sparse ones.

The degree distribution for the ER random network follows the Poisson distribution:

$$P_{ER}(k) \approx \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}. \tag{9}$$

From Ref. [30], the degree distribution for the SW network from the random rewiring mechanism follows Eq. (10).

$$P_{SW}(k) = \begin{cases} \displaystyle\sum_{n=0}^{\min\{k-K/2,K/2\}} C_{K/2}^{n}(1-p)^n p^{K/2-n} \frac{(pK/2)^{k-K/2-n}}{(k-K/2-n)!} e^{-pK/2}, & k \ge K/2, \\ 0, \quad k < K/2 \end{cases} \tag{10}$$

where K is an even number, represents the node degree of the regular network that before randomization. From Eq. (10), one can conclude that the degrees of all nodes in the SW network are roughly the same, similar to the ER networks. The degree distribution for the SF network follows:

$$P_{SF}(k) \propto k^{\gamma}. \tag{11}$$

γ represents the power-law exponent. The SF networks are heterogeneous, nodes with large degrees only take a small fraction, while the vast majority of nodes are with low degrees.

Fig. 5 shows the degree distributions of some random generated networks. The degree heterogeneity [32] of complex networks is defined as:

$$H = \frac{\langle k^2 \rangle}{\langle k \rangle^2}. \tag{12}$$

For the three kinds of networks as shown in Fig. 6, their average degree heterogeneities are $H_{ER} = 1.0004$, $H_{SF} = 1.5938$, $H_{SW} = 1.0476$, respectively. The degree distribution and heterogeneity for the ER and SW networks are very similar, therefore, the SCGC properties for the two types of networks also have some similarities. The differences between Figs. 3 and 4 may result to their differences on the average degrees and the numbers of unique degree values. The average degrees of the ER networks are far larger than that for the SW networks. Moreover, since the SW networks are rewired from the regular networks, the number of their unique degree values may be far less than that for the ER networks, especially under small rewiring probability values. If a network with very low number of unique degree values, then a large fraction of nodes may share the same degree. The nodes with the same degree tend to have similar eigenvector components, and thus will be merged during the coarse graining processes with high probability. From this point of view, it is not difficult to understand why the controllability properties of the SW networks are so robust during the spectral coarse graining processes, and why the ER networks are with certain robustness.

The SF networks are heterogeneous, and most of the nodes are with low degrees. However, these nodes with the same degree may be neighbors of different hubs, which result to the differences of their eigenvector components. Therefore, it is easy to understand the fluctuation of the SCGC curves as shown in Fig. 4. Furthermore, it is reported that inhomogeneous networks are difficult to be controlled, and homogeneous networks can be controlled with a few driver nodes [1]. For example, for the SF networks with $N = 2000, \gamma = -3$, one needs to control 20% to 50% nodes, while for the SW and the ER networks with $N = 2000, p = 0.5$, one only needs to control no more than 0.5% nodes. Therefore, our results on SCGC agree with the existing finding.

Hereinafter, for real-world networks, we classify the physical mechanisms of our observations. For simplicity, we only analyze the Facebook network [28] and the Minnesota road network [27]. From the network information as shown in Ref. [28], the Facebook network is highly heterogeneous, its maximum degree is 769, which connects with almost 27% nodes of the network, while more than 1500 nodes are with degree one. Furthermore, the Facebook network is disassortative, with Pearson correlation coefficient [15,16] −0.6682, which indicates large degree nodes tend to connect with low degree ones. The massive degree one nodes and the disassortativity indicate the Facebook network are with some local star structures. However, the undirected local star network with N nodes needs $N - 2$ driver nodes [9]. Therefore, the Facebook network is difficult to be controlled. The Minnesota road network [27] (the largest component) is homogeneous and small-world, with degree heterogeneity equals to $H = 1.0893$ and SW index [26] equals to 4.1046. Therefore, it is not difficult to understand why the controllability of the Minnesota road network is so robust to the spectral coarse graining processes.

6. Discussion and conclusion

Social, biological, technological and many other systems can be described by complex networks, and complex network theory has wide applications in the associated systems. In recent years, there are some fundamental questions in complex network theory. On one hand, structural controllability of complex networks is a hot topic [1–10]. A fundamental question is how many nodes needed to be controlled to achieve full control of the network. On the other hand, how to reduce the size of a complex network while keeping certain properties is another fundamental question. The latter question is the so called coarse graining of complex networks [17,18]. The coarse graining of complex networks becomes more and more important with the arriving of the era of big data.

