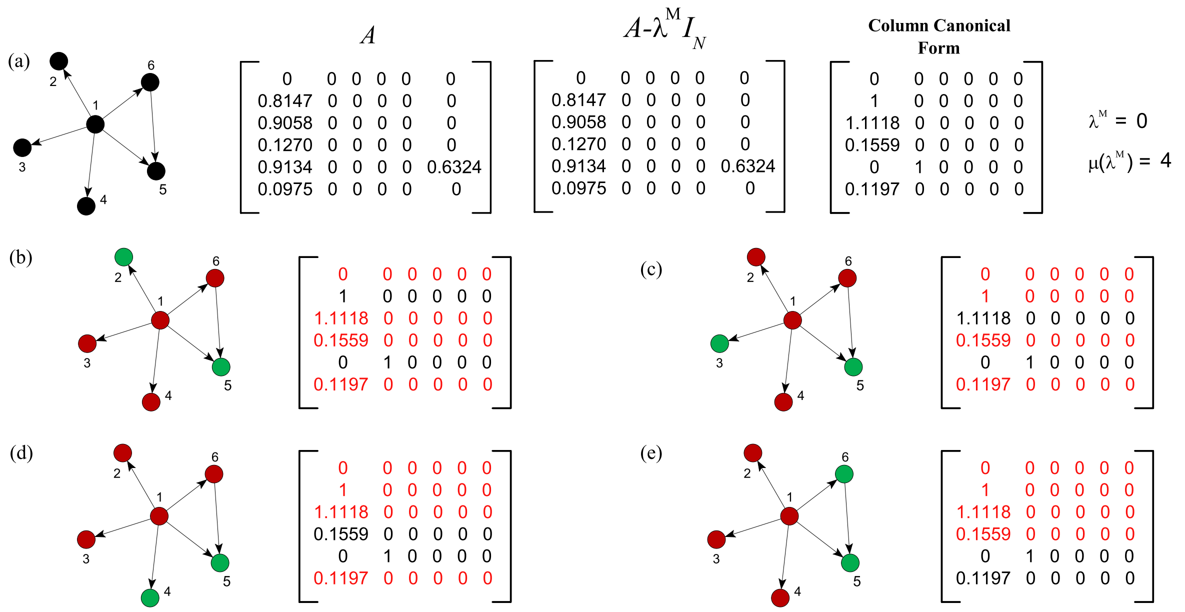
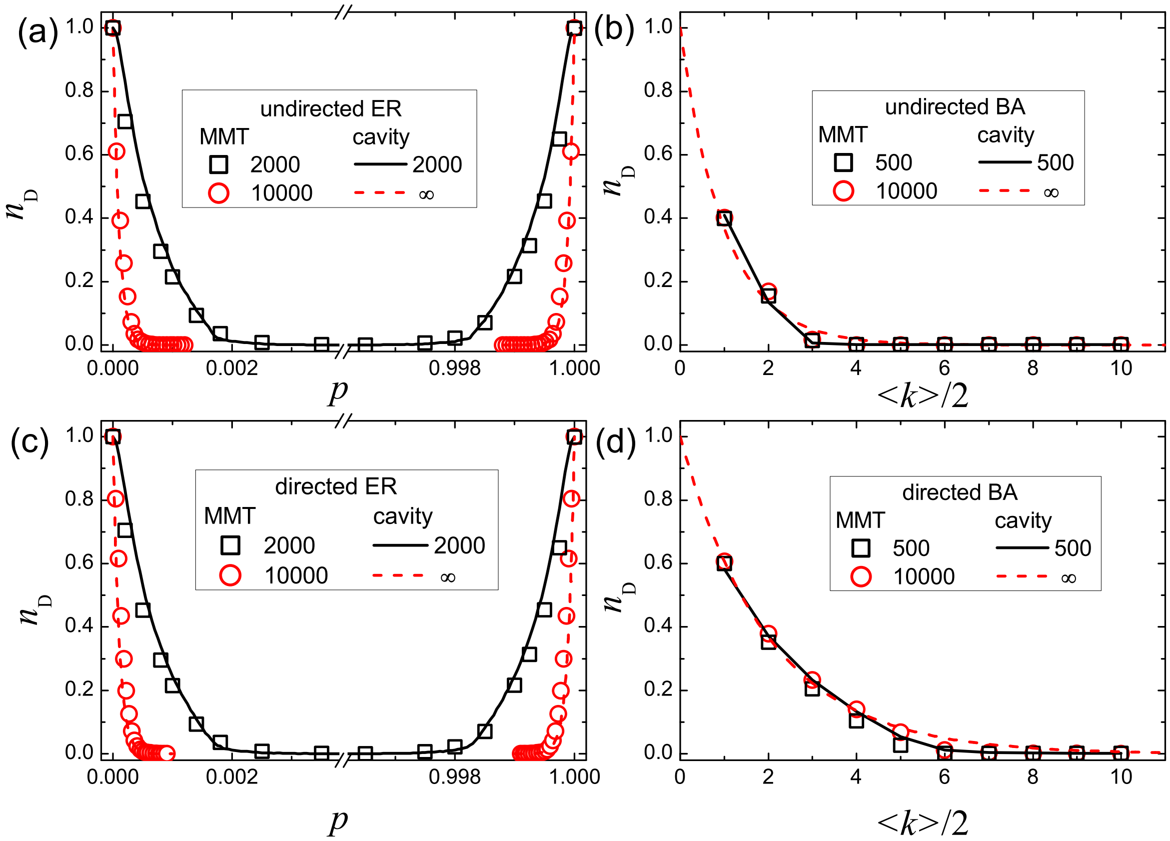
# Supplementary Figures



