A novel structure-based control method for controlling complex large scale nonlinear dynamical networks

Wei-Feng Guo[a], Shao-Wu Zhang[a,*], Tao Zeng[b], Yan Li[a], Jianxi Gao[c,*], Luonan Chen[b,d,*]

[a] *Key Laboratory of Information Fusion Technology of Ministry of Education, School of Automation, Northwestern Polytechnical University, Xian, 710072, China;*
[b] *Key Laboratory of Systems Biology, Shanghai Institutes for Biological Science, Chinese Academy Science, Shanghai 200031, China;*
[c] *Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA;*
[d] *Collaborative Research Center for Innovative Mathematical Modeling, Institute of Industrial Science, University of Tokyo, Tokyo 153-8505, Japan.*

Abstract

Exploring large scale complex systems requires the concepts and approaches delivered by structure-based network control, which investigates the control objective of complex networks through minimum number set of input nodes. However most of the existing network control methods focus on the controllability and ignore the exist of multiple input nodes for the actual control in complex large scale nonlinear dynamic networks. Considering that the selection of input nodes may depend on algorithms and network representations,we design and implement our algorithm focusing on nonlinear control of undirected networks (NCUA) for finding multiple input nodes configurations in large scale nonlinear networks with symmetric edges.The main idea of NCUA is to evaluate the controllability and actual control for driving the undirected network from the undesired attractor to the desired attractor, which is different from the traditional linear network control approaches. We valiate our NCUA algorithm in two main respects. One is that we apply our NCUA to both synthetic networks and real-world networks to investigate how the network parameters, such

*Corresponding author
Email addresses: zhangsw@nwpu.edu.cn (Shao-Wu Zhang), gaoj8@rpi.edu (Jianxi Gao), lnchen@sibs.ac.cn (Luonan Chen)

as the scaling exponent and the degree heterogeneity, affect the controllability of networks with nonlinear dynamics. Another respect is to apply the NCUA to analyze the actual control of the patient-specific molecular networks corresponding to patients across multiple datasets from The Cancer Genome Atlas (TCGA). The experimental results demonstrate the advantages of the nonlinear control method to characterize and quantify the patient-state change over the other state-of-the-art control methods. Thus, our model opens a new way to control the undesired transition of cancer states and also provides a powerful tool for theoretical research on structure based network control.
Keywords: `Structure based control`, Nonlinear dynamics, Undirected networks, FVS, Single patient system

---

1. Introduction

From a network perspective, most of physical, social, biological and computer systems can be represented as networks[1, 2, 3, 4, 5]. Analyzing networks from the controllability viewpoint will lead us to a deeper understanding about the dynamics of complex systems, especially large-scale complex systems [6, 7, 8, 9]. Since the control process is dominated by the intrinsic structure and dynamic propagation within the system, the concepts and approaches of structure-based network control are emergently required to investigate the controllability of complex networks through a minimum set of input nodes[9, 10, 11, 12, 13, 14]. The analysis of complex systems from the structure-based control viewpoint provides a deeper understanding of the dynamics of complex large-scale biological systems [15, 16, 17, 18]. So far, the studies exploiting the structure-based control of complex networks can be mainly divided into two categories according to the styles of the network dynamics, that is, the approaches focusing on linear dynamic networks and the methods focusing on nonlinear dynamic networks. For linear dynamic networks, many researchers have developed structural control tools including the Liu's Maximum Matching Sets (MMS) based control methods[9, 10, 11, 18] and the exact control method[12] and the Min-

imum Dominating Sets (MDS) based control methods [13, 19] to identify the minimum number of input nodes that need to be controlled by external signals for the system to achieve the desired control objectives. The diagram of EC control scheme and Lius control scheme and MDS control scheme are shown in Figure S1.For the nonlinear dynamic networks, an analytical tool called a feedback vertex set based control (FC) has been shown to study the control of large networks in a reliable and nonlinear manner, where the network structure is prior-known and the functional form of the governing equations is not specified, but must satisfy some properties [20, 21]. This formalism identifies the nodes and feedback vertex sets (FVS) in networks, uniquely determining the long-term dynamics of the entire network. With such a scheme, Zanudo et al. identified the source FVS as the input nodes to control the direct networks with nonlinear dynamics [14]. In fact, recently Bao et al. present an algorithm to compute and evaluate the critical, intermittent, and redundant vertices for controlling direct networks under the FC framework[22]. However, they still cannot find multiple input nodes configurations and the selection of input nodes configurations in FC is still a big challenge.

In fact, the selection of input nodes configurations in FC may depend on algorithms and input network representations (i.e. direct networks or undirected networks). Finding multiple sets of input nodes in direct networks with nonlinear dynamics is much more difficult and is still an unsolved problem. However we can solve the multiple solution discovery problem by considering the special symmetric edges in undirected networks. In this paper, we first formalize the nonlinear control problem of undirected networks (NCU), that is, how to choose the proper input nodes to drive the network from one attractor to a desired attractor in the networks with nonlinear and undirected dynamics.The NCU focus on the control and controllability of large scale nonlinear dynamic networks with symmetric edges information. To solve this problem, we developed a novel graphic-theoretic algorithm (NCUA) to measure the controllability and the actual control of undirected networks based on the feedback vertex sets. Specifically, (i) we assume that each edge in a network is bidirectional which

forms a feedback loop; (ii) we construct a bipartite graph from the original undirected network, in which the nodes of the top side are the nodes of the original graph and the nodes of the bottom side are the edges of the original graph (Figure 1 (b)); (iii) we adopt an equivalent optimization procedure for determining the MDS of the top side nodes to cover the bottom side nodes in the bipartite graph that can control the whole network using mathematical terms; and (iv) we apply random Markov chain sampling to obtain the distribution of the input nodes set and uncover the possible sets of the input nodes to control the undirected network.

Since most real world networks have a statistically significant power-law distribution, we generally have defined the control characteristics as the fraction of identified minimum input nodes and applied NCUA for multiple synthetic scale-free (SF) networks and real-world networks, and obtained several counter-intuitive findings: i) the fraction of input nodes in the network increases when the degree exponent value increases for fixed average degree, indicating that control characteristics is affected by degree heterogeneity; ii) new degree heterogeneity is defined and the fraction of input nodes decreases monotonically when degree heterogeneity becomes larger for fixed average degree. Furthermore, the degree heterogeneity and the average degree determine the minimum number of control input nodes; iii) the set of input nodes tends to be highly target-connected nodes, whereas the previous linear control study suggested that driver nodes tend to avoid high-degree nodes [9, 10, 11, 12].

We also investigated the network transition between the disease state and normal state identifiable with the stable network states (dynamical attractors) in personalized patient networks. For each sample of each cancer patient from 10 kinds of cancer sites in TCGA, we constructed a personalized differential network between the normal state and disease state, and applied the NCUA for finding their key control genes on pathologic phenotype transitions. We found that (i) although most of the cancer samples have a similar nonlinear controllability, the determining control genes still differ for different cancer samples; (ii) we identified the controllability of the reconstructed individual networks

for single samples across 10 cancer datasets, and we found the high confidence cancer-specific key genes have significant enrichments in the cancer genes census (CGC) set and the FDA-approved drug target genes (DTG) set. Compared with the traditional control model of linear networks including the EC linear control scheme [12] and Liu linear control scheme[9], our results imply that a single-patient system in cancer may be more controllable than predicted on linear dynamical networks due to the ubiquity of the nonlinear features in individual patient system. In contrast to another model on the network control of undirected networks called MDS[13], our NCUA also showed a higher performance in identifying the cancer-specific key control genes in the CGC and DTG, which were underestimated by the MDS. In conclusion, our model provides a new powerful tool for theoretical and empirical study of structure based network control.