By combining the concept of structural controllability and the recently proposed spectral coarse graining theory of complex networks, we investigate the SCGC of complex networks. By considering the ER random networks, the SF random networks, the SW random networks and some real-world complex networks, we find different networks have different performance during the spectral coarse graining processes. The SW networks are very robust, while the SF networks are sensitive during the coarse graining processes. For the dense ER networks, during the coarse graining processes, there exists a threshold value of the coarse grained network sizes, which separates the networks into robust and sensitive. Investigations on some real-world complex networks indicate that the SCGC properties are also varied among different categories of networks and different kinds of networks, some highly organized networks are more difficult to be controlled, while some other networks can keep the controllability properties during the coarse graining processes. Furthermore, based on the analysis of degree distribution and degree heterogeneity, we discuss the physical mechanisms of the observed phenomenon, and we speculate that the SCGC of complex networks may highly depend on their degree distributions. The associated observations also indicate the SCGC of real-world networks can help us to judge the type of the networks.

The coarse graining processes considered in this paper are based on the spectral coarse graining theory proposed by Gfeller and Rios in 2007 and 2008 [17,18]. From the spectral coarse graining theory, the Laplacian matrices of the coarse grained networks can keep the smallest eigenvalues roughly unchanged, which are crucial in keeping the synchronization ability of the coarse grained networks. A fundamental idea for the spectral coarse graining theory is to merge nodes with the similar eigenvector components, where the eigenvectors correspond to the lowest two or three eigenvalues are considered. In our paper, since we are not care about the synchronization of complex networks, the coarse graining processes are based on the stochastic random walk matrix, and the eigenvectors correspond to the second to the fourth largest eigenvalues are considered. From our investigation, we find the SCGC of the SW networks is very robust during the coarse graining processes, while the SF networks and the ER networks are not so robust. Our future works will consider to design some other coarse graining strategies while keeping the controllability property of as many types of networks. Furthermore, our current investigations only consider the undirected unweighted complex networks, another interesting future research direction is to consider directed or weighted complex networks.

The associated investigations help us understand the different performances of the SCGC of different networks. It can guide us to reduce the complexity of large-scale complex networks and has potential implications in the control of large-scale complex networks [33–36]. The SCGC can also help us to understand the structure of complex networks.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant 61304151. The Key Scientific Research Projects in Colleges and Universities of Henan under grant 17A120002. The Basal Research Fund of Henan University under grant yqpy20140049.

