Vol.2013-MPS-95 No.16
2013/9/27

A Dominating Set Approach to Structural Controllability of Unidirectional Bipartite Networks

Jose C. Nacher[1,a)]   Tatsuya Akutsu[2,b)]

Abstract: In this report, we present a dominating set (DS)-based approach to structural controllability of unidirectional bipartite networks, where we assume that driver nodes (i.e., control nodes) are selected from one side of nodes and the purpose is to control all nodes in the other side. We show that if DS is selected as a set of driver nodes, the system can be structurally controllable under the assumption that each driver node can control its outgoing edges independently. We also show a relationship between the size of the minimum dominating set and the exponent of the degree distribution in scale-free networks.

1. Introduction

Control of complex networks has recently been a hot topic in the field of network science. After a pioneering study by Lombardi and Hörnquist [6] that applied controllability theory to complex networks, Liu et al. established a relationship between structural controllability of a complex network and the maximum matching of the corresponding bipartite graph [4]. They analytically showed that the minimum number of driver nodes that are required to control the entire system is small for homogeneous networks (i.e., random networks, power-law networks with $\gamma > 2$) whereas it is large for heterogeneous networks (i.e., power-law networks with $\gamma < 2$). Following their seminal work, many studies have been done [1], [3], [5], [12], [16], [17].

Although most of these studies focused on nodal dynamics, Nepusz and Vicsek analyzed the control problem from the angle of edge dynamics [12]. Independently, we have recently introduced the minimum dominating set (MDS) approach to control complex networks [9], [10], which conceptually has similarities with the controlling link dynamics [12]. It is to be noted that a dominating set (DS) is a well-known concept in graph theory and has already been applied to the design and/or control of various kinds of discrete systems, which include mobile ad hoc networks (MANET) [14], transportation routing and computer communication networks [2]. Our theoretical findings suggest that scale-free networks with small scaling exponent values ($\gamma < 2$), where high-degree nodes are present, require relatively few nodes to be controlled.[*1] Molnàr et al. studied further on the size of MDS in other types of scale-free networks [8].

1 Department of Information Science, Faculty of Science, Toho University, Miyama 2-2-1, Funabashi, Chiba 274-8510, Japan
2 Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, 611-0011, Japan
a) nacher@is.sci.toho-u.ac.jp
b) takutsu@kuicr.kyoto-u.ac.jp
*1 Analysis in [10] is much more accurate than in [9].

The above mentioned studies focused on unipartite networks. However, bipartite networks often appear in the real world, which include the Facebook-like forum, the firms-world city network, the cond-mat scientific collaboration, and the human drug-target protein network [11]. Therefore, we study the control problem on bipartite networks. Since the nodal dynamics approach requires $O(n)$ driver nodes as shown in Section 3, we emply the MDS-based approach, where n is the total number of nodes in the network. In addition to previous analyses [9], [10], we take the effect of the maximum degree H into account. As the main result, we analytically derive that $O(1/H^{(2-\gamma)(\gamma-1)})$ driver nodes are enough to structurally control a bipartite network if $1 < \gamma < 2$, under some reasonable assumptions. In this technical report, we focus on mathematical analysis. Results on random bipartite networks, computer simulation, and analysis of real networks can be seen in [11].

2. Dynamics Model

In this section, we introduce a dynamics model for bipartite networks, following the formalization for unipartite networks [4]. Let $G(V_\top, V_\bot; E)$ be a bipartite network consisting of two sets of nodes $V_\top$ and $V_\bot$ and a set of edges E. It is to be noted that the directions of all edges are from $V_\top$ to $V_\bot$ in this definition. This assumption is reasonable for such networks as the drug-target networks and the Facebook-like forum because activities of nodes in $V_\top$ are usually not affected by those in $V_\bot$.

Let $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_{n_2}(t))^T$ be the state of nodes in $V_\bot$ at time t, where M^T denotes the transposed matrix of M.

Suppose that $U = \{u_1, u_2, \ldots, u_h\}$ is the set of driver nodes (in our sense) selected from $V_\top$ and each u_i has d_i edges. We define a state vector $\mathbf{u}$ for all edges from U by

$$\begin{aligned}\mathbf{u}(t) \;=\; & (u_{1,1}(t), u_{1,2}(t), \ldots, u_{1,d_1}(t),\\ & u_{2,1}(t), u_{2,2}(t), \ldots, u_{2,d_2}(t),\\ & \ldots,\end{aligned}$$

$u_{h,1}(t), u_{h,2}(t), \ldots, u_{h,d_h}(t))^T$.