2. Methods

2.1. Formulation of the NCU

Network dynamics are commonly nonlinear, especially at the level of nodes or small groups of nodes in the network [23]. Here, we focus on the nonlinear control problem of undirected networks. Given an undirected network *G (V, E)*, we generally consider the following broader class of the model [26] to be the following:

$$dx_i/dt = F_i(x_i, x_{I_i}) + B_i u, i = 1, ..., N \tag{1}$$

where x_i denotes the state variable of the i-th node at time t. The set I_i is a set of neighborhood nodes of node i; $\mathbf{B}_i \in R^{N*N_C}$ characterizes the driving by the NC controllers with the network. $F_i(x_i, x_{I_i})$ shows the enhancement of the activity of node i. We formalize the concept of the Nonlinear Control of the Undirected networks (NCU), which is how we chose proper sets of input nodes that are injected by input signal with the minimum cost to control the above equation (1) from an initial attractor to a desired attractor. Our NCU

not only evaluate the ability of controlling the undirected networks but also choose the proper set of input nodes instead of only choosing one random of the multiple sets of input nodes for controlling the undirected networks with nonlinear dynamic. Therefore, our NCU is a more practical problem than the existing structure based control researches of undirected networks [12, 13]. Note that the NCU control problem for direct networks with nonlinear dynamics is still an unsolved problem and need to be further studied in the future because current structure based control methods cannot find multiple sets of input nodes for controlling directed networks with nonlinear dynamics. In Figure 1 (a), we give a diagram illustration of our NCU with a simple example. (more details in Supplementary Note 1 of Additional File 1)

2.2. Algorithm for the Nonlinear Control of an Undirected Network (NCUA)

In many complex systems, there is adequate knowledge of the underlying wiring diagram, but not of the specific functional forms [23]. Analyzing such complicated systems requires concepts and approaches of structure-based control, which investigates the controllability of complex networks through a minimum set of input nodes. The traditional structure-based control methods, such as the Structural Controllability-based Control [9, 10, 11, 12] and MDS-based Control [13, 19] and Feedback Control (FC) [20, 21] focus on the structure of complex systems with linear dynamics, and may not always match the meaning of control in real systems, where control tends to involve only naturally-occurring system states. In past decades, the focus of network control research has shifted from linear dynamics to nonlinear dynamics [23, 24, 25, 26, 27]. Only one of these methods, namely the feedback vertex set control (FC) [20, 21] , can be reliably applied to large scale complex networks in which the structure is well known and the functional form of the governing equations is not specified but must satisfy that (i) is the continuous differentiability of $F_i(x_i, x_{I_i})$, that is,$F_i(x_i, x_{I_i}) \in C^1$, and (ii) dissipativity, that is, for any initial condition $x(0)$ and for a finite time t , the dynamical state $x(t)$ is bounded by a positive constant C: $\|x_n(t)\| \leq C$; and iii) the decay condition is $\partial_1 F_i(x_i, x_{I_i}) < 0$

(more details are shown in Supplementary Note 1 of Additional file 1). To drive the state of a network to any one of its naturally occurring end states (i.e., dynamical attractors), FC needs to manipulate a set of nodes (i.e., the FVS) that intersects every feedback loop in the network. Mochizuki et al. have mathematically proved that, for a direct network with the nonlinear dynamics of Equation (1), the control action of overriding the state variables of the FVS ensures that the network will asymptotically approach the desired dynamical attractor, regardless of the specific form of the functions. For this FC, it requires only a few conditions (continuous, dissipative, and decaying) on the nonlinear functions that are very mild and satisfied by a wide range of nonlinear dynamical systems [20, 21] . In fact, for a directed network, the FVS can be calculated by an algorithm which selects flip-flops in partial scan designs to break all feedback cycles (more details in Supplementary Note 1 of Additional File 1) [28]. Zanudo et al. applied the FC to study dynamic models of direct networks, illustrating that the controllability is determined by the FVS and the source nodes of the network[14]. Note that the minimum FVS in undirected network G must exist but may have multiple solutions under our assumption that each edge in undirected network is a feedback loop. Therefore our NCU aims to find multiple solutions of minimum input nodes. However, the above FVS based control algorithms cannot search multiple possible proper sets of input nodes for controlling the networks. That is, we still lack an analytical framework for the NCU problem. Therefore to solve the above proposed NCU, we developed a novel algorithm, NCUA for discovering the possible minimum sets of input nodes in the undirected networks. shows the enhancement of the activity of node i. We formalize the concept of the Nonlinear Control of the Undirected networks (NCU), which is how we chose proper sets of input nodes that are injected by input signal u with the minimum cost to control the above equation (1) from an initial attractor to a desired attractor. Our NCU not only evaluate the ability of controlling the undirected networks but also choose the proper set of input nodes instead of only choosing one random of the multiple sets of input nodes for controlling the undirected networks with nonlinear

dynamic. Therefore, our NCU is a more practical problem than the existing structure based control researches of undirected networks [12, 13]. Note that the NCU control problem for direct networks with nonlinear dynamics is still an unsolved problem and need to be further studied in the future because current structure based control methods cannot find multiple sets of input nodes for controlling directed networks with nonlinear dynamics. In Figure 1 (a), we give a diagram illustration of our NCU with a simple example. (more details in Supplementary Note 1 of Additional File 1)

The definition of FVS is a subset of nodes in the graph, such that the removal of the set leaves the graph without cycles. Therefore our NCUA searches the possible minimum nodes set whose removal leaves the graph without bi-directed edges in undirected networks (each bi-directed edge is considered as a feedback loop). This algorithm consists of three main steps: (i) constructing a bipartite graph from the original undirected network, in which the nodes of the top side are the nodes of the original graph and the nodes of the bottom side are the edges of the original graph, shown in Figure 1 (b), (ii) determining the MDS of the top side nodes to cover the bottom side nodes in the bipartite graph by using Integer Linear Programming (ILP), and (iii) designing random Markov chain sampling to obtain different input node sets. In Figure 1 (b), we give a diagram to illustrate the process of our NCUA for discovering the possible input nodes. The input of our NCUA algorithm is an undirected interaction network and the output is the number of input nodes and all the possible sets of input nodes for controlling the undirected networks. The details of the NCUA are introduced below.

Constructing a bipartite graph from the original undirected network. For a given undirected network $G(V, E)$, we assume that each edge is bi-direct and convert $G(V, E)$ into a bipartite graph $G(V_{\mathrm{T}}, V_{\perp}, E_1)$, where $V_{\mathrm{T}} \equiv V$ and $V_{\perp} \equiv E$. If $v_i \in V_{\mathrm{T}}$ is 1 of the nodes for $v_j \in V_{\perp}$, we add 1 edge e_1 connecting $v_i \in V_{\mathrm{T}}$ and $v_j \in V_{\perp}$ into the set E_1 .