**Supplementary Figure S1: Identifying all minimum sets of driver nodes in a directed network associated with structural matrix**. (a) A directed network with structural matrix $A$, matrix $A - \lambda^{\mathrm{M}} I_N$ and the column canonical form of matrix $A - \lambda^{\mathrm{M}} I_N$, where $\lambda^{\mathrm{M}}$ is eigenvalue of $A$ resulting in the maximum geometric multiplicity $\mu(\lambda^{\mathrm{M}})$. (b)-(d), four configurations of drivers as marked by red and relevant column canonical form of matrix $A - \lambda^{\mathrm{M}} I_N$ in which the rows that are linearly dependent on others are marked by red. There is a one-to-one correspondence between the drivers and the rows that are linearly dependent on others in the column canonical form. The number of drivers are fixed, as determined by $\mu(\lambda^{\mathrm{M}}) = 4$. We have checked in the four configurations, all the drivers identified by our method are unmatched nodes according to the maximum matching algorithm. There are totally four maximum matchings for this network, which are exactly the same as our findings.




**Supplementary Figure S2: Controllability measure $n_{\mathrm{D}}$ of (a) undirected ER random network, (b) undirected BA network, (c) directed ER network and (d) directed BA network of different network sizes**. $p$ is the probability of having an undirected or a directed link between any pair of nodes in the ER networks, and $\langle k \rangle$ is the average degree of the BA networks. For the directed BA network, we randomly assign each link a direction and the average degree is the sum of average in- and out- degrees. The data points are obtained from the maximum multiplicity theory and the curves are the results from the cavity method. For ER networks with small and large values of $p$ and BA networks, the results from the cavity method are based on Eqs. (S26) and (S31), respectively. In the thermodynamic limit, the results of cavity method are from Eqs. (S37) and (S38). All numerical results are obtained by averaging over 10 independent network realizations.

