SCIENTIFIC REPORTS

OPEN

Control efficacy of complex networks

Xin-Dong Gao[1], Wen-Xu Wang[1,2] & Ying-Cheng Lai[3,4]

Controlling complex networks has become a forefront research area in network science and engineering. Recent efforts have led to theoretical frameworks of controllability to fully control a network through steering a minimum set of driver nodes. However, in realistic situations not every node is accessible or can be externally driven, raising the fundamental issue of control efficacy: if driving signals are applied to an arbitrary subset of nodes, how many other nodes can be controlled? We develop a framework to determine the control efficacy for undirected networks of arbitrary topology. Mathematically, based on non-singular transformation, we prove a theorem to determine rigorously the control efficacy of the network and to identify the nodes that can be controlled for any given driver nodes. Physically, we develop the picture of diffusion that views the control process as a signal diffused from input signals to the set of controllable nodes. The combination of mathematical theory and physical reasoning allows us not only to determine the control efficacy for model complex networks and a large number of empirical networks, but also to uncover phenomena in network control, e.g., hub nodes in general possess lower control centrality than an average node in undirected networks.

A frontier area of research in network science and engineering is controlling complex networks[1–17]. Nearly two decades of efforts have resulted in tremendous advances[18–23] in our understanding of complex networked dynamical systems, beginning from the discoveries of the small-world[24] and scale-free[25] topologies in a large variety of natural, technological, and social systems. The efforts have created a knowledge foundation based on which the problem of control can be investigated. Ideally, to make controlling complex networked systems practically significant, one must consider nonlinear dynamical processes, due to the ubiquity of nonlinearity in the real world. However, to develop a general and mathematically rigorous control framework for complex networks hosting nonlinear dynamics is at present not achievable. A "stepping stone" is to consider linear dynamical processes on complex networks, an approach pioneered by Liu *et al.*[4], who developed a framework based on Lin's classic structural controllability theory (SCT)[26]. In particular, SCT answers the following question: given a complex, directed network, what is the minimal set of inputs (driver nodes) required to fully control the network in the sense that the entire network can be driven from an arbitrary initial state to an arbitrary final state in finite time? This was accomplished through a systematic methodology to find a minimum set of driver nodes to realize full control using the concept of maximum matching[27–29]. Subsequently, an exact controllability theory (ECT)[10,14] was developed based on the concept of maximum multiplicity[30] in linear algebra to identify the minimum set of driver nodes required to fully control the network. The ECT is applicable to a broader class of complex networks: weighted, directed or undirected, with or without any loop structure, etc. These efforts stimulated a great deal of interest in the linear controllability and observability framework of complex networks[5–17,31–33], addressing problems such as linear edge dynamics[5,6], energy cost of control[8,16,17], the role of nodal dynamics in network controllability[34,35].

In this paper, we address a fundamental and outstanding issue in controlling complex networks: *control efficacy*, the meaning of which can be understood and its significance can be appreciated, as follows. The SCT or ECT framework provides a solution of a minimum set of input signals to fully control any complex network. However, given an *arbitrary external control signal*, typically it is not possible to control the whole but only a part of the network. (The need to consider an arbitrary control signal lies in the fact that, for a network in the real world, the set of minimum input signals from the SCT or ECT framework may be physically or experimentally unrealizable.) In this regard, a related concept is control centrality[36] derived from the SCT, which characterizes the ability of a

[1]School of Systems Science, Beijing Normal University, Beijing, 100875, P. R. China. [2]Business School, University of Shanghai for Science and Technology, Shanghai 200093, China. [3]School of Electrical, Compute and Energy Engineering, Arizona State University, Tempe, Arizona, 85287 USA. [4]Department of Physics, Arizona State University, Tempe, Arizona 85287, USA. Correspondence and requests for materials should be addressed to W.-X.W. (email: wenxuwang@bnu.edu.cn)

single node to control a fraction of nodes in a directed network. Here we shall consider the more general situation where multiple input signals imposed on more than a single node are unable to control the entire network but but only a part of it, for broader classes of complex networks including weighted, undirected networks, with or without local loop structure. The main result is a control efficacy theorem and its proof, which provides a rigorous assessment of the role of an arbitrary set of nodes in partial control of the underlying network. In particular, for any given control inputs, our theorem gives the corresponding set of nodes in the network that are controllable. Our theorem of control efficacy can measure control centrality of a single node in terms of imposing a single external input signal on the node. The control centrality can be generally evaluated in arbitrary undirected networks with any distribution of link weights. We anticipate our finding to be practically significant for situations where the underlying network system is not fully accessible from the standpoint of delivering control signals. For example, for a social network, only a very limited set of nodes may be manipulated for control. For a neuronal network, only a small set of nodes can be perturbed externally. Our theory gives, in these realistic applications, a quantitative picture of what portion of the network may be controlled.

Results

Control efficacy of complex networks.
Consider an undirected network of N nodes described by the following linear time invariant (LTI) dynamical system[37–39]:

$$\dot{\mathbf{x}} = \mathcal{A} \cdot \mathbf{x} + \mathcal{B} \cdot \mathbf{u}, \tag{1}$$

where the vector $\mathbf{x} \equiv (x_1, \ldots, x_N)^{\mathrm{T}}$ represents the state of all nodes at time t, $\mathcal{A} = (a_{ij})$ is an $N \times N$ coupling matrix of the network with the element a_{ij} representing the weight of a directed link from node j to node i ($a_{ij} = a_{ji}$ for an undirected network), $\mathbf{u}(t)$ is the input signal of m controllers: $\mathbf{u} = (u_1, \ldots, u_m)^{\mathrm{T}}$, and $\mathcal{B}$ is the $N \times m$ control matrix with $b_{ij} \in \mathbb{R}$ representing the strength of the input signal $u_j(t)$ on node i. According to the classical control theory[40], the controllability of the system is determined completely and uniquely by the combined matrix $(\mathcal{A}, \mathcal{B})$. For an initial state $\mathbf{x}(t_0) = \overline{\mathbf{x}}$, if there exists a control input $\mathbf{u}(t)$ that can drive the system to the final state, say $\mathbf{x}(t_1) = 0$, within the finite time interval $[t_0, t_1]$, we say that the state $\overline{\mathbf{x}}$ is a controllable state of the system and denote it as $\mathbf{x}_+$. The classic Kalman rank condition[40] stipulates that the linear system Eq. (1) is controllable if and only if the $N \times Nm$ controllability matrix $\mathcal{C}_{(\mathcal{A},\mathcal{B})} = [\mathcal{B}, \mathcal{A} \cdot \mathcal{B}, \mathcal{A}^2 \cdot \mathcal{B}, \ldots, \mathcal{A}^{N-1} \cdot \mathcal{B}]$ has full rank. When the system is not fully controllable or, equivalently, when the state space is not filled entirely with the controllable states $\mathbf{x}_+$, there can still be a controllable subspace spanned by the column vectors of the Kalman controllability matrix $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$. The dimension of the controllable subspace is the rank of $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$: $R = \mathrm{rank}(\mathcal{C}_{(\mathcal{A},\mathcal{B})})$, which characterizes the control efficacy of the system.

There are two difficulties in determining the rank of the controllability matrix: (1) the task is often computationally prohibitive for large networks, and (2) the controllability matrix is typically nearly singular due to the dramatic differences among its elements, making the numerical rank computation inaccurate or even divergent. We are thus led to develop a feasible and effective method to calculate the rank R. The starting point is a non-singular linear transformation. In particular, for an arbitrary matrix $\mathcal{A}$ in the system equation [Eq. (1)], there exists[30] a non-singular matrix $\mathcal{P}$ such that $\mathcal{A} = \mathcal{P} \cdot \mathcal{J} \cdot \mathcal{P}^{-1}$ or $\mathcal{P}^{-1} \cdot \mathcal{A} \cdot \mathcal{P} = \mathcal{J}$ with $\mathcal{J} = \mathrm{diag}[\mathcal{J}(\lambda_1), \mathcal{J}(\lambda_2), \ldots, \mathcal{J}(\lambda_l)]$, where $\lambda_i (i = 1, \ldots, l)$ are the distinct eigenvalues of $\mathcal{A}$ and $\mathcal{J}(\lambda_i)$ is the Jordan block matrix of $\mathcal{A}$ associated with the eigenvalue λ_i. For an undirected network, the coupling matrix $\mathcal{A}$ is diagonalizable and the matrix $\mathcal{J}$ reduces to the diagonal matrix with elements being all the eigenvalues[41].