We rename $u_{i,j}$s by $u'_1, u'_2, \ldots, u'_l$ and and rewrite $\mathbf{u}(t)$ by

$$\mathbf{u}(t) = (u'_1(t), u'_2(t), \ldots, u'_l(t))^T,$$

where each node u'_i has only one outgoing edge that corresponds to the original edge.

Here, we assume that the dynamics is given by

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

where A is the diagonal matrix (i.e., $A_{i,j} = 0$ for all $i \neq j$) and B satisfies that $B_{i,j} \neq 0$ only if u'_j is connected to w_i. It is to be noted that A can be the null matrix (i.e., any node in $V_\perp$ does not have a self-loop).

Then, it is clear that this system is structurally controllable if each node $w \in V_\perp$ has at least one incoming edge (i.e., U is a dominating set).

Furthermore, there can be connections between nodes in $V_\perp$ because addition of non-zero elements to A does not impair the structural controllability. However, in this case, the number of required driver nodes may be fewer than that derived from our model.

In the above, we assumed that $u_i \notin V_\top - U$ does not have any effect on $v_j \in V_\perp$. However, this assumption can be removed if we can know the signals from nodes in $V_\top - U$ to $V_\perp$. Suppose that v_j has incoming edges from $u_{i_0}, u_{i_1}, \ldots, u_{i_k}$ where $u_{i_0} \in U$ and $u_{i_1}, \ldots, u_{i_k} \notin U$ (this case can be trivially extended for the case where there exist edges from multiple nodes in U). Let the state vector from these nodes to v_j be $(u^j_0(t), u^j_1(t), \ldots, u^j_k(t))^T$ and the vector of corresponding weights be $(b_0, b_1, \ldots, b_k)^T$. Then, in order to remove the effects from $u_{i_1}, \ldots, u_{i_k}$ to v_j, it is enough to add the following term to $u^j_0(t)$:

$$-\frac{1}{b_0}(b_1 u^j_1(t) + \cdots b_k u^j_k(t)).$$

3. Structural Controllability of Bipartite Networks

In this work, we use a modified version of the dominating set, in which a set must be selected from $V_\top$ and it is enough to dominate all nodes in $V_\perp$ (i.e., for all node $w \in V_\perp$, there exists a node $v \in V_\top$ such that $(v, w) \in E$). This corresponds to a *set cover* problem by associating a set $S_v = \{w | (v, w) \in E\}$ for each $v \in V_\top$. We use MDS to denote the minimum dominating set (i.e., the dominating set with the minimum number of nodes) in the above sense (see also Fig. 1).

As proved in [9], a unipartite network is structurally controllable if a dominating set is selected as a set of control nodes under the assumption that each control node can control its outgoing edges separately.

We can also consider structural controllability under the assumption in [4] that each driver node can control only its value. In such a case, the number of driver nodes is determined by the number of nodes in V^R not appearing in a maximum matching of the adjunct bipartite graph $G'(V^L, V^R; E')$ [4]. However, in this case, all nodes in V^R corresponding to $V_\top$ remain unmatched

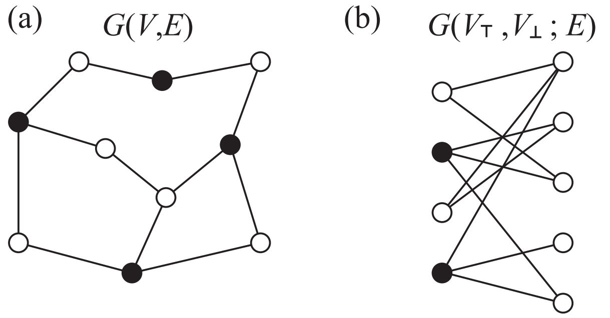

Fig. 1 Minimum dominating set (MDS) for a unipartite network (a) and our definition of MDS for a bipartite network (b), where black circles denote nodes in MDS. MDS in (b) corresponds to a set cover for $V_\perp$.

because there is no edge connecting to any of these nodes (see Fig. 2). Therefore, we have

Proposition 1 The number of driver nodes in the sense of [4] is at least $|V_\top|$ for a bipartite network $G(V_\top, V_\perp; E)$ such that $E \subseteq V_\top \times V_\perp$.