# Supplementary Table

**Supplementary Table S1: Summary of the real unweighted and weighted networks analyzed in the paper**.

| | Index | Name | $N$ | $L$ | Class | Desciption |
|---|---|---|---|---|---|---|
| Trust | 1 | Prison inmate [60, 61] | 67 | 182 | Unweighted | Social networks of positive sentiment(prisoninmates). |
| | 2 | WikiVote [62] | 7115 | 103689 | Unweighted | Who-vote-whom network of Wikipedia users. |
| Food web | 3 | St.martin [63] | 45 | 224 | Unweighted | Food Web in YthanEstuary. |
| | 4 | Seagrass [64] | 49 | 226 | Unweighted | Food Web in Seagrass. |
| | 5 | Grassland [65] | 88 | 137 | Unweighted | Food Web in Grassland. |
| | 9 | Ythan [65] | 135 | 601 | Unweighted | Food Web in Ythan. |
| | 6 | Mangrove [66] | 97 | 1492 | Weighted | Food Web in Mangrove. |
| | 7 | Florida Baydry [66] | 128 | 2137 | Weighted | Food Web in Baydry. |
| | 8 | Florida Baywet [66] | 128 | 2106 | Weighted | Food Web in Florida. |
| | 10 | Silwood [67] | 154 | 370 | Unweighted | Food Web in Silwood. |
| | 11 | Littlerock [68] | 183 | 2494 | Unweighted | Food Web in Littlerock. |
| Electronic circuits | 12 | s208a [69] | 122 | 189 | Unweighted | Electronic sequential logic circuit. |
| | 13 | s420a [69] | 252 | 399 | Unweighted | Same as above. |
| | 14 | s838a [69] | 515 | 819 | Unweighted | Same as above. |
| Neuronal | 15 | C.elegans [70] | 297 | 2359 | Unweighted | Neural network of C.elegans. |
| Citation | 16 | Small World [71] | 233 | 1988 | Unweighted | Citation network in S.Milgram's Small World(1967). |
| | 17 | SciMet [71] | 2729 | 10416 | Unweighted | Citation network in Scientometrics(1978-2000). |
| | 18 | Kohonen [72] | 3772 | 12731 | Unweighted | Citation network in T.Kohonen's Small World. |
| World Wide Web | 19 | Polblogs [73] | 1224 | 19090 | Unweighted | Hyper links between web logs on US politics. |
| Internet | 20 | P2P-1 [74] | 10876 | 39994 | Unweighted | Gnutella peer-to-peer file sharing network. |
| | 21 | P2P-2 [74] | 8846 | 31839 | Unweighted | Same as above. |
| | 22 | P2P-3 [74] | 8717 | 31525 | Unweighted | Same as above. |
| Organizational | 23 | Freeman-1 [75] | 34 | 695 | Unweighted | Social network of network researchers. |
| | 24 | Consulting [76] | 46 | 879 | Unweighted | Social network from a consulting company. |
| Language | 25 | Word-English [77] | 7381 | 46281 | Unweighted | The words network in English. |
| | 26 | Word-French [77] | 8325 | 24295 | Unweighted | The words network in French. |
| Transportation | 27 | USA top 500 [78] | 500 | 5960 | Weighted | Flight network in USA. |
| Co-authorships | 28 | Coauthorships [79] | 1461 | 2742 | Weighted | The Co-authorships between the scientists. |
| Social communication | 29 | Facebook-like [80] | 899 | 142760 | Weighted | The online social network as similar as Face-book. |
| | 30 | UCIonline [81] | 1899 | 20296 | Weighted | Online message network of students at UC,Irvine. |
| Metabolic | 31 | C.elegans [82] | 453 | 2040 | Weighted | Metabolic network of C.elegans. |

## Supplementary Note 1: Rank of $\lambda_i I_N - J$

We detail the calculation of the rank of matrix $\lambda_i I_N - J$, where $\lambda_i$ denotes distinct eigenvalues of the Jordan block matrix $J$. Recall the linear dynamical system [1] as described by

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}, \tag{S1}$$

where $A \in R^{N\times N}$ denotes the system's coupling matrix, in which $a_{ij}$ denotes the weight of a directed link from node $j$ to $i$ (for undirected networks, $a_{ij} = a_{ji}$), $\mathbf{u}$ is the controller vector with $\mathbf{u} = (u_1, u_2, \cdots, u_m)^{\mathrm{T}}$, and $B$ is the $N \times m$ control matrix. Using the nonsingular transformation $\mathbf{y} = P^{-1}\mathbf{x}$ and $Q = P^{-1}B$, system (S1) can be rewritten in the following Jordan form:

$$\dot{\mathbf{y}} = J\mathbf{y} + Q\mathbf{u}, \tag{S2}$$

where $J$ is the Jordan matrix. Systems (S1) and (S2) possess the same controllability in the sense that $\mathrm{rank}(\lambda I_N - A, B) = \mathrm{rank}(\lambda I_N - J, Q), \forall\lambda \in \sigma(A)$ with $\mathrm{rank}(B) = \mathrm{rank}(Q)$.

For an arbitrary Jordan matrix [45,57]

$$J = \mathrm{diag}(J(\lambda_1), J(\lambda_2), \cdots, J(\lambda_l)), \tag{S3}$$

we have

$$\lambda_i I_N - J = \mathrm{diag}(\lambda_i I_1 - J(\lambda_1), \lambda_i I_2 - J(\lambda_2), \cdots, \lambda_i I_l - J(\lambda_l)), \tag{S4}$$

where the unit matrix $I_i$ $(i = 1, 2, \cdots, l)$ is of the same order as $J(\lambda_i)$. If $\lambda_i \neq \lambda_j$, $\lambda_i I_j - J(\lambda_j)$ is a nonsingular matrix. In this case, rank deficiency can only appear in $\lambda_i I_i - J(\lambda_i)$. Note that each basic Jordan block can be described by