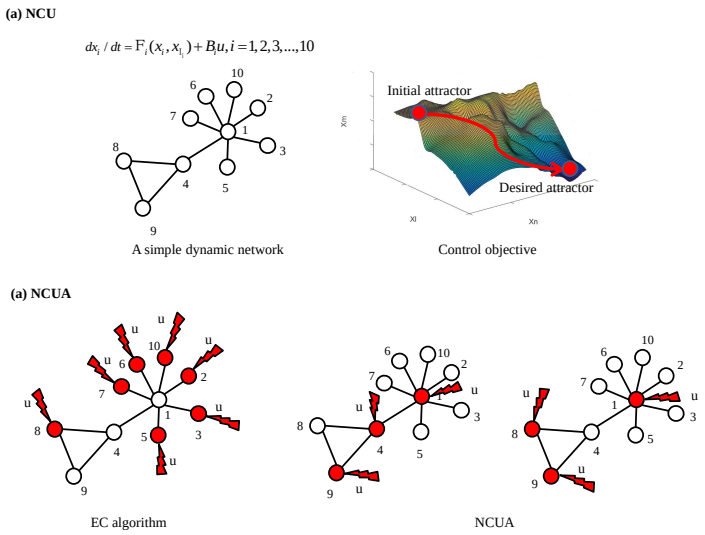

Figure 1: A schematic diagram illustrating NCU and our NCUA algorithm.(a) NCU:Our NCU not only evaluate the ability of controlling the networks but also choose the proper set of input nodes instead of only choosing one random of the multiple sets of input nodes for controlling the large scale nonlinear dynamic with symmetric edges.For this simple network, by controlling the three minimum feedback vertex nodes$\{v_1, v_4, v_9\}$ or $\{v_1, v_8, v_9\}$and ensuring that the removal of the three nodes leaves the graph without cycles or edges, the system is guaranteed to be controllable from initial attractor to desired attractor. By using EC linear control, they identified one random set $\{v_1, v_2, v_3, v_5, v_6, v_7, v_9, v_{10}\}$ to control the network from initial state to desired state. (b) NCUA: We assume that each edge in a network is bi-directional and construct a bipartite graph from the original undirected network, in which the nodes of top side are the nodes of original graph and the nodes of the bottom side are the edges of the original graph. Then, we adapt an equivalent optimization procedure for obtaining the initial input nodes $M_1 = \{v_1, v_4, v_9\}$ as a initial Markov chain within the top side nodes to cover the bottom side nodes in the bipartite graph that are sufficient to control the whole network with nonlinear dynamics in mathematical term. Finally, we generate a new Markov chain $M_2 = \{v_1, v_8, v_9\}$ by replacing node v_4 with node v_8 in the Markov chain M_1 which can also cover edge and generated the new Markov chain M_2 and repeat this process until the terminated condition is satisfied.The diagram of EC control scheme and our NCUA are shown in Figure S1.

Obtaining the cover set with minimum cost by using Integer Linear Programming (ILP). After we obtain the bipartite graph $G(V_{\mathrm{T}}, V_{\perp}, E_1)$, we adopt a modified version of the dominating set, in which the dominating set S must be selected from V_{T} and is also sufficient to dominate all of the nodes in $V_{\perp}$. We use a minimum dominating set cover problem for determining the nodes to control the whole network, that is, how to select a proper node set S, in which for each node $v_j \in V_{\perp}$, there exists a node $s_i \in S \subseteq V_{\mathrm{T}}$ such that $(v_j, s_i) \in E_1$. This problem can be solved by solving the following ILP model,

$$\begin{aligned} &\min \sum_{v \in V_{\perp}} x_v \\ &s.t. \sum_{\{v,u\} \in E} x_v \geq 1 (everyu \in V_{\mathrm{T}}), x_v \in \{0, 1\} \end{aligned}$$

where it will take the value $x_i = 1$ when node i belongs to the cover set; the object is to obtain the minimum number of nodes to cover set $V_{\perp}$. Although it is an NP-hard problem[29], the optional solution is obtained efficiently for moderate sizes of graphs with up to a few tens of thousands of nodes by utilizing an algorithm that uses the LP-based classic branch and bound method [12, 13, 30, 31] to determine the optimal solution.

Obtaining different input nodes by using random Markov chain sampling. Here, we define the minimum dominating nodes in the bipartite graphas a Markov chain. The state space G of the Markov Chain (MC) is the set of all the possible minimum dominating nodes of the bipartite graph. The different MCs need to be sampled from the state space G so that a random MC method is used. The MC approach samples different sets of minimum dominating nodes randomly. Thus, the MC method can give different sets of input nodes to control the undirected network with nonlinear dynamics. The basic idea of the MC method is to build a Markov Chain whose states are collections of the minimum dominating nodes in the top nodes covering the bottom nodes in the bipartite graph $G(V_{\mathrm{T}}, V_{\perp}, E_1)$, as shown in Figure 1 (b). The MC method is described as follows:

Initialization:. By using ILP, obtain the initial Markov Chain M_0.

Iteration:. By using ILP, obtain the initial Markov Chain M_0. For $t = 1, 2, ,$ obtain M_{t+1} from M_t as follows:

- Choose a node w uniformly at random in M_t. Then, delete node w and add a new node which can cover the edges connected by node w in the bipartite graph $G(V_{\mathrm{T}}, V_{\perp}, E_1)$. A new Markov Chain $M_{t+1} = M_t - \{w\} + \{v\}$ has been obtained.

- Accept the new Markov Chain M_{t+1} randomly.

We terminate the procedure of the MC sampling when the absolute percentage error, $MAPE = \sum_i |\phi_t(i) - \phi(i)| / n$, is less than the tolerance error$\tau = 0.01$, where $\phi_t(i)$ is the frequency of node i at time t, $\phi(i)$ is the average frequency of node i in the process of sampling, and n is number of nodes, satisfying $0 < \phi_i < 1$. Otherwise, the search process is terminated if the iteration time exceeds the fixed default value $N_{max} = 10,000$.

Complexity analysis of the NCUA. The NCUA consist of three steps: the first step is to convert the undirected network $G(V, E)$ into a bipartite graph $G(V_{\mathrm{T}}, V_{\perp}, E_1)$, where the top nodes V_{T} and the bottom nodes $V_{\perp}$ are the nodes and the edges (feedback loop) in the undirected network, respectively; the second step is to obtain the minimum set among the top nodes to cover the bottom nodes in the bipartite graph $G(V_{\mathrm{T}}, V_{\perp}, E_1)$, where $V_{\mathrm{T}} \equiv V$ and $V_{\perp} \equiv E$; the third step is to obtain different input nodes by using random Markov Chain sampling. The computational complexity of our NCUA method stems from the second part and the third part. In the phase of the minimum set cover problem, we utilize a branch-and-bound algorithm on the bipartite graph, an automatic method with the computational complexity $O(\log \|V + E\|)$[36]; in the phase of the random Markov Chain sampling with N_m times, we employ a greedy algorithm on the bipartite graph with the computational complexity, where $N_{\mathrm{T}} \equiv \|V\|$ is the network size, and N_D is the number of the input nodes. In fact, the computational complexity $O(N_m * (N - N_D))$ can be approximately considered as $O(N_m * N)$. Therefore, the overall complexity of our NCUA

approach is $O\left(\log \|V+E\|\right)+O\left(N_m * N\right)$.

3. Results

Controllability of the SF network revealed by the NCU in synthetic networks

In order to evaluate the control characteristics of the NCU, we applied our NCUA to the synthetic SF networks generated by the static model [9, 32] (more details are listed in Supplementary Note 1 of Additional File 1). We assumed the degree distribution of the undirected network $G(V,E)$ follows$P(k) \propto k^{-\gamma}$. We first defined the fraction of the input nodes $n_d = \frac{\|S\|}{\|V\|}$, where $\|S\|$ denotes the set of input nodes to control the whole network and $\|V\|$ denotes the number of connected nodes in the network. Then, we applied our NCUA to estimate the minimum number of input nodes to control the networks with nonlinear dynamics. For a given γ and average degree $< k >$, 100 networks of 10,000 nodes were constructed. The results of the NCUA were averaged over all realizations. We list the numerical results of our NCUA for the synthetic networks in Figure 2.