Since $|V_\top|$ is usually a very large number and the size of MDS is expected to be much smaller than $|V_\top|$, we focus on the structural controllability in terms of MDS.

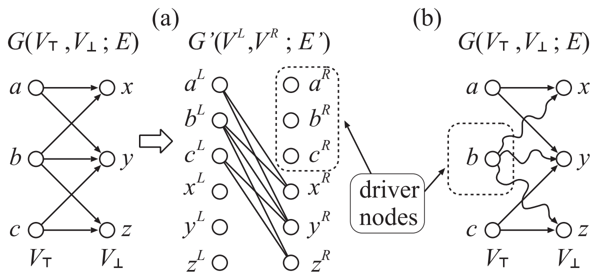

Fig. 2 Comparison of the model by Liu et al. [4] (a) with the MDS model (b) for bipartite networks. In this example, $\{b\}$ is the dominating set (i.e., set cover) of $G(V_\top, V_\perp; E)$, whereas $\{a^R, b^R, c^R\}$ cannot appear in the maximum matching of $G'(V^L, V^R; E')$ and thus $\{a, b, c\}$ must be the set of driver nodes in the sense of [4].

Now, we formally describe a relationship between DS and the structural controllability of a bipartite network. As a direct consequence of Theorem 1 in [9] and the discussions in Section 2, we have the following proposition.

Proposition 2 Suppose that we need to control the states of nodes only in $V_\perp$ and that every node in the DS ($\subseteq V_\top$) can control all of its outgoing links separately. Then, the network is structurally controllable by selecting the nodes in the DS as the driver nodes.

Recall that we assumed that all of the nodes in a dominating set DS must be selected from $V_\top$ and that it is necessary to dominate all nodes in $V_\perp$ (we need not dominate the nodes in $V_\top$), which means that DS is a *set cover* for $V_\perp$. In [9], [10], structural controllability was studied in terms of MDS for unipartite graphs. In this report, we present the analytically derived estimations for the minimum number of drivers using the MDS controllability approach for bipartite networks in which all edges are directed from $V_\top$ to $V_\perp$.

4. Theoretical Analysis of the MDS Size in Bipartite Networks

We assume that the degree distribution of $V_\top$ and $V_\perp$ follow $P_\top(k) \propto k^{-\gamma_1}$ and $P_\perp(k) \propto k^{-\gamma_2}$, respectively. We let $n_1 = |V_\top|$ and $n_2 = |V_\perp|$.

4.1 The case of $\gamma_1 > 2$

We assume $P_\top(k)$ follows $\alpha_1 k^{-\gamma_1}$ with cut off at $k = n_1$, where $\gamma_1 > 2$. From $\alpha_1 n_1 \int_1^{n_1} k^{-\gamma} dk = n_1$, we have $\alpha_1 \approx \gamma_1 - 1$.

For $S \subseteq V_\top$, $\Gamma(S)$ denotes the set of edges between S and $V_\perp$ (i.e., $\Gamma(S) = \{(u,v) \mid u \in S \text{ and } v \in V_\perp\}$). The following property is trivial:

$$\text{if } |\Gamma(S)| < n_2,\ \ S \text{ can not dominate } V_\perp.$$

Let S be the set of nodes whose degree is greater than or equal to K. It is to be noted that S is chosen so that the total degree (i.e., the number of edges incident to S) is maximized among the sets with the same cardinality.