$$j = \lambda I_v + (0, e_1, e_2, \cdots, e_{v-1}), \tag{S5}$$

where $e_i$ $(i = 1, 2, \cdots, v)$ is the $i$th column of $I_v$. We thus have

$$\lambda I_v - j = -(0, e_1, e_2, \cdots, e_{v-1}) \tag{S6}$$

and

$$\mathrm{rank}(\lambda I_v - j) = v - 1. \tag{S7}$$

Using the fact

$$J(\lambda_i) = \mathrm{diag}(j_1, j_2, \cdots, j_{\mu(\lambda_i)}), \tag{S8}$$

where $\mu(\lambda_i)$ is the geometric multiplicity of $\lambda_i$ and is equal to the number of basic Jordan block in $J(\lambda_i)$, we can conclude that $\lambda_i I_N - J$ has $\mu(\lambda_i)$ zero columns and $N - \mu(\lambda_i)$ independent columns. Thus,

$$\mathrm{rank}(\lambda_i I_N - J) = N - \mu(\lambda_i). \tag{S9}$$

For example, assuming a Jordan block of the following form

$$J = \begin{bmatrix} 1 & 1 & & & & & & & \\ & 1 & & & & & & & \\ & & 1 & & & & & & \\ & & & 2 & 1 & & & & \\ & & & & 2 & 1 & & & \\ & & & & & 2 & & & \\ & & & & & & 2 & & \\ & & & & & & & 2 & \\ & & & & & & & & 0 \end{bmatrix},$$

where $\mu(1) = 2, \mu(2) = 3, \mu(0) = 1$, and all missing elements are zero, we can check that for $\lambda_1 = 1$, the matrix $\lambda_1 I_N - J = 1 \times I_9 - J$ has 2 zero columns which are column 1 and 3; for $\lambda_2 = 2$, $2 \times I_9 - J$ has 3 zero columns, i.e., column 4, 7 and 8; for $\lambda_3 = 0$, $0 \times I_9 - J$ has only 1 zero column 9. These results verify that the number of zero columns for an eigenvalue $\lambda$ is nothing but the geometric multiplicity $\mu(\lambda)$. Subsequently, we have rank$(1 \times I_9 - J) = 9 - 2$, rank$(2 \times I_9 - J) = 9 - 3$, and rank$(0 \times I_9 - J) = 9 - 1$.

## Supplementary Note 2: Determination of control matrix $B$ and transformed matrix $Q$

### A special case: undirected chain graph

In order to achieve actual control of a complex network system, we need to identify the key nodes to apply external control to. Mathematically, this entails finding $B$ and the transformed matrix $Q$. Take an undirected chain graph, $P_N$, as an example. The coupling matrix of $P_N$ is

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{S10}$$

where the $N$ distinct eigenvalues are displayed in Table I in the main text. According to the PBH rank condition, $P_N$ is controllable when all elements of $Q = P^{\mathrm{T}} B$ are nonzero, where $P$ is the matrix whose columns consist of an orthonormal basis of eigenvectors of $A$. The eigenvector $\alpha$ associated with the eigenvalue $\lambda$ satisfies the relation $(\lambda I_N - A)\alpha = 0$ which, when written per component, is [58, 59]

$$\begin{aligned} \lambda\alpha_1 - \alpha_2 &= 0, \\ -\alpha_{k-1} + \lambda\alpha_k - \alpha_{k+1} &= 0, \qquad 2 \le k \le N-1, \\ -\alpha_{N-1} + \lambda\alpha_N &= 0. \end{aligned}$$

We have $\alpha_1 \neq 0$ and $\alpha_N \neq 0$. (Otherwise, we can infer that $\alpha_1 = \alpha_2 = \cdots = \alpha_N = 0$, which contradicts to the definition of eigenvector that must be nonzero.) We can set $B = (1, 0, \cdots, 0)^{\mathrm{T}}$ or

$B = (0, \cdots, 0, 1)^{\mathrm{T}}$ to guarantee that all elements of $Q = P^{\mathrm{T}}B$ are non-zero. This indicates that we can control a chain simply by driving a single node at either end of the chain. This is valid for weighted undirected chain as well.