In fact, we plotted the NCU size as a function of the degree exponents and the average degrees and list the results in Figure 2 (a-c). In Figure 2(a), we show that for $\gamma < 2$, the number of input nodes increases as γ increases, while the number of input nodes does not depend on the average degree $< k >$. However, if the value of γ is above 2, the number of input nodes is governed by both γ and $< k >$. Furthermore, SF networks with a large value of γ above 2 or large value of $< k >$ are hard to control, as shown in Figure 2(a, c). These results are complemented by Figure 2 (b-c), where it shows that, compared with the Erdos-Renyi random networks (ER), only a few nodes are needed to control the entire network if the power law degree exponent γ is smaller than or around 2, whereas it is more difficult to controlled with a value of γ above 2. This result gives insight into which SF network will be easier to control with the minimum number of input nodes. To more clearly visualize the impact of the network structure on the number of input nodes, we plot the NCU size as a

function of the network degree heterogeneity for fixed $< k >$ in Figure 2(d). We observe that for the fixed average degree $< k >$, the NCU size decreases as the degree heterogeneity defined in Methods increases. These results illustrate that heterogeneous networks are not difficult to control, which is opposite to the conclusions of the EC linear control scheme [12] and Liu's linear control scheme[9]; however, these results are in agreement with the results of the MDS control scheme [13].In Figures 2(e) and (f), S3, and S4, we list the number of input nodes in the function of the network size for a fixed degree exponent with $\gamma = 1.4, 1.6, 1.8$, and $\gamma = 2.4, 2.6, 3.4, and 3.6$. We find that the number of input nodes decreases with the increasing network size for $\gamma < 2$, while for $\gamma > 2$, the number of input nodes are not affected by the network size.

Counterintuitive findings of the controllability from the NCU on real-world networks

We collected 17 networks with 11 categories, which were chosen for their diversity in applications and scopes (Additional File 2). By calculating the p-value of the Kolmogorov-Smirnov goodness-of-fit statistic [33], whose results are listed in Table S3 of Additional File 1, we found that the above networks are significantly subject to the power-law distribution; the detailed results are shown in Supplementary Note 5 of Additional File 1. First, we focused on how the number of input nodes correlates with the topological features, such as the scaling exponent in the power-law degree distributions and their average degree. The results of the application of the NCU on real-world networks are shown in Figure S5 and Figure 3. In Figure 3(a), we show the scatter plot of the number of input nodes as a function of the degree exponents of SF networks and their average degrees. As shown in Figure S4, we show that the number of input nodes has a tendency to increase as the exponent and the average degree increase. Furthermore, in Figure S4, we can evaluate the value of scaling exponent approximately by fitting its control characteristic to that on the synthetic networks.

We observed that the degree heterogeneity (defined in Methods) becomes

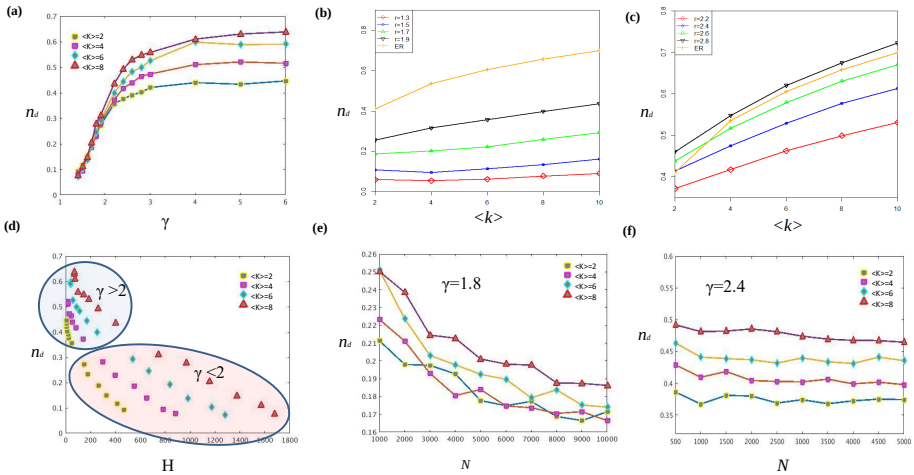

Figure 2: Nonlinear Controllability of synthetic scale free undirected networks with 10,000 nodes. (a) The number of input nodes for NCU control cost n_d in function of the average degree $< k >$ and the degree exponent γ for SF networks. All results are averaged over 10 independent realizations of the networks with 1000 nodes. (b-c) The number of input nodes for NCU control cost n_d in function of the average degree $< k >$ compared with the ER networks, for the degree exponent γ larger than 2 and less than 2 respectively. It shows that compared with the ER networks, few nodes are needed to control the entire network if the power law degree exponent γ is smaller than around 2, whereas more difficultly it is to be controlled, with larger value of γ above 2. (d) The NCU size as a function of the network degree heterogeneity at fixed $< k >$.(e-f) The number of input nodes in function of network size for different average degree exponent with $\gamma = 1.8$ and $\gamma = 2.4$ respectively

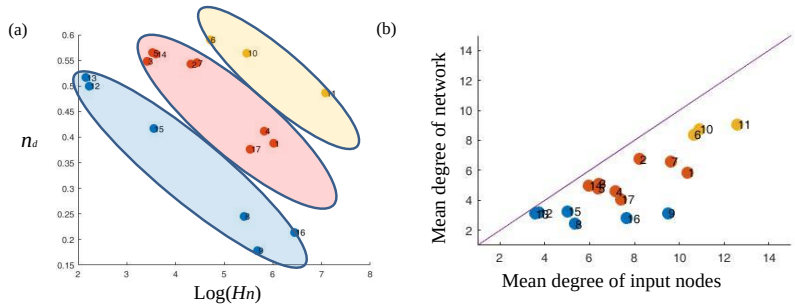

Figure 3: Results of NCU on real-world networks. (a) For each real network in Additional file 2,the scatter plot for each real network with the fraction of input nodes in NCU vs the new degree heterogeneity with eliminating the effect of network number ($n = 10,000$). Note that for the three networks with $\gamma < 2$, we estimate the value of the degree exponent by mapping the controllability on the synthetic networks with the same number of given real-world network, while for other networks with$\gamma > 2$, we estimate the value of the degree exponent by mapping the controllability on the synthetic networks with N=1000. (b) Scatter plot with the mean degree of the whole networks vs mean degree of the input nodes for each real network by using our NCUA. The bold line denotes the positions in the plot where the mean degree of the whole networks is equal to that of the input nodes;

larger as the number of network nodes increases (Figure S3 and Figure S4). To clearly visualize the impact of the degree heterogeneity and the average degree on the number of input nodes, we first eliminated the effect of the network number on the degree heterogeneity by using the following formula,

$$H_N = H_0 - H_{n_0} + H_N^{'}$$

where H_0 is value of the degree heterogeneity in the given network and H_{n_0} is the mean value of the degree heterogeneity in the synthetic network with same value of the scaling exponent whose values are evaluated by matching the controllability on the synthetic scale free networks and average degree and number of network nodes compared with the given network. $H_N^{'}$ is the mean value of the degree heterogeneity in the synthetic network with the number of network nodes $N = 10,000$, but with the same values of the scaling exponent and average degree compared with the given network.

We list the results of the number of input nodes in the function of the average degree and the new converted degree heterogeneity measure in Figure 3(a). As shown in Figure 3 (a), we find that networks with a lower average degree and higher degree heterogeneity are easier to control than those with a large average degree. The control characteristics of networks can be fully discriminated by the new converted degree heterogeneity and the average degree. We also find that the set of input nodes tends to highly target connected nodes, whereas the previous linear control study suggested that driver nodes tend to avoid high degree nodes, as shown in Figure 3(b) [12].