We estimate the size of $\Gamma(S)$ as follows.

$$\begin{aligned}|\Gamma(S)| &< \alpha n_1 \int_K^{n_1} k \cdot k^{-\gamma_1} dk \approx n_1(\gamma_1 - 1)\int_K^{n_1} k^{-\gamma_1+1} dk \\ &= n_1 \cdot \left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \left(\frac{1}{K^{\gamma_1 - 2}} - \frac{1}{n_1^{\gamma_1 - 2}}\right) \\ &< n_1 \cdot \left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \frac{1}{K^{\gamma_1 - 2}}.\end{aligned}$$

If S is a DS, the last term should be no less than n_2. Therefore, the following inequality should be satisfied:

$$n_1 \cdot \left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \frac{1}{K^{\gamma_1 - 2}} > n_2.$$

By solving this inequality, we have

$$K < \left[\left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \left(\frac{n_1}{n_2}\right)\right]^{1/(\gamma_1 - 2)}.$$

Then, the size of S is estimated as

$$\begin{aligned}|S| &\approx \alpha n_1 \int_K^{n_1} k^{-\gamma_1} dk \approx n_1 \left(\frac{1}{K^{\gamma_1 - 1}} - \frac{1}{n_1^{\gamma_1 - 1}}\right) \approx n_1 \cdot \frac{1}{K^{\gamma_1 - 1}} \\ &> \left[\left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right)\right]^{-\frac{\gamma_1 - 1}{\gamma_1 - 2}} \cdot \left(\frac{n_2}{n_1}\right)^{\frac{\gamma_1 - 1}{\gamma_1 - 2}} \cdot n_1.\end{aligned}$$

From this inequality and the fact that $V_\top$ is a trivial dominating set, we can see that the size of the minimum dominating set is $\Theta(n)$ (for fixed γ_1) and the coefficient increases as γ_1 increases if $n_2 \approx n_1$.

4.2 Case of $\gamma_1 < 2$

In this section, we focus on degree distribution for $V_\top$ and thus we let $\gamma = \gamma_1$, $n = n_1$, and $m = n_2$.

We assume that the maximum degree is H. Then, we have

$$n = \alpha n \int_1^H k^{-\gamma} dk = \frac{\alpha n}{\gamma - 1}(1 - H^{1-\gamma}) \approx \frac{\alpha n}{\gamma - 1}$$

from which $\alpha = \gamma - 1$ follows.

Let DS be the set of nodes with degree between B and H. Then, the number of nodes N_{DS} in DS is given by

$$N_{DS} = \alpha n \int_B^H k^{-\gamma} dk = n\left(B^{1-\gamma} - H^{1-\gamma}\right) = O(nB^{1-\gamma}).$$

On the other hand, the total number of edges E_G is

$$\begin{aligned}E_G &= \alpha n \int_1^H k \cdot k^{-\gamma} dk = \frac{\gamma - 1}{2 - \gamma} \cdot n \cdot (H^{2-\gamma} - 1) \\ &\approx \frac{\gamma - 1}{2 - \gamma} \cdot nH^{2-\gamma} = \langle k \rangle n,\end{aligned}$$

from which $\langle k \rangle \approx \frac{\gamma - 1}{2 - \gamma} \cdot H^{2-\gamma}$ follows.

The number of edges E_{NDS} not covered by DS is

$$E_{NDS} = \alpha n \int_1^B k \cdot k^{-\gamma} dk \approx \frac{\gamma - 1}{2 - \gamma} \cdot n \cdot B^{2-\gamma}$$

Therefore, the probability that an arbitrary edge is not covered by DS is

$$\frac{E_{NDS}}{E_G} \approx \left(\frac{B}{H}\right)^{2-\gamma}$$

Let $V_\perp \ominus DS$ denote the set of nodes in $V_\perp$ that are not dominated by DS. Since a node is dominated by DS if at least one edge connecting to the node is covered by DS, the expected number of nodes (denoted by $N_{V_\perp \ominus DS}$) of $V_\perp \ominus DS$ is bounded as

$$N_{V_\perp \ominus DS} \le O\left(m \cdot \left(\frac{B}{H}\right)^{2-\gamma}\right),$$

where m is the number of nodes in $V_\perp$.

In order to dominate $V_\perp \ominus DS$, it is enough to select at most $N_{V_\perp \ominus DS}$ nodes from $V_\top$. Therefore, the size of a minimum dominating set is bounded by

$$|DS| + N_{V_\perp \ominus DS} \le O\left(nB^{1-\gamma} + m\left(\frac{B}{H}\right)^{2-\gamma}\right).$$

Then, in order to find B minimizing this order, we let

$$nB^{1-\gamma} = m\left(\frac{B}{H}\right)^{2-\gamma},$$

which results in $B = \frac{n}{m} \cdot H^{2-\gamma}$.