Applying the non-singular transformations $\mathbf{y} = \mathcal{P}^{-1} \cdot \mathbf{x}$ and $\mathcal{Q} = \mathcal{P}^{-1} \cdot \mathcal{B}$, we can rewrite Eq. (1) in the following form:

$$\dot{\mathbf{y}} = \mathcal{D} \cdot \mathbf{y} + \mathcal{Q} \cdot \mathbf{u}, \tag{2}$$

where $\mathcal{D}$ is a diagonal matrix (see Method). The controllability matrix $\mathcal{C}$ for the transformed system is

$$\mathcal{C}_{(\mathcal{D},\mathcal{Q})} = [\mathcal{Q}, \mathcal{D} \cdot \mathcal{Q}, \ldots, \mathcal{D}^{N-1} \cdot \mathcal{Q}] = [\mathcal{P}^{-1} \cdot \mathcal{B}, \mathcal{P}^{-1} \cdot \mathcal{A} \cdot \mathcal{B}, \ldots, \mathcal{P}^{-1} \cdot A^{N-1} \cdot \mathcal{B}] = \mathcal{P}^{-1} \cdot \mathcal{C}_{(\mathcal{A},\mathcal{B})} \tag{3}$$

We can verify that the systems Eqs (1) and (2) possess the same degree of controllability in the sense that $\mathrm{rank}(\mathcal{C}_{(\mathcal{D},\mathcal{Q})}) = \mathrm{rank}(\mathcal{C}_{(\mathcal{A},\mathcal{B})}) \equiv \mathrm{rank}(\mathcal{C})$, i.e., the rank of the controllability matrix of the original system is equal to $\mathrm{rank}(\mathcal{C}_{(\mathcal{D},\mathcal{Q})})$, which can be calculated reliably and accurately. For an arbitrary undirected network, we can prove that $\mathrm{rank}(\mathcal{C}_{(\mathcal{D},\mathcal{Q})})$ is determined by the corresponding element values in the transformed control input matrix $\mathcal{Q} = \mathcal{P}^{-1} \cdot \mathcal{B}$ associated with the distinct eigenvalues (Method). A schematic illustration of our method to calculate the rank for undirected networks with self loops is presented in Fig. 1, and an explicit example is given in Supplementary Note 1. Our key analytic results are as follows.

Single control input. When there is only a single controller (i.e., when the control input matrix $\mathcal{B}$ is a column vector), the task of calculating $\mathrm{rank}(\mathcal{C})$ is reduced to counting the corresponding nonzero elements in the matrix $\mathcal{Q} = \mathcal{P}^{-1} \cdot \mathcal{B}$. Letting the element corresponding to the eigenvalue λ_i be x_{λ_i} in $\mathcal{Q}$, we have

$$R = \mathrm{rank}(\mathcal{C}) = \sum_{i=1}^{l} \eta(\mathcal{Q}_{\lambda_i}), \tag{4}$$

where
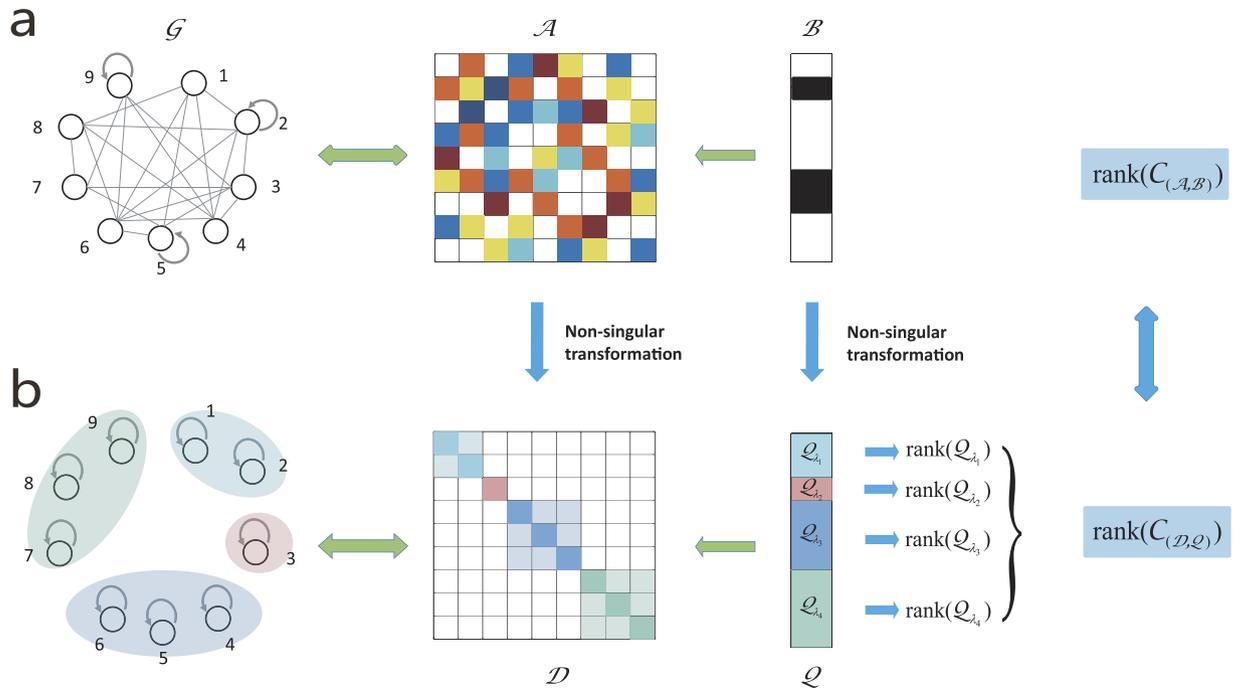
Figure 1. Schematic illustration of rank calculation for undirected networks with self-loops. (a) The adjacency matrix $\mathcal{A}$ and the control input matrix $\mathcal{B}$ for a simple undirected network with self-loops. Each colored lattice point in $\mathcal{A}$ represents an element, where the colors from white to black correspond to element values from zero to one, respectively. A similar color scheme applies to the matrix $\mathcal{B}$. **(b)** Through a nonsingular matrix transformation, the system $(\mathcal{A},\ \mathcal{B})$ is converted into the equivalent system $(\mathcal{D},\ \mathcal{Q})$, where $\mathcal{D}$ is a diagonal matrix. Distinct eigenvalues of $\mathcal{A}$ correspond to different subblocks marked with different colors. The rank of the controllability matrix for the transformed system $(\mathcal{C}_{(\mathcal{D},\mathcal{Q})})$ is equal to the sum of the rank values of the corresponding subblocks in the transformed control input matrix $\mathcal{Q}$, which is identical to the rank of the controllability matrix of the original system $(\mathcal{A},\ \mathcal{B})$.

$$\eta(\mathcal{Q}_{\lambda_i}) = \begin{cases} 1, & x_{\lambda_i} \neq 0, \\ 0, & x_{\lambda_i} = 0. \end{cases} \tag{5}$$

For each distinct, non-degenerate eigenvalue, the corresponding nonzero elements in $\mathcal{Q}$ contribute equally to the value of $\mathrm{rank}(\mathcal{C})$. For a degenerate eigenvalue, if there are no corresponding nonzero elements in $\mathcal{Q}$, the contribution of this eigenvalue to $\mathrm{rank}(\mathcal{C})$ is zero. If the corresponding elements in $\mathcal{Q}$ are nonzero, the degenerate eigenvalue contributes one to $\mathrm{rank}(\mathcal{C})$ (see Method).

Multiple control inputs. When there are m control inputs, the matrix $\mathcal{B}$ has the dimension $N \times m$. In this case, the control efficacy is determined by the sum of the rank values of the sub-block matrices composed of the corresponding rows in the transformed control input matrix $\mathcal{Q}$ for each distinct eigenvalue (see Method). We have

$$R = \mathrm{rank}(\mathcal{C}) = \sum_{i=1}^{l} \mathrm{rank}(\mathcal{Q}_{\lambda_i}), \tag{6}$$

Control Centrality. Given a complex network, it is often necessary to quantify the relative importance of the nodes with respect to a specific function. For this purpose, various kinds of centrality measures[23] were proposed in the past, such as the degree centrality, the closeness centrality[42], the betweenness centrality[43], the eigenvector centrality[44,45], and PageRank[46]. Control centrality has been defined in directed networks for quantifying the relative importance of nodes in effecting control[36], i.e., if an external driver signal is applied to a node in a directed network, how many other nodes can be controlled? Our task is to extend the definition of control centrality to undirected networks and offer a control centrality measure.