We observe that most types of biological networks (e.g., gene regulatory, PPI, and genetic networks) require the control of a smaller fraction of nodes than social networks (trust and social communication networks); the fraction of input nodes is between 10% and 30% in biological networks vs. more than 40% in social networks. These predictions match well with recent experimental results in cellular reprogramming and large scale social network experiments[34, 35]. Note that this prediction stands in contrast with those of linear control [12] on the same type of networks, and to some extent, can address the initial arguments

on network controllability[34, 36].

To ensure that our NCU is physically significant, we then focused on the required control energy and the control time to achieve control for networks with nonlinear dynamics. We applied a 3-dimensional stable nonlinear Lorenz oscillator system[37, 38] on the real-world network to control the networked system to the desired attractor $\{8.484, 8.484, 27\}$. The enhancement of the activity of node i in Equation (1),$F_i(x_i, x_{I_i})$, is defined as

$$F_i(x_i, x_{I_i}) \equiv f_i(x_i) - \sum_{j=1}^{N} l_{ij} \Gamma x_j$$

where $L = (l_{ij})$ is the Laplacian matrix determined by the network topological structure, and describes the internal coupling configuration at each node; By assuming that $F_i(x_i, x_{I_i})$ satisfies the conditions (continuous, dissipative, and decaying) of FC, we firstly apply our NCUA to identify the input nodes of the undirected real-world networks. Then we obtain the energy cost and time cost by applying the feedback controllers [37, 38] and closed-loop controller [39] on input nodes to control the networked system to the desired attractor.

Figure 4 demonstrates that the networks with a smaller number of input nodes would demand a greater energy cost for the pinning control to the real-world networks with the feedback controllers and 1 type of closed loop controllers ($\alpha = 0.5$). Note that in Figure 4, the energy cost and time cost of a given network are the average energy and time cost of different input nodes obtained by our NCUA, respectively.The details of the feedback controllers and the closed-loop controller and time cost and energy cost are shown in Methods.

Finally, we evaluated the differences between closed-loop controller and linear feedback controller on the nonlinear network control. Here, we adopt the local feedback controllers[37, 38] and closed-loop controllers [39] on the real-world networks. Figure 4 shows that closed-loop controllers demand a greater number of determining nodes, but require less control time and control energy than the traditional linear feedback controllers.

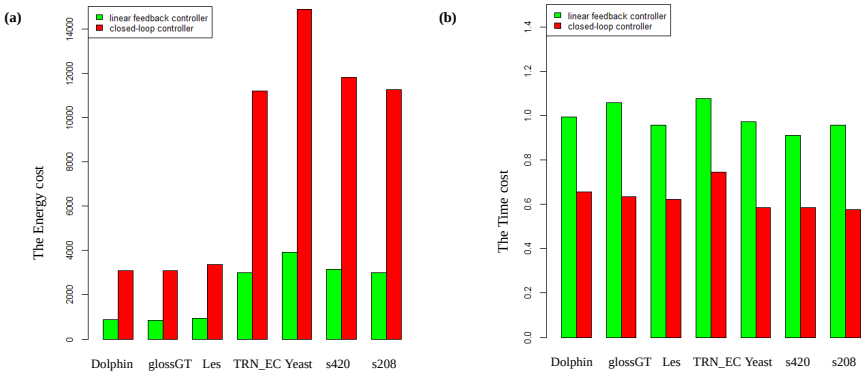

Figure 4: The compared results in terms of control energy and control time for linear feedback controller and the closed-loop controller. All the energy cost and the time cost is averaged in the process of sampling different input nodes to control the networks. (a) Result with the average energy cost of different input nodes set by using NCUA in linear feedback controller and the closed-loop controller respectively in eleven real-world networks (b) Result with the average time cost of different input nodes set by using NCUA in linear feedback controller and the closed-loop controller respectively in eleven real-world networks.

Advanced discovery of individual phenotype-transition genes in cancer samples using the NCUA

Cancer is a complex disease that generally results from a dysfunction of the relevant system or network with nonlinear dynamics, which dynamically changes with time and conditions [40, 41]. We also investigated the network transition between the disease state and normal state identifiable with the stable network states (dynamical attractors) in personalized patient networks.To further evaluate the merits of our NCU, we applied the NCUA to estimate the nonlinear controllability of a network corresponding to a single patient with cancer and to find the key control genes related with the phenotype transition from the normal state to the tumor state. Ten cancer datasets were analyzed, including the datasets for breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD), kidney chromophobe (KICH) and kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), stomach adenocarcinoma (STAD), and uterine corpus endometrial carcinoma (UCEC). According to the requirements of the NCUA, only normal-disease paired samples were obtained from TCGA data portal[47,48] (112 paired samples for BRCA, 50 paired samples for COAD, 23 paired samples for KICH, 72 paired samples for KIRC, 31 paired samples for KIRP, 50 paired samples for LIHC, 57 paired samples for LUAD, 49 paired samples for LUSC, 32 paired samples for STAD, and 23 paired samples for UCEC cancer). Each individual had paired samples (a control sample and a tumor sample). More detailed information is given in Table S1 in Additional File 1. We constructed the patient-manner interaction network by integrating the tumor and normal expression data for each patient and gene interaction network data[41, 42]. First, we chose all the normal data as the reference data and constructed the tumor network and normal network, respectively, based on the reference data with SSN method[41]. Next, we constructed the patient-manner differential expression network where the edge between gene i and gene j will exist if the p-value of the edge is less than (greater than) 0.05 in the tumor network, but greater

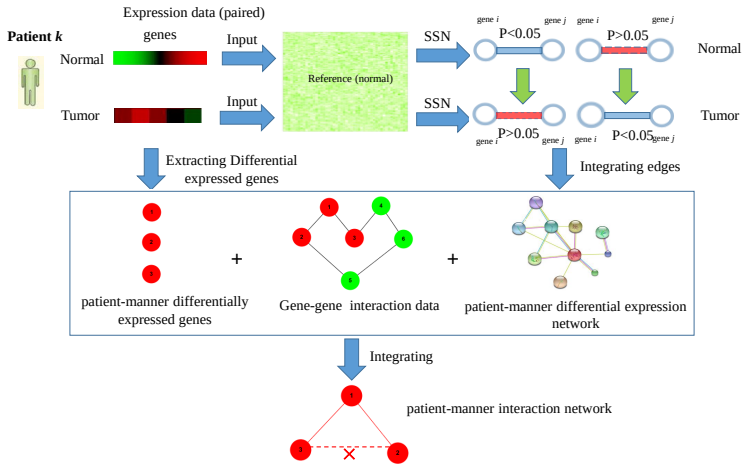

Figure 5: The scheme diagram to demonstrate the framework to construct the patient-manner interaction network by integrating the tumor and normal expression data and gene-gene interaction network; i) choose all the normal data as the reference data and construct the tumor network and normal network respectively based on the reference data with SSN method; ii) Construct the patient-manner differential expression network where edge between gene i and gene j will exist if p-value of the edge is less than (greater than) 0.05 in the tumor network but greater than (less than) 0.05 in the normal network; ii) Calculate the differential expression genes with the +/-1 log2 fold change and obtain the sub-network which consists of the differential expressed genes and where edge exist in our constructed patient-manner interaction network if edges exist in both PPI network and the patient-manner differential expression network. Note that the method for constructing network for each patient needs at least three control (normal) samples.

than (less than) 0.05 in the normal network. We calculated the patient-manner differential expression genes with the $+/-1$ log2 fold change and defined the patient-manner interaction network as the sub-network, which consists of the differential expressed genes and where the edge exists in the both gene-gene interaction network and the patient-manner differential expression network for each patient (Figure 5).