References

[1] Y. Liu, J.J. Slotine, A.L. Barabási, Controllability of complex networks, Nature 473 (2011) 167–173.
[2] F. Pasqualetti, S. Zampieri, F. Bullo, Controllability metrics, limitations and algorithms for complex networks, IEEE Trans. Control Netw. Syst. 1 (2014) 395–400.
[3] S. O'Rourke, B. Touri, Controllability of random systems: universality and minimal controllability, arXiv, 2015. arXiv:1506.03125.
[4] A. Olshevsky, Minimal controllability problems, IEEE Trans. Control Netw. Syst. 1 (2014) 249–258.
[5] M. Pósfai, Y. Liu, J.J. Slotine, A.L. Barabási, Effect of correlations on network controllability, Sci. Rep. 3 (2013) art.no. 1067.
[6] T. Jia, Y. Liu, E. Csóka, M. Pósfai, J.J. Slotine, A.L. Barabási, Emergence of bimodality in controlling complex networks, Nature Commun. 4 (2013) 1–6.
[7] T. Jia, A.L. Barabási, Control capacity and a random sampling method in exploring controllability of complex networks, Sci. Rep. 3 (2013) art. no. 2354.
[8] J. Ruths, D. Ruths, Control profiles of complex networks, Science 343 (2014) 1373–1376.
[9] Z. Yuan, C. Zhao, Z. Di, W. Wang, Y. Lai, Exact controllability of complex networks, Nature Commun. 4 (2013) 1–6.
[10] C. Zhao, W. Wang, Y. Liu, J.J. Slotine, Intrinsic dynamics induce global symmetry in network controllability, Sci. Rep. 5 (2015) art.no. 8422.
[11] R.E. Kalman, Mathematical description of linear dynamical systems, J. Soc. Ind. Appl. Math. Ser. A 1 (1963) 152–192.
[12] C.T. Lin, Structural controllability, IEEE Trans. Automat. Control 19 (1974) 201–208.
[13] W.J. Rugh, Linear System Theory, Prentice Hall, Englewood Cliffs, NJ, USA, 1996.
[14] L. Xiang, J.J.H. Zhu, F. Chen, G. Chen, Controllability of weighted and directed networks with nonidentical node dynamics, Math. Probl. Eng. 2013 (2013) art. no. 405034.
[15] P. Wang, X. Yu, J. Lü, Identification and evolution of structurally dominant nodes in protein protein interaction networks, IEEE Trans. Biomed. Circuits Syst. 8 (2014) 87–97.
[16] P. Wang, Y. Chen, J. Lü, Q. Wang, X. Yu, Graphical features of functional genes in human protein interaction network, IEEE Trans. Biomed. Circuits Syst. 10 (2016) 707–720.
[17] D. Gfeller, P. De Los Rios, Spectral coarse graining of complex networks, Phys. Rev. Lett. 99 (2007) art.no. 038701.
[18] D. Gfeller, P. De Los Rios, Spectral coarse graining and synchronization in oscillator networks, Phys. Rev. Lett. 100 (2008) art.no.174104.
[19] J. Chen, J. Lu, X. Lu, X. Wu, G. Chen, Spectral coarse graining of complex clustered networks, Commun. Nonlinear Sci. Numer. Simul. 18 (2013) 3036–3045.
[20] P. Erdös, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960) 17–61.
[21] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[22] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[23] B. Bollobás, O.M. Riordan, Mathematical results on scale-free random graphs, in: S. Bornholdt, H.G. Schuster (Eds.), Handbook of Graphs and Networks: from the Genome to the Internet, Wiley-VCH, Berlin, 2002.
[24] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2003) 167–256.
[25] M.E.J. Newman, D.J. Watts, Renormalization group analysis of the small-world network model, Phys. Lett. A 263 (1999) 341–346.
[26] M.D. Humphries, K. Gurney, Network small-world-ness: a quantitative method for determining canonical network equivalence, PLoS One 3 (2008) art.no. e0002051.
[27] T.A. Davis, Y. Hu, The University of Florida sparse matrix collection, ACM Trans. Math. Software 38 (2011) 1–25.
[28] J. Kunegis, KONECT-The Koblenz network collection, in: Proc. Int. Conf. on World Wide Web Companion, May 13–17, 2013, pp. 1343–1350.
[29] S. Sun, L. Ling, N. Zhang, G. Li, R. Chen, Topological structure analysis of the protein-protein interaction network in budding yeast, Nucleic Acids Res. 31 (2003) 2443–2450.
[30] A. Barrat, M. Weigt, On the properties of small world networks, Eur. Phys. J. B 13 (2000) 547–560.
[31] H. Yu, et al., High-quality binary protein interaction map of the Yeast interactome network, Science 322 (2008) 104–110.
[32] H. Hu, X. Wang, Unified index to quantifying heterogeneity of complex networks, Physica A 387 (2008) 3769–3780.
[33] W. Wang, Q.-H. Liu, S.-M. Cai, M. Tang, L.A. Braunstein, H.E. Stanley, Suppressing disease spreading by using information diffusion on multiplex networks, Sci. Rep. 6 (2016) 29259.
[34] G. Wen, Z. Duan, G. Chen, W. Yu, Consensus tracking of multi-agent systems with Lipschitz-Type node dynamics and switching topologies, IEEE Trans. Circuits Syst. I 61 (2014) 499–511.
[35] S. Wen, Z. Zeng, M.Z.Q. Chen, T. Huang, Synchronization of switched neural networks with communication delays via the event-triggered method, IEEE Trans. Neural Netw. Learn. Syst. (2017) http://dx.doi.org/10.1109/TNNLS.2016.2580609.
[36] W. Yu, W. Ren, W. Zheng, G. Chen, J. Lü, Distributed control gains design for consensus in multi-agent systems with second-order dynamics, Automatica 49 (2013) 2107–2115.