Specifically, For node i in an undirected network, its control centrality is nothing but the dimension of the controllable subspace. When a single driving signal is applied to i, the corresponding control input matrix $\mathcal{B}$ is effectively reduced to a vector $\mathbf{b}_{(i)}$ with a single non-zero element. For convenience, we set this element to be unity, let $R_{(i)}$ be the rank of the controllability matrix, and rewrite the system as
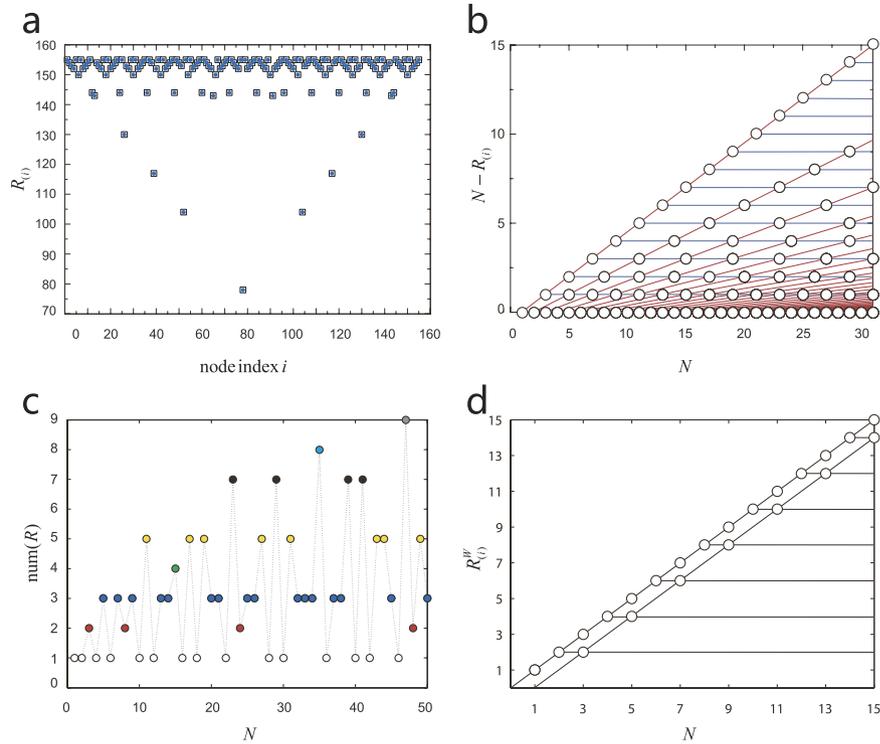
Figure 2. Control centrality of an undirected chain graph. (**a**) Nodal control centrality $R_{(i)}$ versus the node index for a one-dimensional undirected chain graph of size 155, where the squares denote the results from Eq. (4) and the blue solid circles are those from Eq. (10). (**b**) All the possible values of $N - R_{(i)}$ versus the system size N. For a fixed value of N, there are a finite number of $R_{(i)}$ values. (**c**) num(R), the number of distinct control centralities $R_{(i)}$ versus N. Each distinct value of num(R) is marked with a different color. It is remarkable that num(R) is related to the prime decomposition, as can be calculated from Eq. (11). For instance, the hollow circles represent that $N + 1$ is a prime number. (**d**) For the undirected chain graph with random weights, the corresponding control centralities $R_{(i)}^W$ versus N. According to Eq. (12), there are two periodic behavior for odd and even number of nodes alternately.

$$\dot{\mathbf{x}} = \mathcal{A} \cdot \mathbf{x} + \mathbf{b}_{(i)} u, \tag{7}$$

where u is the strength of the input signal. The value of $R_{(i)}$ can be used to characterize node i's ability to control the whole network. In Method, we provide a proof for the following inequality, which gives the upper bound of $R_{(i)}$:

$$R_{(i)} \leq \mathrm{num}(\lambda) \tag{8}$$

where num(λ) is the number of the distinct eigenvalues of the matrix $\mathcal{A}$. If $R_{(i)} = N$, then node i alone can control the whole system. However, for $R_{(i)} = 1$, node i is not able to control any other node in the networks. A value of $R_{(i)}$ between 1 and N gives the dimension of the controllable subspace of node i. To compare the control centrality in networks with different size, the normalized control centrality $r_{(i)}$ can be defined as the ratio of $R_{(i)}$ to the network size N. Then the average value, maximum and minimum values of $r_{(i)}$ are

$$\bar{r} = \sum_{i=1}^{N} \frac{R_{(i)}}{N}, \quad r_{\max} = \frac{\max(R_{(i)})}{N}, \quad r_{\min} = \frac{\min(R_{(i)})}{N}. \tag{9}$$

For the networks with random weights, the control centrality and the normalized control centrality can be denoted by $R_{(i)}^W$ and $r_{(i)}^W$ respectively.

We employ the criteria of control efficacy to explore undirected chains. To our surprise, complex phenomena associated with prime numbers emerge in the extremely simple regular network. Figure 2(a) shows, for an undirected chain graph of size $N = 155$ with identical link weights, that the values of the nodal control centrality are distributed symmetrically. For certain node (e.g., 1 or 155), the chain is fully controllable with a single input signal. For majority of the nodes, the control centrality measure is less than N. In fact, we can show analytically that the control centrality value of each node is given by (see Supplementary Note 2)

$$R_{(i)} = (N + 1) - \mathrm{GCD}(i, N + 1 - i) \tag{10}$$

Network	Eigenvalues	num(λ)	$R_{(i)}$	$R_{(i)}^W$
Star network	$0(N-2)$, $\pm\sqrt{N-1}$ (1)	3	$R_{(1)} = 2, R_{(i)} = 3, (i = 2, ..., N)$	$R_{(1)}^W = 2, R_{(i)}^W = 3, (i = 2, ..., N)$
Fully connected network	$N-1(1), -1(N-1)$	2	$R_{(i)} = 2, (i = 1, ..., N)$	$R_{(i)}^W = N, (i = 1, ..., N)$
Ring network	$2\cos\left(\frac{2\pi(i-1)}{N}\right)$	$\left[\frac{N}{2}\right] + 1$	$R_{(i)} = \left[\frac{N}{2}\right] + 1, (i = 1, ..., N)$	$R_{(i)}^W = N, (i = 1, ..., N)$
Undirected chain	$2\cos\left(\frac{\pi i}{N+1}\right)$	N	$R_{(i)} = (N+1) - \text{GCD}(i, N+1-i),$ $(i = 1, ..., N)$	$R_{(i)}^W = N$ or $N-1$

Table 1. The distinct eigenvalues, their numbers num(λ), and the nodal control centrality for regular undirected graphs. Here, $R_{(i)}^W$ is associated with random link weights and the other variables are associated with identical weights. The algebraic multiplicity of the eigenvalues is indicated for the star and fully connected networks.

where GCD(m, n) is the greatest common divisor of the positive integers m and n. Figure 2(b) shows, the distribution of the control centrality values versus the network size N. Two clusters of periodic behavior of $R_{(i)}$ present as N is increased. The periodic phenomena can be verified in terms of Eq. (10).

The combination of the two clusters of periodic behavior lead to the emergence of complex control centrality in undirected chains. Let num(R) be the number of the distinct control centrality values in a chain with a certain size. The dependence of num(R) on N is shown in Fig. 2(c). Analytically, num(R) can be determined from the following equation where, for fixed N, num(R) is the total number of all integer solutions of f_a and f_b that satisfy $f_a \cdot f_b = N+1$ (see Supplementary Note 2):

$$N + 1 = f_a \cdot f_b, \quad \left[f_a = 1, ..., \left[\frac{N}{2}\right] + 1, f_b = 2, ..., N + 1\right]. \tag{11}$$

Thus, the solution of num(R) is related with prime numbers, accounting for the complex result of num(R) in a simple chain structure. Specifically, if $N+1$ is a prime number, there is only one integer solution of Eq. (11): $f_a = 1$ and $f_b = N+1$, leading to num(R) = 1 [the hollow circles in Fig. 2(c)]. When $N+1$ is the square of a prime number, the integer solutions are $(f_a, f_b) = (1, N+1)$ and $(f_a, f_b) = (\sqrt{N+1}, \sqrt{N+1})$, accounting for num($R$) = 2, as shown by the red circles in Fig. 2(c). For num(R) > 2, the situation will become more complicated, because of the inequality of exchanging f_a and f_b. For example, if $N+1$ is the product of two different prime number, num(R) will be 3. A typical case is $N+1=6$, for which there are three integer solutions: $(f_a, f_b) = (1, 6)$, $(f_a, f_b) = (2, 3)$ and $(f_a, f_b) = (3, 2)$. However, the scenario that $N+1$ is cube of a prime number can result in num(R) = 3 as well. For instance, when $N+1 = 2^3 = 8$, there are three integer solutions: $(f_a, f_b) = (1, 8)$, $(f_a, f_b) = (2, 4)$ and $(f_a, f_b) = (4, 2)$. As a result, Fig. 2(c) exhibits rich behavior of num(R) as the length N of an undirected chain is increased.