### Identifying drivers in directed network with structural matrix

We have offered a general method to identifying a minimum set of driver nodes in arbitrary networks by relying on the PBH rank condition and the elementary column transformation. To be concrete, according to the exact-controllability theory, $N_{\mathrm{D}}$ is determined by the maximum geometric multiplicity $\mu(\lambda^{\mathrm{M}})$ associated with the eigenvalue $\lambda^{\mathrm{M}}$. Thus, the control matrix $B$ to ensure full control should satisfy the PBH rank condition by substituting $\lambda^{\mathrm{M}}$ for the complex number $c$, as follows:

$$\mathrm{rank}[\lambda^{\mathrm{M}} I_N - A, B] = N. \tag{S11}$$

The question becomes how to find the minimum number of drivers identified in $B$ to satisfy Eq. (S11). Note that the rank of the matrix $[\lambda^{\mathrm{M}} I_N - A, B]$ is contributed by the number of linearly-independent rows. In this regard, we implement elementary column transformation on the matrix $\lambda^{\mathrm{M}} I_N - A$, which reveals a set of linearly-dependent rows that violates the full rank condition (S11). The controllers should be imposed on the identified rows to eliminate all linear correlations to ensure condition (S11). The nodes corresponding to the linearly-dependent rows are the drivers with number $N - \mathrm{rank}(\lambda^{\mathrm{M}} I_N - A)$, which is nothing but the maximum geometric multiplicity $\mu(\lambda^{\mathrm{M}})$.

Here, we test the validity of this approach in comparison with the structural-controllability framework for a directed network with structural matrix. The assumption of structural matrix can be ensured by assigning each link in a directed network a random parameter, as shown in Supplementary Fig. S1(a). In the network matrix $A$, all the values associated with links are completely independent of others. According to the exact-controllability theory, we calculate the eigenvalues $\lambda$ of $A$ and their geometric multiplicity $\mu(\lambda)$. In this specific case, all the eigenvalues are zero and the maximum geometric multiplicity $\mu(\lambda^{\mathrm{M}}) = 4$. The eigenvalue $\lambda^{\mathrm{M}}$ related with $\mu(\lambda^{\mathrm{M}})$ is zero as well. Subsequently, according to the PBH rank condition, we construct the matrix $A - \lambda^{\mathrm{M}} I_N$, as shown in Supplementary Fig. S1(a), the column canonical form of which resulting from the elementary column transformation reveals the linear dependence among rows. We find that the column canonical form in this case is unique, but there are more than one possible choice of rows that are linearly dependent on others. Each set of options results in a distinct configuration of the set of drivers, but the number of drivers are fixed and determined by $\mu(\lambda^{\mathrm{M}})$. Supplementary Fig. S1(b) to (d) show four combinations of linearly-dependent rows chosen from the column canonical form. In the column canonical form, the 1st row must be chosen and three rows from the 2nd, 3rd, 4th and 6th rows should be chosen to be controlled to eliminate all linear correlations. We thus totally have four different combinations of drivers as shown in Supplementary Fig. S1(b) to (d).

To validate the four identified configurations of drivers, we use the framework of structural controllability based on the maximum matching algorithm to find all possible configurations of drivers for comparison. For the directed network with structural matrix as guaranteed by assigning random parameters to directed links, our method will offer the same result as that from the structural-controllability framework, if our method is correct. We have enumerated all possible combinations of unmatched nodes that according to the structural-controllability theory are the drivers. We acquire four configurations of drivers that are indeed exactly the same as the results shown in Supplementary Fig. S1(b) to (d). This test provides strong evidence to the validity of our PBH-based method for identifying driver nodes.

### How to determine the simplest $Q$

The transformed matrix $Q$ can be used to study control issues related to energy consumption and prove the reachability of the minimum number of controllers and independent drivers as guaranteed by the maximum multiplicity theory. However, $Q$ cannot be immediately derived from the relation of $B = PQ$, since we can only identify independent driver nodes with minimum number in $B$, rather than the complete matrix $B$. Here we present a method [45] to construct $Q$ without requiring full information about $B$. In particular, for the transformed system (S2) with the simplest form of $J$, we can construct $Q$ by assigning value one to the row of $Q$ corresponding to the last row of every Jordan sub-block of $J(\lambda_i)$, i.e., $I_{\mu(\lambda_i)}$, and leaving others rows empty. This is the simplest form of $Q$ with the smallest number of nonzero elements. For diagnolizable matrix $A$ or undirected network, $Q$ can be set according to the multiplicity of eigenvalue $\lambda_i$, i.e., $I_{\delta(\lambda_i)}$. For example, we can set $Q_1$ of the following $J_1$ with $N_{\rm D} = 3$ which we highlight the last row's element of each Jordan sub-block matrix by a box:

$$J_1 = \begin{bmatrix} 1 & 1 & & & & & & & \\ & \boxed{1} & & & & & & & \\ & & \boxed{1} & & & & & & \\ & & & 2 & 1 & & & & \\ & & & & 2 & 1 & & & \\ & & & & & \boxed{2} & & & \\ & & & & & & \boxed{2} & & \\ & & & & & & & \boxed{2} & \\ & & & & & & & & \boxed{0} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & & \\ & 1 & \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & & \\ & 1 & \\ & & 1 \\ 1 & & \end{bmatrix} = Q_1$$