For each sample of each cancer patient from 10 kinds of cancer sites in TCGA, we constructed a personalized differential network between the normal state and disease state, and applied the NCUA for finding their key control

genes on pathologic phenotype transitions. We found that (i) although most of the cancer samples have a similar nonlinear controllability, the key control genes still differ for different cancer samples; the NCU controllability for the patient-manner interaction network in different types of cancer, that is the ability to alter the normal state and tumor state, will be much easier to control than the controllability of the network linear dynamics, including the EC linear control scheme [12] and Liu's linear control scheme[9] (Figures 6 (a) and S6 of Additional File 1). This result reveals that for the cancer patient, we only need a small fraction of genes to change the network state between the stable states, which is not applicable for controlling the biological network from initial states to any states in linear dynamics. This observation is in agreement with previous biological conclusions[34, 36]. Although Figures 6 (b) and S7 of Additional File 1 show that NCU and MDS [13] can obtain a small fraction of input nodes, our NCU can obtain more key cancer-specific genes discussed in next section.(ii) we identified the controllability of the reconstructed individual networks for single samples across 10 cancer datasets, and we found the high confidence cancer-specific key genes have significant enrichments in the cancer genes census (CGC) set and the FDA-approved drug target genes (DTG) set. Compared with the traditional control model of linear networks including the EC linear control scheme [12] and Liu's linear control scheme[9], our results imply that a single-patient system in cancer may be more controllable than predicted on linear dynamical networks due to the ubiquity of the nonlinear features in individual patient system.

The personalized key control genes in the patient-manner interaction network were further investigated using the NCUA method. In fact, the NCUA method provides a ranking of the personalized input genes according to the value of the frequencies of the genes, in which the personalized input genes are ordered by decreasing the sampling frequency in the random Markov chain sampling. We first defined the personalized key control genes as the genes that appear as the personalized input genes with a high frequency ($f > 0.6$) in the patient-manner network. Then, we calculated the frequency of the personalized

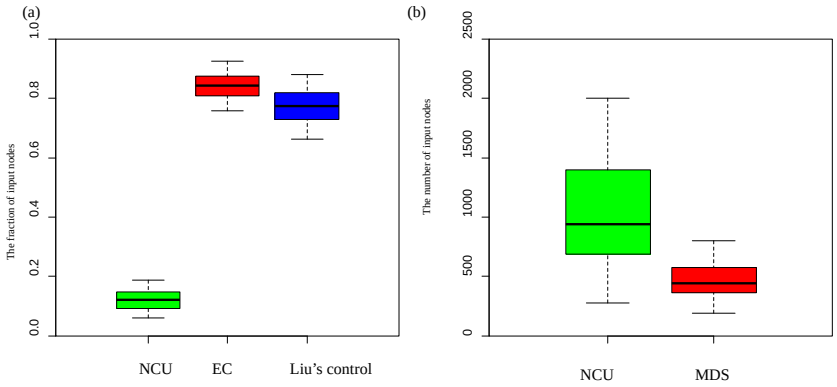

Figure 6: NCU controllability for the patient-manner interaction network in different types of cancer. (a) Box plot with the distribution of the fraction of input nodes in the NCU and linear control (EC control and Liu's control scheme) methods, respectively, for each single sample network for the STAD cancer data. Other cancer datasets used here are shown in Figure S5. (b) Box plot with the distribution of the number of input nodes in the NCU and MDS control methods, respectively, for each single sample network for the STAD cancer data. Other cancer datasets used here are shown in Figure S6.

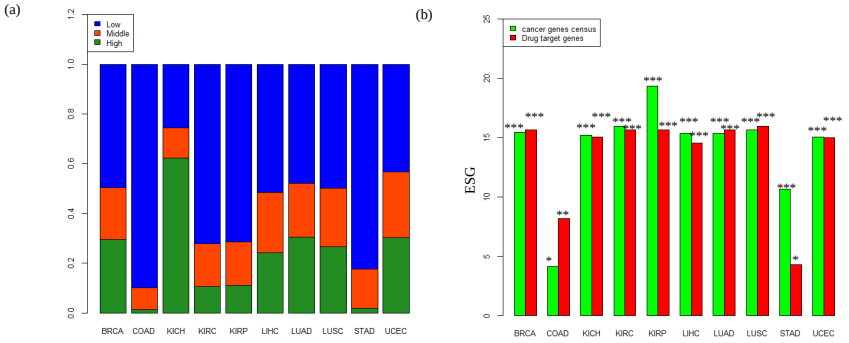

Figure 7: Applying the NCUA to discover the individual phenotype-transition genes in cancer. (b1) The frequency distribution of the key control genes in all patients of different cancers. (b2) The p-value of the high-confidence cancer specific key control genes enriching the cancer genes census set or FDA-approved drug targeted genes set.

key control genes in all patients for each cancer dataset. We defined the high confidence cancer-specific key control genes ($f > 0.6$), middle confidence cancer-specific key control genes ($0.3 < f <= 0.6$), and low confidence cancer-specific key control genes ($f <= 0.3$). The computational results of 10 cancer datasets are listed in Figure 7 (a). Figure 7 (a) shows that the low cancer-specific key control genes account for the majority, demonstrating that the key control genes vary for most of the cancer patients. The high confidence cancer-specific key control genes in each cancer dataset are in listed in Additional File 3.

Finally, we computed the p-value of the high-confidence cancer-specific key control genes enriching the cancer genes census set or FDA-approved drug targeted genes set [43, 44] by using the hyper-geometric test [45]. If the calculated p-value was less than 0.05, then we regarded that this cancer gene set is significantly enriched in the Cancer Genes Census set and FDA-approved DTG set. Figure 7 (b) shows that the high-confidence key control genes for different cancer datasets have a good enrichment in the cancer genes census set and the FDA-approved DTG set. Furthermore, we find that the set of input nodes tends

to target highly connected nodes, as shown in Figure S8 of Additional File 1. These results are in agreement with previous biological observations[34, 36].

Comapare NCUA and MDS in undirected networks

Nacher and Akutsu introduced the MDS to study the controllability of undirected networks by assuming that each node in the MDS can control all of its outgoing edges separately [13]. However, the MDS-based model assumes that more powerful control is possible (because each driver node can control its outgoing links independently), which has the possible drawback of requiring higher costs and may not be possible in many kinds of networks. Even if such powerful controllers exist, the MDS-based model still suffers from the underestimated nonlinear control of complex systems (networks). Despite its success and widespread application in searching for the important genes in the protein interaction network [15, 46, 47, 48], the MDS-based model may give an incomplete view of the undirected network control properties. In the case of a network with nonlinear dynamics, the definition of control (full control; from any initial to any final state) for the MDS-based model does not always match the meaning of control in biological, technological, and social systems, where control tends to involve only naturally occurring system states.