For an undirected chain graph with random weights, the control centrality is simpler than that of a directed chain graph. We can as well offer theoretical results (see Supplementary Note 2). Specifically, when $N=2n+1$ is odd, we have

$$\begin{cases} R_{(i)}^W = N, & (i = 1, 3, 5, \cdots, 2n + 1); \\ R_{(i)}^W = N - 1, & (i = 2, 4, 6, \cdots, 2n). \end{cases} \tag{12}$$

When $N=2n$ is even, we have

$$R_{(i)}^W = N, \quad (i = 1, 2, 3, \cdots, 2n + 1). \tag{13}$$

The results are graphically shown in Fig. 2(d), which differs from the results in Fig. 2(c) with identical weights. The fact that random weights eliminates lots of linear correlations in the network matrix accounts for the simplification of the control centrality.

For undirected regular graphs with identical weights, the eigenvalues can be calculated analytically[10,47,48], so can be $R_{(i)}$, as listed in Table 1. We see that the control centrality of the star network is either 2 or 3, and that of the fully connected network is 2, regardless of the network size. For a ring network, almost half of the network can be controlled by a single input. The results for chain networks are relatively more complicated, for which the control centrality values are symmetrically distributed, which can be obtained from Eq. (10) (see Supplementary Note 2). In general, the control centrality $R_{(i)}^W$ of the undirected regular graphs with random weights is simpler than that of the undirected regular graphs with identical weights, because of the elimination of linear correlation by random weights (see Table 1).

Figure 3(a) shows the control centrality versus a key structural parameter, the connecting probability, for undirected Erdös-Rényi (ER) random networks with identical link weights. We see that, regardless of the network size, in the regime of small values of the connecting probability p, the value of $\bar{r}$ increases monotonically with p, indicating that making the network more dense can on average enhance the control efficacy. However, in the extreme regime where p is close to unity (e.g., exceeding 0.99), $\bar{r}$ begins to decrease toward zero due to the effect of identical link weights. The control centrality of ER random networks associated with random link weights differs significantly when the network becomes very dense. Specifically, Fig. 3(a) show that $\bar{r}$ is always one, as p
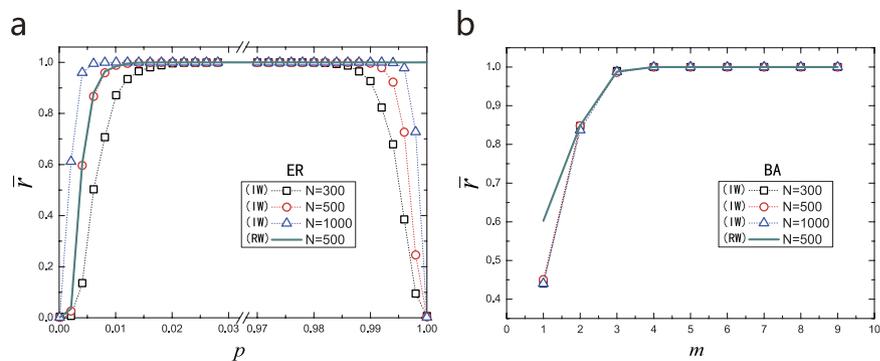
Figure 3. Control centrality of undirected model networks. (**a**) Average control centrality $\bar{r}$ versus the connecting probability p for Erdös-Rényi (ER) random networks. (**b**) $\bar{r}$ versus half of the average degree $m = \langle k \rangle / 2$ for Barabási-Albert (BA) scale-free networks. IW and RW represent identical link weights and random link weights, respectively. All the networks are undirected with symmetric coupling matrices. The data points are averaged over 50 independent network realizations. The representative network sizes are $N = 300$, 500 and 1000.

Data Sets	Nodes	Edges	$\bar{r}$	r_{max}	r_{min}	$\bar{r}^{W}$	r_{max}^{W}	r_{min}^{W}
Adjnoun[56]	112	425	0.97377	0.98214	0.97321	0.97481	0.98214	0.97321
Dolphins[57]	62	159	0.96878	0.98387	0.96774	0.97633	0.98387	0.96774
Football[58]	115	615	1	1	1	1	1	1
karate club[59]	34	78	0.69983	0.73529	0.67647	0.80536	0.82353	0.79412
Lesmis[60]	77	254	0.61899	0.63636	0.61039	0.84702	0.85714	0.84416
Netscience[56]	1589	2742	0.04060	0.16174	0.00063	0.07580	0.26306	0.00063
Polbooks[61]	105	441	1	1	1	1	1	1
Power[24]	4941	6594	0.84098	0.84739	0.80935	0.87344	0.88241	0.82372
Hep-th[62]	8361	15751	0.39915	0.57218	0.00012	0.45460	0.65184	0.00012
Email[63]	1133	5451	0.96125	0.96204	0.96028	0.96163	0.96293	0.96117
Jazz[64]	198	2742	0.96500	0.96970	0.96464	0.96740	0.96970	0.96465
USAir[65]	332	2126	0.76639	0.77410	0.76205	0.76944	0.77410	0.76506

Table 2. Control centrality of empirical networks. For each network, its size, the total number of links, the average value, its maximum and minimum values for identical link weights and random link weights, are given. The structural data of all the networks are available online (see Supplementary Table 1). Superscript W represents random weights.

approaches unity. The difference is as well attributed to the elimination of linear correlation by random weights, but such effect of random is negligible in sparse ER network.

Figure 3(b) shows, for Barabási-Albert (BA) scale free networks, $\bar{r}$ versus half of the average degree $m = \langle k \rangle / 2$, where we see that $\bar{r}$ increases rapidly toward unity as $\langle k \rangle$ is increased, regardless of the number of new links associated with the addition of a new node into the network during its growing process. We see that, qualitatively similar to ER networks, making a scale free network more densely connected can enhance its control efficacy. We also see that because of the general sparsity of the BA network, random link weights have negligible effect on $\bar{r}$ for $m \leq 2$ compared to identical link weights.