Therefore, an upper bound of the size of the dominating set is estimated as

$$O\left(\frac{n^{2-\gamma} \cdot m^{\gamma-1}}{H^{(2-\gamma)(\gamma-1)}}\right).$$

It is to be noted that $(2 - \gamma)(\gamma - 1) \le 0.25$.

By using $\langle k \rangle \approx \frac{\gamma - 1}{2 - \gamma} \cdot H^{2-\gamma}$, this upper bound can also be written as

$$O\left(\frac{n^{2-\gamma} \cdot m^{\gamma-1}}{\langle k \rangle^{(\gamma-1)} \left(\frac{2-\gamma}{\gamma-1}\right)^{\gamma-1}}\right).$$

On the other hand, if $H = n$, the upper bound becomes

$$O\left(n^{(2-\gamma)^2} \cdot m^{\gamma-1}\right).$$

If $m = cn$ where c is a constant, this order is $O(n^{\gamma^2 - 3\gamma + 3})$, which takes the minimum order ($O(n^{0.75})$) when $\gamma = 1.5$.

Readers might wonder why γ_2 is ignored in the above analysis. We briefly discuss this point. The number of nodes with degree 1 or 2 in $V_\perp$ is approximated by

$$\alpha_2 n_2 \int_1^2 k^{-\gamma} dk = \frac{\alpha_2 n_2}{\gamma_2 - 1}\left(1 - \frac{1}{2^{\gamma_2 - 1}}\right).$$

Since we have $\alpha_2 = \gamma_2 - 1$ as in the case of $V_\top$, this number is equal to

$$n_2\left(1 - \frac{1}{2^{\gamma_2 - 1}}\right).$$

Therefore, a constant fraction of nodes in $V_\perp$ have degree 1 or 2 if γ_2 is a constant. Here we note that in the analysis of the MDS size, we only used for $V_\perp$ a property that every node in $V_\perp$ has degree at least 1. If most nodes in $V_\top$ were of high degree, the size of MDS would be much less. However, as mentioned above, a constant fraction of nodes in $V_\perp$ have degree 1 or 2 and thus we can only expect a reduction of a constant factor (i.e., not an exponent) of the MDS size even if we make extensive use of γ_2.

5. Discussions

In this report, we focused on bipartite networks in which all edges are directed from one side of nodes ($V_\top$) to the other side of nodes ($V_\perp$) and thus driver nodes can be selected only from $V_\top$. However, there exist other types of bipartite graphs. One typical example is a metabolic network. In this network, there exist two kinds of nodes [13], [15]: nodes corresponding to chemical reactions and nodes corresponding to chemical compounds. In this case, it is reasonable to assume also that only values (activities) of chemical reactions can be controlled because activities of chemical reactions may be modified by controlling concentrations of corresponding enzymes via knockout or overexpression of genes whereas it seems difficult to directly control concentrations of chemical compounds in a cell. However, there exist edges in both directions and thus we require a smaller number of driver nodes. In an example shown in Fig. 3, the transformed bipartite network has a complete matching and thus we need only one driver node under the model of [4]. However, if we assume that all edges are directed from left to right, there exist three unmatched nodes in the transformed network and thus we need three driver nodes under the same model. Therefore, structural controllability in bipartite networks strongly depends on directions of edges. Theoretical and simulation analyses of bi-directional bipartite networks would be more complicated than those of unidirectional bipartite networks because we should consider four degree distributions (indegree and outdegree of $V_\top$; indegree and outdegree of $V_\perp$), and thus are left as future work.

Acknowledgments J.C.N. was partially supported by MEXT, Japan (Grant-in-Aid 23700352), and T.A. was partially supported by MEXT, Japan (Grant-in-Aid 22650045). This work was also partially supported by research collaboration projects by Institute for Chemical Research, Kyoto University.

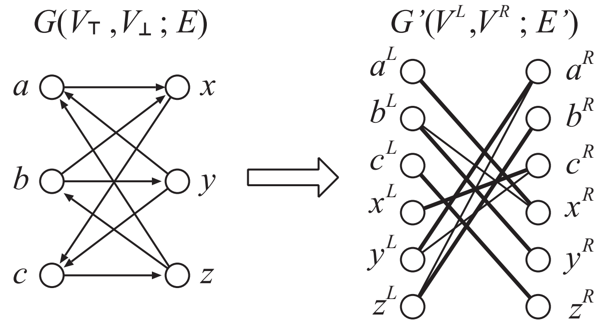

Fig. 3 Relationship between a bi-directional bipartite network (left) and its transformed network (right). In this example, $G'(V^L, V^R; E')$ has a complete matching and thus we need only one driver node.