In this work, our control model NCU drives the whole-networked system from the initial state toward its desired dynamical attractors (e.g., the steady states and limit state cycles) by steering the input nodes to the desired dynamic attractors. Our NCU algorithm (NCUA) predicts the input nodes whose override (by an external controller or drive signals) can steer a network's dynamics toward its desired long-term dynamic behaviors (its desired dynamical attractors). Furthermore, we used the NCU control model on biological, technological, and social networks, and we identified the topological characteristics underlying the predicted node overrides. We also identified that the networks with a low average degree are easier to control than those with a large average degree, which is opposite to the previous observation from the MDS theory, as shown in Figure S8. We summarize the difference between the MDS-based method and

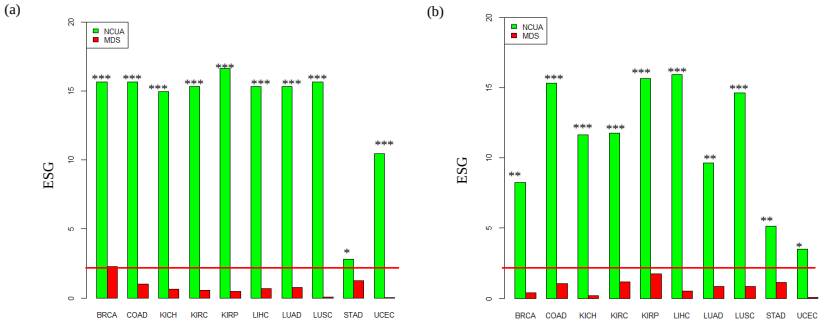

Figure 8: The advantage of the NCU method over the MDS method for finding key cancer specific genes.(a) The significant enrichment in the Cancer Genes Census set of the unique NCU input nodes (green) and the unique MDS input nodes (red) respectively. (b) The significant enrichment in the FDA-approved drug targeted genes set of the unique NCU input nodes (green) and the unique MDS input nodes (red) respectively. Note: *scores of the ESG are larger than 5, but less than 10; **scores of the ESG are larger than 10, but less than 15; ***scores of the ESG are larger than 15.

the NCU method in Table S2 of Additional File 1. The NCU and MDS methods are very different methods, so one should be careful about extending their predictions beyond their realm of applicability. In fact, in the case of network with MDS's assumption, the key nodes identified using our NCU control model can provide sufficient conditions to control the system from any initial state to any desired final steady state. For example, in Figure 1, the key nodes identified using our NCU control model dominate the nodes of the whole network for the MDS model, but the key nodes identified using the MDS model cannot cover the edges of the whole network for our NCU control model.

To further emphasize the advantage of the NCU method over the MDS method, for patient-manner interaction network we firstly obtain personalized input genes of NCU and MDS and defined NCU key cancer specific genes and MDS key cancer specific genes as all personalized input genes in cancer specific samples identified by NCU model and MDS model respectively. Then the unique

NCU key cancer specific genes (green color in Figure 8), are defined as NCU key cancer specific genes, but not of the MDS key cancer specific genes, while the unique MDS key cancer specific genes (red color in Figure 8), are defined as the MDS key cancer specific genes, but not of the NCU key cancer specific genes. Finally we provide the enrichment results from the CGC set (Figure 8 (a)) and DTG set (Figure 8 (b)) of the unique NCU key cancer specific genes and the unique MDS key cancer specific genes in the 10 cancer datasets. Figure 8 shows that the NCUA can identify the key genes in the CGC set and the FDA-approved DTG set, which are missed using the MDS method. That is, our NCU capture the nonlinear dynamic of undirected networks while MDS focus on the linear dynamic of undirected networks. Furthermore the output of NCU is all possible input nodes while the output of MDS is one random set of input nodes. Therefore the NCU model provides us with a more complete insight into the control of undirected network-based systems.

4. Conclusions

Controllability and actual control are two key issues associated with controlling large scale nonlinear dynamic networks. In the past decades, existing control methods can be divided into two main categories: one is the linear dynamic focus control approaches, which focus on the large scale linear networks and ignore the nonlinear dynamic on complex networks[9, 49, 11, 18]; another is the nonlinear control methods which evaluating the actual control and controllability of small scale networks such as boolean networks[25, 26, 27, 50] and cannot be generalized to large scale networks such as complex gene interaction network in biological fields. Recently, a new control approach based on the feedback vertex set has been proposed to find key nodes in large scale networks in which the target states are restricted to steady states. However the existing FVS based control methods only evaluate the controllability of complex networks with nonlinear dynamics and ignore the multiple input nodes exist for the actual control of complex dynamic networks. A theoretical control frame-

work is urgently required to solve the control and controllability problem in the nonlinear networks. Allowing the symmetric edges on dynamic networks, we here designed and implemented a novel and general graphic-theoretic algorithm (NCUA) from the perspective of the feedback vertex set to measure the controllability and discover the multiple minimum input nodes congurations for actual control of the nonlinear large scale networks. Instead of focusing on how to obtain the state transitions of the undirected network with linear dynamics such as the exact controllability and MDS-based model, a new concept, the nonlinear control of undirected networks (NCU), is introduced to understand how we can estimate the ability of the proper set of input nodes and how to achieve the control from the initial attractor to the desired attractor in undirected networks.

The NCUA has been evaluated on multiple synthetic SF networks and real complex networks, and it has exhibited the novel control characteristics of the undirected SF networks with nonlinear dynamics. The NCUA has also been applied to investigate the networks and their nonlinear control of cancer samples from TCGA by screening known driver genes and known drug targets as controls of their phenotype transitions, as well as to provide meaningful predictions with biological significance. Interestingly, we find that the control performance of our nonlinear control method in the single-patient system in cancer is much better than that of the traditional linear control methods, which are limited to a canonical linear time-invariant approximation. The key control genes for the individual cancer samples have significant enrichments both in the CGC set and the FDA-approved DTG set. Furthermore, it is worth exploring how to solve the NCU model with more constrained conditions (such as the target control and constrained target control[9, 10, 11]) and how to extend our method to the edge dynamics [51] to create new avenues to tackle complex systems. Note that finding multiple sets of input nodes in direct networks with nonlinear dynamics is much more difficult and is still an unsolved problem. Although the NCUA is applied to the analysis of undirected networks, we believe that, in the future, it can be extended to the analysis of directed or semi-directed networks after the implementation of a module processing technique on directed or semi-directed

networks with a network community detection algorithm from the microcosmic perspective [5, 46].

5. Acknowledgments

This paper was supported by the National Natural Science Foundation of China (61873202, 61473232, 91430111, 91439103, 91529303, 31771476, 81471047 and 11871456) and National Key R&D Program (2017YFA0505500),Strategic Priority Research Program of the Chinese Academy of Sciences (No.XDB13040700), National Key R&D Program (Special Project on Precision Medicine) (2016YFC0903400), and Natural Science Foundation of Shanghai (17ZR1446100). We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

References

[1] M. Newman, A.-L. Barabasi, D. J. Watts, The structure and dynamics of networks, Vol. 19, Princeton University Press, 2011.

[2] A.-L. Barabási, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, Nature reviews genetics 12 (1) (2011) 56.

[3] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, science 286 (5439) (1999) 509–512.

[4] D. J. Watts, S. H. Strogatz, Collective dynamics of small-worldnetworks, nature 393 (6684) (1998) 440.

[5] W.-F. Guo, S.-W. Zhang, A general method of community detection by identifying community centers with affinity propagation, Physica A: Statistical Mechanics and its Applications 447 (2016) 508–519.

[6] R. E. Kalman, Mathematical description of linear dynamical systems, Journal of the Society for Industrial and Applied Mathematics, Series A: Control 1 (2) (1963) 152–192.

[7] C.-T. Lin, Structural controllability, IEEE Transactions on Automatic Control 19 (3) (1974) 201–208.

[8] A. Lombardi, M. Hörnquist, Controllability analysis of networks, Physical Review E 75 (5) (2007) 056110.

[9] Y.-Y. Liu, J.-J. Slotine, A.-L. Barabási, Controllability of complex networks, Nature 473 (7346) (2011) 167.

[10] J. Gao, Y.-Y. Liu, R. M. D'souza, A.-L. Barabási, Target control of complex networks, Nature communications 5 (2014) 5415.

[11] W.-F. Guo, S.-W. Zhang, Z.-G. Wei, T. Zeng, F. Liu, J. Zhang, F.-X. Wu, L. Chen, Constrained target controllability of complex networks, Journal of Statistical Mechanics: Theory and Experiment 2017 (6) (2017) 063402.