We characterize the control efficacy for a number of real world (empirical) networks. The results are listed in Table 2. (For the empirical networks with random weights, its corresponding control centrality are slightly higher than the origin network topology.) An issue is whether the hub nodes carry a stronger control centrality in undirected networks. We find that the average control centrality of hub nodes is generally smaller than that of the other nodes in undirected networks, which is consistent with the finding that driver nodes avoid hubs in directed networks[49]. To demonstrate this counterintuitive phenomenon, we divide the nodes into three groups in terms of their degrees: low, medium and high. Figure 4(a) shows, for model ER and BA networks, that the control centrality is generally higher for low-degree nodes than that for the hubs. Figure 4(b) shows the mean degree of the nodes with the maximum control centrality versus the mean degree $\langle k \rangle$ of all nodes, for each empirical network in Table 2. We see that the values of $\bar{k}_{r_{max}}$, the degree value at which maximum control efficacy is achieved, are significantly smaller than or comparable to $\langle k \rangle$, indicating the nodes with large values of control centrality are generally not hubs. To provide further evidence for the determining role of nodal degree in the control efficacy, we randomize each empirical network by converting it into an ER random network, keeping the network size N and its diameter L unchanged. As shown in Fig. 4(c), for some networks there is no correlation between the values of $\bar{r}$ for the original and randomized networks, indicating that the full randomization process has effectively eliminated any topological features of the original network that determine the control efficacy. We then apply a
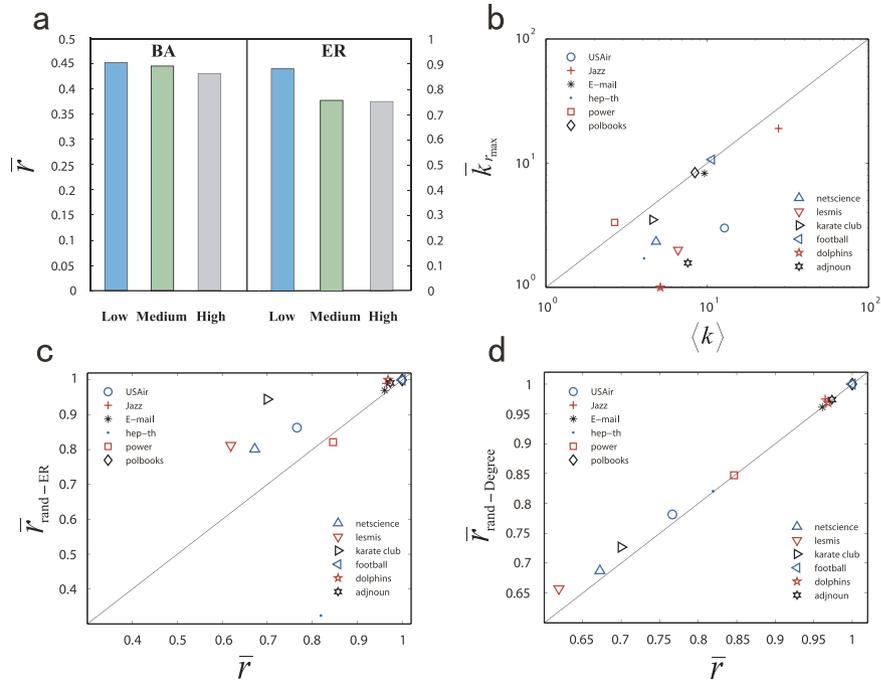
Figure 4. Role of hubs on control centrality in model and empirical networks. (**a**) The average control centrality (bars) for the low-, medium- and high-degree nodes in ER and SA networks of size $N = 500$ and average degree $\langle k \rangle = 2$, where the control centrality of hubs is generally less than that for smaller degree nodes. The results are averaged over 500 network realizations. For the ER networks, different connected components are considered separately. (**b**) Mean degree of the nodes with the maximum control centrality r_{max} as compared with the mean degree of all nodes for a number of empirical networks. It can be seen that for these real-world networks the nodes with relatively large values of the control centrality are not hub nodes, which is consistent with the results in (**a**). (**c**) For randomized empirical networks and (**d**) for the randomized networks but with the degrees preserved, the values of $\bar{r}$ in comparison with these from the original networks.

degree-preserving procedure[4,50,51] that randomly rewires the links but keeps the degree of each node unchanged. Contrary to the case of full randomization [Fig. 4(c)], when the nodal degrees are preserved, there is little change in the value of $\bar{r}$, indicating strongly that degree is the key characteristic that determines the control efficacy.

Identification of controllable nodes. For an arbitrary undirected network, given a control input matrix, we can obtain the dimension of the controllable subspace by calculating the control efficacy. An issue of practical importance is how to identify the actual set of nodes that can be controllable, i.e., the set of controllable nodes for a given control input configuration. Here, we offer a general method based on network diffusion dynamics to address this issue. Specifically, note that the $N \times Nm$ controllability matrix $\mathcal{C}_{(\mathcal{A},\mathcal{B})} = [\mathcal{B}, \mathcal{A} \cdot \mathcal{B}, \mathcal{A}^2\mathcal{B}, \ldots, \mathcal{A}^{N-1}\mathcal{B}]$ can be expressed iteratively as

$$\mathcal{C}_{(\mathcal{A},\mathcal{B})} = [\mathcal{B}, \mathcal{A} \cdot \mathcal{B}, \mathcal{A} \cdot (\mathcal{A} \cdot \mathcal{B}), \ldots, \mathcal{A} \cdot ((\mathcal{A}^{N-2}) \cdot \mathcal{B})]. \tag{14}$$

For the $N \times N$ matrix $\mathcal{A}^s$, between any pair of nodes (e.g., i and j), there exists a path of length s:

$$(\mathcal{A}^s)_{ij} = \sum_{k_1, \ldots, k_{s-1}} a_{ik_1} \ldots a_{k_{s-1}j} \neq 0.$$

Regarding the nonzero elements of $\mathcal{B}$ as sources of diffusion, the controllability matrix $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$ can be viewed as being formed by a diffusion process from the nodes with control matrix $\mathcal{B}$ to all the controllable nodes in the network in $(N-1)$ time steps, generating the corresponding diffusion mode for each column of $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$. At time step s, the matrix product $\mathcal{A}^{s-1} \cdot \mathcal{B}$ is a linear combination of the mode at the s step and the modes from all prior forward steps. The rank of $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$ is determined by the number of the distinct modes of diffusion. In general, unless $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$ has a full rank, there is interdependence among its columns. Using this fact, we can prove that, for fixed rank$(\mathcal{C}) = r$, the distinct diffusion modes are fully contained in the former r iteration steps (Supplementary Note 3). Consequently, we can implement the following elementary column transformation on the controllability matrix to obtain

$$\mathcal{C}_r = [\mathbf{b}_1, \mathcal{A} \cdot \mathbf{b}_1, \ldots, \mathcal{A}^{r-1} \cdot \mathbf{b}_1; \mathbf{b}_2, \mathcal{A} \cdot \mathbf{b}_2, \ldots, \mathcal{A}^{r-1}\mathbf{b}_2; \ldots; \mathbf{b}_m, \mathcal{A} \cdot \mathbf{b}_m, \ldots, \mathcal{A}^{r-1} \cdot \mathbf{b}_m] \tag{15}$$
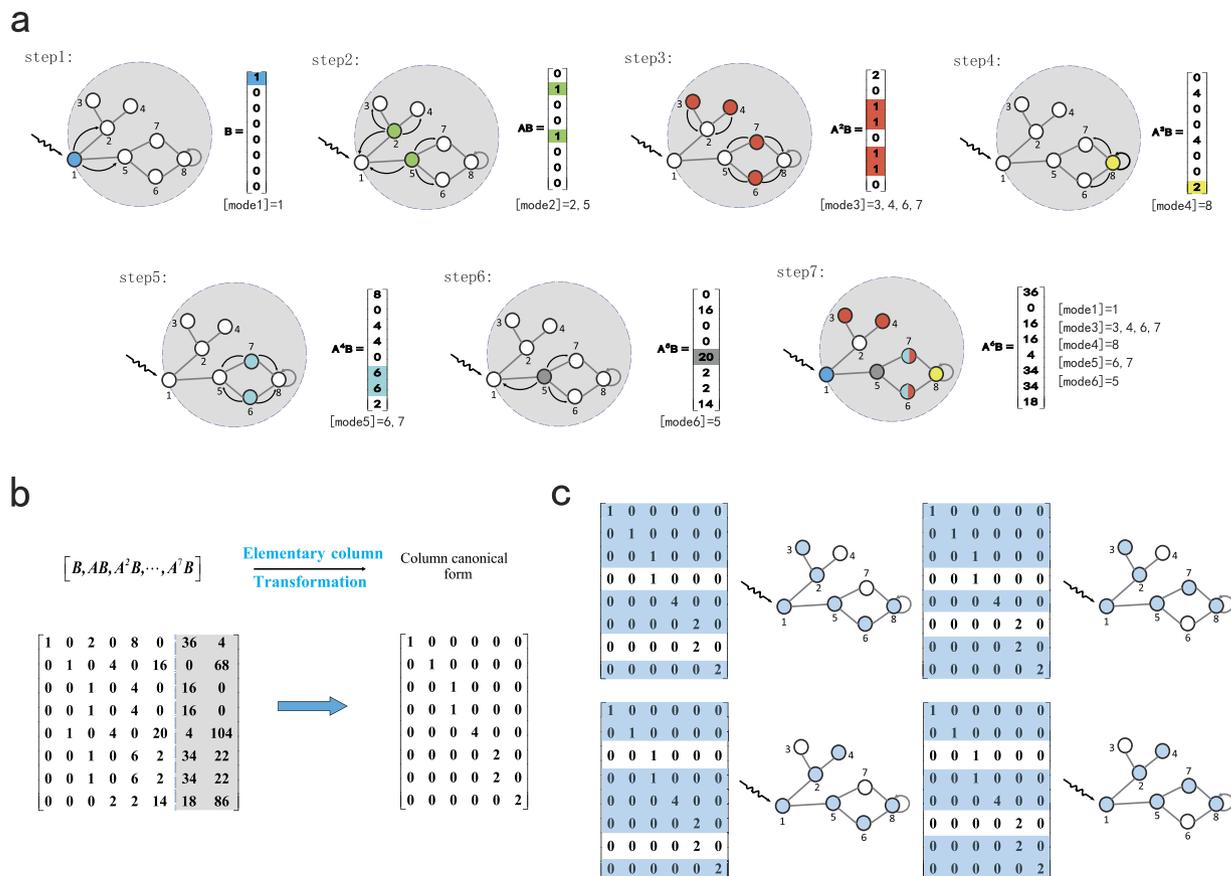
Figure 5. Illustration of our method to identify the set of controllable nodes. Take as an example a simple undirected network with self-loops. (**a**) A step by step illustration of the diffusion process over the network from the driver node, where a control signal is applied at node 1. The newly appeared mode at each step is marked with different colors. At step 7, the iteration column vector $\mathcal{A}^6\mathcal{B}$ can be expressed as a linear combination of the former modes, so the corresponding value of the control efficacy is 6. (**b**) For the controllability matrix $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$, its column canonical form generated by the elementary column transformation. For a fixed value of the control efficacy measure r, the column canonical form can be performed only for r iterations of the column vector. (**c**) There is a one-to-one correspondence between the controllable nodes and the rows that are linearly dependent upon others in the column canonical form. In the specific case shown, there are four distinct configurations of the controllable nodes (marked in blue). Nevertheless, the number of controllable nodes is fixed and solely determined by rank($\mathcal{C}_{(\mathcal{A},\mathcal{B})}$).