References

[1] Cowan, N. J., Chastain, E. J., Vilhena, D. A., Freudenberg, J. S., and Bergstrom, C. T.: Nodal Dynamics, Not Degree Distributions, Determine the Structural Controllability of Complex Networks, *PLoS ONE*, Vol. 7, e38398 (2012).

[2] Haynes, T. W., Hedetniemi, S. T., and Slater, P. J.: *Fundamentals of Domination in Graphs*, Pure Applied Mathematics, Chapman and Hall/CRC, New York (1998).

[3] Jia, T., Liu, Y.-Y., Csókam E., Pósfai, M., Slotine, J.-J., and Barabàsi, A.-L.: Emergence of Bimodality in Controlling Complex Networks, *Nature Communications*, Vol. 4, 2002 (2013).

[4] Liu, Y.-Y., Slotine, J.-J., and Barabàsi, A.-L.: Controllability of Complex Networks. *Nature*, Vol. 473, pp. 167–173 (2011).

[5] Liu, Y.-Y., Slotine, J.-J., and Barabàsi, A.-L.: Observability of Complex Systems, *Proceedings of the National Academy of Sciences, USA*, Vol. 110, pp. 2460–2465 (2013).

[6] Lombardi, A. and Hörnquist, M.: Controllability Analysis of Networks, *Physical Review E*, Vol. 75, 056110 (2007).

[7] M. Mitzenmacher and E. Upfal, *Probability and Computing. Randomized Algorithms and Probabilistic Analysis*, Cambridge Univ. Press, Cambridge, MA (2005).

[8] Molnár Jr., F., Sreenivasan, S., Szymanski, B. K., and Korniss, G.: Minimum Dominating Sets in Scale-Free Network Ensembles, *Scientific Reports*, Vol. 3, 1736 (2013).

[9] Nacher, J. C. and Akutsu, T.: Dominating Scale-free Networks with Variable Scaling Exponent: Heterogeneous Networks Are Not Difficult to Control, *New Journal of Physics*, Vol. 14, 073005 (2012).

[10] Nacher, J. C. and Akutsu, T.: Analysis on Controlling Complex Networks Based on Dominating Sets, *Journal of Physics: Conference Series*, Vol. 410, 012104 (2013).

[11] Nacher, J. C. and Akutsu, T.: Structural Controllability of Unidirectional Bipartite Networks, *Scientific Reports*, Vol. 3, 1647 (2013).

[12] Nepusz, T. and Vicsek, T.: Controlling Edge Dynamics in Complex Networks, *Nature Physics*, Vol. 8, pp. 568–573 (2012).

[13] Smart, A. G., Amaral, L. A. N., and Ottino, J. M.: Cascading Failure and Robustness in Metabolic Networks, *Proceedings of the National Academy of Sciences, USA*, Vol. 105, pp. 13223–13228 (2008).

[14] Stojmenovic, I., Seddigh, M., and Zunic, J.: Dominating Sets and Neighbor Elimination-based Broadcasting Algorithms in Wireless Networks, *IEEE Trans. Parallel Distributed Systems*, Vol. 13, pp. 14-25 (2002).

[15] Takemoto, K., Tamura, T., Cong, Y., Ching, W-K., Vert, J-P., and Akutsu, T.: Analysis of the Impact Degree Distribution in Metabolic Networks Using Branching Process Approximation, *Physica A*, Vol. 391, 379–387 (2012).

[16] Wang, W-X., Ni, X., Lai, Y-C., and Grebogi, C.: Optimizing Controllability of Complex Networks by Minimum Structural Perturbations, *Physical Review E*, Vol. 85, 026115 (2012).

[17] Yan, G., Ren, J., Lai, Y-C., Lai, C-H., and Li, B.: Controlling Complex Networks: How Much Energy is Needed ? *Physical Review Letters*, Vol. 108, 218703 (2012).