[12] Z. Yuan, C. Zhao, Z. Di, W.-X. Wang, Y.-C. Lai, Exact controllability of complex networks, Nature communications 4 (2013) 2447.

[13] J. C. Nacher, T. Akutsu, Dominating scale-free networks with variable scaling exponent: heterogeneous networks are not difficult to control, New Journal of Physics 14 (7) (2012) 073005.

[14] J. G. T. Zañudo, G. Yang, R. Albert, Structure-based control of complex networks with nonlinear dynamics, Proceedings of the National Academy of Sciences 114 (28) (2017) 7234–7239.

[15] S. Wuchty, Controllability in protein interaction networks, Proceedings of the National Academy of Sciences 111 (19) (2014) 7156–7160.

[16] W.-F. Guo, S.-W. Zhang, L.-L. Liu, F. Liu, Q.-Q. Shi, L. Zhang, Y. Tang, T. Zeng, L. Chen, Discovering personalized driver mutation profiles of sin-

gle samples in cancer by network control strategy, Bioinformatics 34 (11) (2018) 1893–1903.

[17] X. Liu, L. Pan, Detection of driver metabolites in the human liver metabolic network using structural controllability analysis, BMC systems biology 8 (1) (2014) 51.

[18] W.-F. Guo, S.-W. Zhang, Q.-Q. Shi, C.-M. Zhang, T. Zeng, L. Chen, A novel algorithm for finding optimal driver nodes to target control complex networks and its applications for drug targets identification, BMC genomics 19 (1) (2018) 924.

[19] J. C. Nacher, T. Akutsu, Structural controllability of unidirectional bipartite networks, Scientific reports 3 (2013) 1647.

[20] B. Fiedler, A. Mochizuki, G. Kurosawa, D. Saito, Dynamics and control at feedback vertex sets. i: Informative and determining nodes in regulatory networks, Journal of Dynamics and Differential Equations 25 (3) (2013) 563–604.

[21] A. Mochizuki, B. Fiedler, G. Kurosawa, D. Saito, Dynamics and control at feedback vertex sets. ii: A faithful monitor to determine the diversity of molecular activities in regulatory networks, Journal of theoretical biology 335 (2013) 130–146.

[22] Y. Bao, M. Hayashida, P. Liu, M. Ishitsuka, J. C. Nacher, T. Akutsu, Analysis of critical and redundant vertices in controlling directed complex networks using feedback vertex sets, Journal of Computational Biology 25 (10) (2018) 1071–1090.

[23] Y.-C. Lai, Controlling complex, non-linear dynamical networks, National Science Review 1 (3) (2014) 339–341.

[24] M. Li, H. Gao, J. Wang, F.-X. Wu, Control principles for complex biological networksli et al. control principles for biological networks, Briefings in bioinformatics.

[25] S. P. Cornelius, W. L. Kath, A. E. Motter, Realistic control of network dynamics, Nature communications 4 (2013) 1942.

[26] A. J. Gates, L. M. Rocha, Control of complex networks requires both structure and dynamics, Scientific reports 6 (2016) 24456.

[27] S. Karl, T. Dandekar, Convergence behaviour and control in non-linear biological networks, Scientific reports 5 (2015) 9746.

[28] S. T. Chakradhar, A. Balakrishnan, V. D. Agrawal, An exact algorithm for selecting partial scan flip-flops, Journal of Electronic Testing 7 (1-2) (1995) 83–93.

[29] C. Lund, M. Yannakakis, On the hardness of approximating minimization problems, Journal of the ACM (JACM) 41 (5) (1994) 960–981.

[30] H. W. Lenstra Jr, Integer programming with a fixed number of variables, Mathematics of operations research 8 (4) (1983) 538–548.

[31] G. L. Nemhauser, L. A. Wolsey, Integer programming and combinatorial optimization, Wiley, Chichester. GL Nemhauser, MWP Savelsbergh, GS Sigismondi (1992). Constraint Classification for Mixed Integer Programming Formulations. COAL Bulletin 20 (1988) 8–12.

[32] K.-I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks, Physical Review Letters 87 (27) (2001) 278701.

[33] A. Clauset, C. R. Shalizi, M. E. Newman, Power-law distributions in empirical data, SIAM review 51 (4) (2009) 661–703.

[34] F.-J. Müller, A. Schuppert, Few inputs can reprogram biological networks, Nature 478 (7369) (2011) E4.

[35] A. D. Kramer, J. E. Guillory, J. T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks, Proceedings of the National Academy of Sciences (2014) 201320040.

[36] J. G. Zanudo, R. Albert, Cell fate reprogramming by control of intracellular network dynamics, PLoS computational biology 11 (4) (2015) e1004193.

[37] W. Huang, H. Sun, W. He, Pinning control of complex networks with general topology, in: Control and Automation (ICCA), 2010 8th IEEE International Conference on, IEEE, 2010, pp. 360–364.

[38] X. F. Wang, G. Chen, Pinning control of scale-free dynamical networks, Physica A: Statistical Mechanics and its Applications 310 (3-4) (2002) 521–531.

[39] Y.-Z. Sun, S.-Y. Leng, Y.-C. Lai, C. Grebogi, W. Lin, Closed-loop control of complex networks: A trade-off between time and energy, Physical review letters 119 (19) (2017) 198301.

[40] X. Yu, J. Zhang, S. Sun, X. Zhou, T. Zeng, L. Chen, Individual-specific edge-network analysis for disease prediction, Nucleic acids research 45 (20) (2017) e170–e170.

[41] X. Liu, Y. Wang, H. Ji, K. Aihara, L. Chen, Personalized characterization of diseases using sample-specific networks, Nucleic acids research 44 (22) (2016) e164–e164.

[42] D. Bertrand, K. R. Chng, F. G. Sherbaf, A. Kiesel, B. K. Chia, Y. Y. Sia, S. K. Huang, D. S. Hoon, E. T. Liu, A. Hillmer, et al., Patient-specific driver gene prediction and risk assessment through integrated network analysis of cancer omics profiles, Nucleic acids research 43 (7) (2015) e44–e44.

[43] P. A. Futreal, L. Coin, M. Marshall, T. Down, T. Hubbard, R. Wooster, N. Rahman, M. R. Stratton, A census of human cancer genes, Nature reviews cancer 4 (3) (2004) 177.

[44] D. S. Wishart, C. Knox, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, M. Hassanali, Drugbank: a knowledgebase for drugs, drug actions and drug targets, Nucleic acids research 36 (suppl_1) (2007) D901–D906.

[45] I. Rivals, L. Personnaz, L. Taing, M.-C. Potier, Enrichment or depletion of a go category within a class of genes: which test?, Bioinformatics 23 (4) (2006) 401–407.

[46] P. G. Sun, X. Ma, Understanding the controllability of complex networks from the microcosmic to the macrocosmic, New Journal of Physics 19 (1) (2017) 013022.

[47] M. Ishitsuka, T. Akutsu, J. C. Nacher, Critical controllability in proteome-wide protein interaction network integrating transcriptome, Scientific reports 6 (2016) 23541.

[48] P. G. Sun, Controllability and modularity of complex networks, Information Sciences 325 (2015) 20–32.

[49] T. Jia, A.-L. Barabási, Control capacity and a random sampling method in exploring controllability of complex networks, Scientific reports 3 (2013) 2354.

[50] L.-Z. Wang, R.-Q. Su, Z.-G. Huang, X. Wang, W.-X. Wang, C. Grebogi, Y.-C. Lai, A geometrical approach to control and controllability of nonlinear dynamical networks, Nature communications 7 (2016) 11323.

[51] T. Nepusz, T. Vicsek, Controlling edge dynamics in complex networks, Nature Physics 8 (7) (2012) 568.