so that the controllable nodes correspond to the maximal linearly independent group of the rows. To illustrate the method explicitly, we present a concrete example, as shown in Fig. 5, where the diffusion process can be seen by noting the newly appeared diffusion mode (color marked) at each time step. We next perform the elementary column transformation on $\mathcal{C}$ to obtain its column canonical form that reveals the linear dependence among the rows, where the rows that are linearly independent of other rows correspond to the controllable nodes. Note that the configuration of drivers is not unique as it depends on the order with which the elementary transform is implemented. While there are many possible choices of the linearly dependent rows, the number of controllable nodes is fixed and determined solely by rank($\mathcal{C}$) = r. Our procedure of finding the driver nodes is rigorous, as guaranteed by our theory of control efficacy and the column canonical form associated with the matrix rank.

Discussion

For complex networks in the real world, from the standpoint of control not every node is externally accessible. Often, control signals can be applied to a limited set of nodes or just a few nodes. If the network structure is known, theoretically it is possible to determine a specific set of nodes to apply the control signals, e.g., through identification of maximum matching in SCT. However, the set of control nodes so determined may not overlap with the set of externally accessible nodes. Under these circumstances it is not possible to control the whole network. Nonetheless, there are situations where full control of the entire network is not necessary. A fundamental question is then, if control is applied to a few nodes or even a single node, what fraction of the network can be controlled? That is, for a complex network of arbitrary structure, what is the control efficacy or, equivalently, the dimension of the controllable subspace of the underlying network?

The issue of control efficacy (or control centrality if control is applied at a single node) was addressed in a previous work[36] but for directed networks. The contribution of the present paper is a rigorous framework based on the theory of exact controllability[10,14] to determine, for undirected complex networks of arbitrary structure (regular, random, or scale-free, weighted or unweighted, with or without self loops, etc.), their control efficacy. From the mathematical control theory, the control efficacy is given by the rank of the Kalman controllability matrix, the determination of which is computationally prohibitive for large networks. Utilizing the non-singular similarity transformation, we discovered a mathematical theorem that enables us to convert rank calculation into a counting problem in terms of the block matrices associated with the distinct eigenvalues of the network coupling matrix. The framework allows us to determine, rigorously, the control efficacy of not only model complex networks, but also a large number of real world networks. Physically, we developed the picture of diffusion, i.e., to view the control process as a signal originated from the driver node and diffused through the controllable subnetwork. The powerful combination of rigorous mathematical theorem and physical reasoning leads to the discovery of striking phenomena in controlling complex networks. For example, more densely connected networks in general have stronger control efficacy, regardless of their topology, and nodal degree is key to control efficacy. However, hub nodes in general have low values of control centrality as compared with majority of the nodes in the network.

From the perspective of fundamental science, our framework of control efficacy represents an important step forward in understanding, quantitatively, the controllability of complex networks at the detailed level of individual nodes. (Extension of our control efficacy framework to analyzing the efficacy of observability of complex networks is straightforward - see Supplementary Note 5). Practically, our theory provides a method and algorithms that can be used to identify efficiently the nodes that possess the strongest possible control centrality. This can have significant applications. For example, given a social network, our framework allows the nodes with the largest control efficacy, i.e., the nodes that can control the largest possible fraction of the network, to be identified. Similarly, for a complex infrastructure network, we can determine a small set of critical nodes to obtain maximum possible control of the network to achieve the highest possible energy efficiency.

Despite our initial success as reported in the present paper, many outstanding issues remain. For example, in real world networks the estimated link weights are not exactly known, which will lead to errors in determining the control efficacy. A mathematical uncertainty or error analysis is needed, but at the present a rigorous treatment seems difficult. Also, our framework of control efficacy relies on complete knowledge of the network structure. What if there is missing information about nodes, links and/or their weights? - at the present we do not have a theory to deal with this practically important issue. Last but not least, our entire theory is based on hypothesizing the underlying complex network as a linear and time invariant dynamical systems. Although much effort has been dedicated to controlling complex networks with nonlinear dynamics[52–54], a general approach for measuring control efficacy remains to be an outstanding problem[55]. The main challenge stems from the fact that the control efficacy is determined by both network structure and dynamics, in contrast to the network governed by linear dynamics. Much further effort is called for in the extremely rapidly developing field of controlling complex networks.

Methods

For an undirected network with arbitrary link weights [Eq. (1)], the matrix $\mathcal{A}$ is symmetric and so is diagonalizable: there exists an orthogonal matrix $\mathcal{P}$ and a diagonal matrix $\mathcal{D}$ such that $\mathcal{A} = \mathcal{P} \cdot \mathcal{D} \cdot \mathcal{P}^{\mathrm{T}} = \mathcal{P} \cdot \mathcal{D} \cdot \mathcal{P}^{-1}$ with $\mathcal{D} = \mathrm{diag}(\mathcal{D}_{\lambda_1}, \mathcal{D}_{\lambda_2}, \ldots, \mathcal{D}_{\lambda_l})$, where λ_i's $(i = 1, \ldots, l)$ are the distinct eigenvalues of $\mathcal{A}$ and $\mathcal{D}_{\lambda_i}$ is the diagonal block matrix of $\mathcal{A}$ associated with λ_i. The size of $\mathcal{D}_{\lambda_i}$ is given by the multiplicity of λ_i. We write

$$\mathcal{D}_{(N \times N)} = \mathcal{P}^{-1} \cdot \mathcal{A} \cdot \mathcal{P} = \begin{bmatrix} \mathcal{D}_{\lambda_1} & & & \\ & \mathcal{D}_{\lambda_2} & & \\ & & \mathcal{D}_{\lambda_3} & \\ & & & \ddots \end{bmatrix}, \quad \mathcal{Q}_{(N \times m)} = \mathcal{P}^{-1} \cdot \mathcal{B} = \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \\ \mathcal{Q}_3 \\ \vdots \end{bmatrix}$$

(16)

For a linear dynamical system, its controllability is invariant under any non-singular transform. The control efficacy of the original system can then be determined by calculating the rank of the transformed Kalman matrix $\mathcal{P}^{-1} \cdot \mathcal{C}_{(\mathcal{A},\mathcal{B})}$ [Eq. (3)].

Single control input. When the system is subject to a single control input, the control matrix $\mathcal{B}$ and the transformed control matrix $\mathcal{Q} = \mathcal{P}^{-1} \cdot \mathcal{B} = [q_1, \ldots, q_N]^{\mathrm{T}}$ is an $N \times 1$ column vector. If $\mathcal{Q}$ has zero element, the corresponding row in the the transformed Kalman matrix $\mathcal{P}^{-1} \cdot \mathcal{C}_{(\mathcal{A},\mathcal{B})}$ is zero. For the nonzero elements in $\mathcal{Q}$, the corresponding eigenvalues can be of two types.

(i) Case I: Distinct eigenvalues. An illustrative example for this case is shown in Fig. 1(a), where the values of q_1, q_2 and q_3 are assumed to be nonzero, corresponding to the eigenvalues λ_1, λ_2 and λ_3, respectively. The corresponding row of (q_1, q_2, q_3) in $\mathcal{P}^{-1} \cdot \mathcal{C}_{(\mathcal{A},\mathcal{B})}$ is a Vandermonde matrix, whose rows are linearly independent. In this case, the rank of the controllability matrix is nothing but the number of the nonzero elements corresponding to distinct eigenvalues of $\mathcal{A}$.

(ii) Case II: Degenerate eigenvalues. When there is eigenvalue multiplicity, the rows of $\mathcal{P}^{-1} \cdot \mathcal{C}_{(\mathcal{A},\mathcal{B})}$ are linearly dependent upon each other. An example of the controllability matrix is

$$\mathcal{C}_{(\mathcal{D},\mathcal{Q})} = \mathcal{P}^{-1}\mathcal{C}_{(\mathcal{A},\mathcal{B})} = [\mathcal{Q}, \mathcal{D}\cdot\mathcal{Q}, ..., \mathcal{Q}^{N-1}\cdot\mathcal{D}] = \begin{bmatrix} q_1 & \lambda_1 q_1 & \lambda_1{}^2 q_1 & \cdots & \lambda_1{}^N q_1 \\ q_2 & \lambda_1 q_2 & \lambda_1{}^2 q_2 & \cdots & \lambda_1{}^N q_2 \\ q_3 & \lambda_2 q_3 & \lambda_2{}^2 q_3 & \cdots & \lambda_3{}^N q_3 \\ q_4 & \lambda_3 q_4 & \lambda_3{}^2 q_4 & \cdots & \lambda_4{}^N q_4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_6 & \lambda_3 q_6 & \lambda_3{}^2 q_6 & \cdots & \lambda_3{}^N q_6 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_m & \lambda_i q_m & \lambda_i{}^2 q_m & \cdots & \lambda_i{}^N q_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_N & \lambda_l q_N & \lambda_l{}^2 q_N & \cdots & \lambda_l{}^N q_N \end{bmatrix}, \qquad (17)$$

where the rows of q_4 and q_6 are linearly dependent. If the linearly dependent rows in $\mathcal{Q}$ have nonzero elements, they together contribute one to the rank of the controllability matrix.

For a single control input, the calculation of the rank of $\mathcal{C}_{(\mathcal{A},\mathcal{B})}$ is thus equivalent to counting the corresponding nonzero elements in $\mathcal{Q}$. We have

$$R = \mathrm{rank}(\mathcal{C}_{(\mathcal{A},\mathcal{B})}) = \mathrm{rank}(\mathcal{P}^{-1}\cdot\mathcal{C}_{(\mathcal{A},\mathcal{B})}) = \sum_{i=1}^{l}\eta(\mathcal{Q}_{\lambda_i}) \qquad (18)$$

Since the control matrix $\mathcal{B}$ has a single column, the control centrality of the input node is given by

$$R_{(i)} \leq \mathrm{num}(\lambda), \qquad (19)$$

where the $\mathrm{num}(\lambda)$ is the number of the distinct eigenvalues of $\mathcal{A}$.

Multiple control inputs. With multiple control input signals, the transformed control matrix $\mathcal{Q}$ has the dimensional $N\times m$. To illustrate our method of rank calculation explicitly, we consider the first two columns in $\mathcal{Q}$. The matrices $\mathcal{D}$ and $\mathcal{Q}$ can be written as

$$\mathcal{D} = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & & \\ & & & \lambda_2 & & \\ & & & & \ddots & \\ & & & & & \lambda_2 \\ & & & & & & \ddots \end{bmatrix}, \quad \mathcal{Q} = \begin{bmatrix} q_{11} & q_{21} & \cdots \\ \vdots & \vdots & \cdots \\ q_{1l_1} & q_{2l_1} & \cdots \\ q_{1l_1+1} & q_{2l_1+1} & \cdots \\ \vdots & \vdots & \cdots \\ q_{1l_2} & q_{2l_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad (20)$$

Adjusting the order of the original column vectors appropriately, we can convert the transformed Kalman matrix $\mathcal{P}^{-1}\cdot\mathcal{C}_{(\mathcal{A},\mathcal{B})}$ into a form in which two single controller inputs are applied sequentially, i.e.,

$$\mathcal{P}^{-1}\cdot\mathcal{C}_{(\mathcal{A},\mathcal{B})} = \begin{bmatrix} q_{11} & q_{11}\lambda_1 & \cdots & q_{11}\lambda_1{}^N & q_{21} & q_{21}\lambda_1 & \cdots & q_{21}\lambda_1{}^N & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ q_{1l_1} & q_{1l_1}\lambda_1 & \cdots & q_{1l_1}\lambda_1{}^N & q_{2l_1} & q_{2l_1}\lambda_1 & \cdots & q_{2l_1}\lambda_1{}^N & \cdots \\ q_{1l_1+1} & q_{1l_1+1}\lambda_2 & \cdots & q_{1l_1+1}\lambda_2{}^N & q_{2l_1+1} & q_{2l_1+1}\lambda_2 & \cdots & q_{2l_1+1}\lambda_2{}^N & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ q_{1l_2} & q_{1l_2}\lambda_2 & \cdots & q_{1l_2}\lambda_2{}^N & q_{2l_2} & q_{2l_2}\lambda_2 & \cdots & q_{2l_2}\lambda_2{}^N & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \qquad (21)$$

For the case where the control matrix $\mathcal{A}$ has distinct eigenvalues, if certain rows of the transformed control matrix $\mathcal{Q}$ contain nonzero elements, the corresponding rows of the transformed Kalman matrix $\mathcal{P}^{-1}\cdot\mathcal{C}$ must be linearly independent of each other. The matrix $\mathcal{P}^{-1}\cdot\mathcal{C}_{(\mathcal{A},\mathcal{B})}$ can be organized into a block matrix form, where each block corresponds to one distinct eigenvalue and its dimension is the multiplicity of the eigenvalue. The rank of such a matrix is the sum of the rank values of the sub-block matrices. In particular, letting the algebraic multiplicity of the eigenvalue λ_1 be l_1, we have

$$\mathrm{rank}\begin{bmatrix} q_{11} & q_{11}\lambda_1 & \cdots & q_{11}\lambda_1{}^N & q_{21} & q_{21}\lambda_1 & \cdots & q_{21}\lambda_1{}^N & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ q_{1l_1} & q_{1l_1}\lambda_1 & \cdots & q_{1l_1}\lambda_1{}^N & q_{2l_1} & q_{2l_1}\lambda_1 & \cdots & q_{2l_1}\lambda_1{}^N & \cdots \end{bmatrix} = \mathrm{rank}\begin{bmatrix} q_{11} & q_{21} & \cdots \\ \vdots & \vdots & \ddots \\ q_{1l_1} & q_{2l_1} & \cdots \end{bmatrix}. \qquad (22)$$

In general, for multiple control inputs, the control efficacy $\mathrm{rank}(\mathcal{C}) = \mathrm{rank}(\mathcal{P}^{-1} \cdot \mathcal{C}_{(\mathcal{A},\mathcal{B})})$ is the sum of the rank values of the sub-block matrices:

$$R = \mathrm{rank}\,(\mathcal{C}) = \sum_{i=1}^{l} \mathrm{rank}(\mathcal{Q}_{\lambda_i}). \tag{23}$$

References

1. Lombardi, A. & Hörnquist, M. Controllability analysis of networks. *Phys. Rev. E* **75**, 056110 (2007).
2. Liu, B., Chu, T., Wang, L. & Xie, G. Controllability of a leader-follower dynamic network with switching topology. *IEEE Trans. Automat. Contr.* **53**, 1009–1013 (2008).
3. Rahmani, A., Ji, M., Mesbahi, M. & Egerstedt, M. Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM J. Contr. Optim.* **48**, 162–186 (2009).
4. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. *Nature* (*London*) **473**, 167–173 (2011).
5. Wang, W.-X., Ni, X., Lai, Y.-C. & Grebogi, C. Optimizing controllability of complex networks by minimum structural perturbations. *Phys. Rev. E* **85**, 026115 (2012).
6. Nepusz, T. & Vicsek, T. Controlling edge dynamics in complex networks. *Nat. Phys.* **8**, 568–573 (2012).
7. Nacher, J. C. & Akutsu, T. Dominating scale-free networks with variable scaling exponent: heterogeneous networks are not difficult to control. *New J. Phys.* **14**, 073005 (2012).
8. Yan, G., Ren, J., Lai, Y.-C., Lai, C.-H. & Li, B. Controlling complex networks: How much energy is needed? *Phys. Rev. Lett.* **108**, 218703 (2012).
9. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Observability of complex systems. *Proc. Natl. Acad. Sci.* (*USA*) **110**, 2460–2465 (2013).
10. Yuan, Z.-Z., Zhao, C., Di, Z.-R., Wang, W.-X. & Lai, Y.-C. Exact controllability of complex networks. *Nature Commun.* **4**, 2447 (2013).
11. Menichetti, G., Dall'Asta, L. & Bianconi, G. Network controllability is determined by the density of low in-degree and out-degree nodes. *Phys. Rev. Lett.* **113**, 078701 (2014).
12. Ruths, J. & Ruths, D. Control profiles of complex networks. *Science* **343**, 1373–1376 (2014).
13. Wuchty, S. Controllability in protein interaction networks. *Proc. Natl. Acad. Sci.* (*USA*) **111**, 7156–7160 (2014).
14. Yuan, Z.-Z., Zhao, C., Wang, W.-X., Di, Z.-R. & Lai, Y.-C. Exact controllability of multiplex networks. *New J. Phys.* **16**, 103036 (2014).
15. Whalen, A. J., Brennan, S. N., Sauer, T. D. & Schiff, S. J. Observability and controllability of nonlinear networks: The role of symmetry. *Phys. Rev. X* **5**, 011005 (2015).
16. Yan, G. *et al.* Spectrum of controlling and observing complex networks. *Nature Phys.* **11**, 779–786 (2015).
17. Chen, Y.-Z., Wang, L.-Z., Wang, W.-X. & Lai, Y.-C. The paradox of controlling complex networks: control inputs versus energy requirement. *arXiv:1509.03196v1* (2015).
18. Strogatz, S. H. Exploring complex networks. *Nature* (*London*) **410**, 268–276 (2001).
19. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys* **74**, 47–97 (2002).
20. Mendes, J. F. F., Dorogovtsev, S. N. & Ioffe, A. F. *Evolution of Networks: From Biological Nets to the Internet and the WWW* (Oxford University Press, Oxford UK, 2003) first edn.
21. Pastor-Satorras, R. & Vespignani, A. *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge UK, 2004) first edn.
22. Barrat, A., Barthélemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge UK, 2008) first edn.
23. Newman, M. E. J. *Networks: An Introduction* (Oxford University Press, Oxford, 2010) first edn.
24. Watts, D. J. & Strogatz, S. H. Collective dynamics of small-worldnetworks. *Nature* **393**, 440–442 (1998).
25. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
26. Lin, C.-T. Structural controllability. *IEEE Trans. Automat. Contr.* **19**, 201–208 (1974).
27. Hopcroft, J. E. & Karp, R. M. An $n^5/2$ algorithm for maximum matchings in bipartite graphs. *SIAM J. Comput.* **2**, 225–231 (1973).
28. Zhou, H. & Ou-Yang, Z.-C. Maximum matching on random graphs. *arXiv preprint cond-mat/0309348* (2003).
29. Zdeborová, L. & Mézard, M. The number of matchings in random graphs. *J. Stat. Mech.* **2006**, P05003 (2006).
30. Horn, R. A. & Johnson, C. R. *Matrix analysis* (Cambridge university press, 1985).
31. Jia, T. *et al.* Emergence of bimodality in controlling complex networks. *Nature Commun.* **4** (2013).
32. Jia, T. & Barabási, A.-L. Control capacity and a random sampling method in exploring controllability of complex networks. *Sci. Rep.* **3** (2013).
33. Gao, J., Liu, Y.-Y., D'Souza, R. M. & Barabási, A.-L. Target control of complex networks. *Nature Commun.* **5** (2014).
34. Zhao, C., Wang, W.-X., Liu, Y.-Y. & Slotine, J.-J. Intrinsic dynamics induce global symmetry in network controllability. *Sci. Rep.* **5** (2015).
35. Cowan, N. J., Chastain, E. J., Vilhena, D. A., Freudenberg, J. S. & Bergstrom, C. T. Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PLoS One* **7**, e38398 (2012).
36. Liu, Y.-Y., Slotine, J.-J. & Barabási, A. L. Control centrality and hierarchical structure in complex networks. *PLoS One* **7**, e44459 (2012).
37. Slotine, J.-J. E. & Li, W. *Applied Nonlinear Control*, vol. 199 (Prentice-Hall Englewood Cliffs, NJ, 1991).
38. Luenberger, D. G. *Introduction to Dynamical Systems: Theory, Models, and Applications* (John Wiley & Sons, Inc, New Jersey, 1999) first edn.
39. Kailath, T. *Linear systems*, vol. 1 (Prentice-Hall Englewood Cliffs, NJ, 1980).
40. Kalman, R. E. Mathematical description of linear dynamical systems. *J. Soc. Indus. Appl. Math. Ser. A* **1**, 152–192 (1963).
41. Strang, G. *Linear Algebra and Its Applications* (Academic Press, 1976).
42. Sabidussi, G. The centrality index of a graph. *Psychometrika* **31**, 581–603 (1966).
43. Freeman, L. C. A set of measures of centrality based on betweenness. *Sociometry* **40**, 35–41 (1977).
44. Bonacich, P. Power and centrality: A family of measures. *Ame. J. Socio.* **92**, 1170–1182 (1987).
45. Bonacich, P. & Lloyd, P. Eigenvector-like measures of centrality for asymmetric relations. *Soc. Net.* **23**, 191–201 (2001).
46. Brin, S. & Page, L. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Comp. Net.* **56**, 3825–3833 (2012).
47. van Mieghem, P. *Graph Spectra for Complex Networks* (Cambridge University Press, 2010).
48. Parlangeli, G. & Notarstefano, G. On the reachability and observability of path and cycle graphs. *IEEE Trans. Autom. Contr.* **57**, 743–748 (2012).
49. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. *Nature* **473**, 167–173 (2011).
50. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).
51. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
52. Cornelius, S. P., Kath, W. L. & Motter, A. E. Realistic control of network dynamics. *Nature Commun* **4**, 1942 (2013).

53. Well, D. K., Kath, W. L. & Motter, A. E. Control of Stochastic and Induced Switching in Biophysical Networks. *Phys. Rev. X* **5,** 031036 (2015).
54. Wang, L.-Z., Su, R.-Q., Huang, Z.-G., Wang, X., Wang, W.-X., Grebogi, C. & Lai, Y.-C. A geometrical approach to control and controllability of nonlinear dynamical networks. *Nature Commun.* **7,** 11323 (2016).
55. Lai, Y.-C. Controlling complex, nonlinear dynamical networks. *Nat. Sci. Rev.* **1,** 339–341 (2014).
56. Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* **74,** 036104 (2006).
57. Lusseau, D. *et al.* The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behav. Ecol. Sociobio.* **54,** 396–405 (2003).
58. Girvan, M. & Newman, M. E. Community structure in social and biological networks. *Proc. Nat. Acad. Sci. (USA)* **99,** 7821–7826 (2002).
59. Zachary, W. W. An information flow model for conflict and fission in small groups. *J. Anthrop. Res.* 452–473 (1977).
60. Knuth, D. E., Knuth, D. E. & Knuth, D. E. *The Stanford GraphBase: A Platform for Combinatorial Computing*, vol. 37 (Addison-Wesley Reading, 1993).
61. Ripeanu, M., Foster, I. & Iamnitchi, A. Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system design. *arXiv preprint cs/0209028* (2002).
62. Newman, M. E. The structure of scientific collaboration networks. *Proc. Nat. Acad. Sci. (USA)* **98,** 404–409 (2001).
63. Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F. & Arenas, A. Self-similar community structure in a network of human interactions. *Phys. Rev. E* **68,** 065103 (2003).
64. Gleiser, P. M. & Danon, L. Community structure in jazz. *Adv. Complex Sys.* **6,** 565–573 (2003).
65. http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm.

Acknowledgements

Author Contributions

W.-X.W. designed research. X.-D.G. and W.-X.W. performed research; all analyzed data; W.-X.W. and Y.-C.L. wrote the paper.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Gao, X.-D. *et al.* Control efficacy of complex networks. *Sci. Rep.* **6**, 28037; doi: 10.1038/srep28037 (2016).