Note that the unwritten elements are zero, the same for $J_2$ and $Q_2$ in the following examples. For undirected networks with the same eigenvalue and algebraic multiplicity of $J_1$, we also can set $Q_2$ with $N_{\rm D} = 5$, as follows:

$$J_2 = \begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & 1 & & & & & & \\ & & & 2 & & & & & \\ & & & & 2 & & & & \\ & & & & & 2 & & & \\ & & & & & & 2 & & \\ & & & & & & & 2 & \\ & & & & & & & & 0 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \\ 1 & & & & \end{bmatrix} = Q_2$$

According to this construction of matrix $Q$, we can ensure that the minimum number $\max_i\{\mu(\lambda_i)\}$ of controllers and independent driver nodes can be achieved as constrained by

$$\text{rank}(B) \geq N - \text{rank}(\lambda_i I_N - J) = \mu(\lambda_i). \tag{S12}$$

Our maximum multiplicity theory thus offers both sufficient and necessary conditions for achieving full control of any complex network, including directed, undirected, weighted, unweighted, connected and disconnected networks in the presence or absence of self-loops or loops.

## Supplementary Note 3: Exact controllability of simple regular graphs

Using our maximum-multiplicity theory, we analytically evaluate the exact controllability of four regular, unweighted and undirected graphs: a chain, a ring graph, a star graph, and a fully connected graph. Especially, for a given graph, we calculate all the eigenvalues and the maximum algebraic multiplicity.

### Undirected chain graph

The coupling matrix of an undirected path of $N$ nodes [58, 59] is

$$A = \begin{bmatrix} 0 & 1 & & & & \\ 1 & 0 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 1 & 0 & 1 \\ & & & 1 & 0 \end{bmatrix}, \tag{S13}$$

where all unwritten elements are zeros. The eigenvector $\alpha$ associated with eigenvalue $\lambda$ of the chain satisfies $(\lambda I_N - A)\alpha = 0$ or $(A - \lambda I_N)\alpha = 0$, which can be written for each component as [58, 59]

$$\begin{aligned} +\lambda\alpha_1 - \alpha_2 &= 0, \\ -\alpha_{k-1} + \lambda\alpha_k - \alpha_{k+1} &= 0, \qquad (2 \le k \le N-1) \\ -\alpha_{N-1} + \lambda\alpha_N &= 0. \end{aligned}$$

The set with $\alpha_0 = \alpha_{N+1} = 0$ can be rewritten as

$$-\alpha_{k+2} + \lambda\alpha_{k+1} - \alpha_k = 0. \qquad (0 \le k \le N-1) \tag{S14}$$

The general solution is $\alpha_k = ar_1^k + br_2^k$, where $r_1$ and $r_2$ are the roots of the corresponding polynomial $x^2 - \lambda x + 1 = 0$, which satisfy

$$\begin{aligned} r_1 + r_2 &= \lambda, \\ r_1 r_2 &= 1. \end{aligned} \tag{S15}$$

The constants $a$ and $b$ in Eq. (S15) are constrained by the boundary requirement $\alpha_0 = \alpha_{N+1} = 0$, yielding

$$\begin{aligned} a + b &= 0, \\ ar_1^{N+1} + br_2^{N+1} &= 0. \end{aligned}$$

We have $a = -b$. The last equation can be modified to $(\frac{r_1}{r_2})^{N+1} = 1$ or $\frac{r_1}{r_2} = e^{\frac{2\pi q\sqrt{-1}}{N+1}} (q = 1, 2, \cdots, N)$. Substituting $r_1 = r_2 e^{\frac{2\pi q\sqrt{-1}}{N+1}}$ into the last equation of (S15) yields

$$r_1 = e^{\frac{\pi q\sqrt{-1}}{N+1}} \qquad \text{and} \qquad r_2 = e^{-\frac{\pi q\sqrt{-1}}{N+1}}. \tag{S16}$$

Since $\lambda = r_1 + r_2$, we have

$$\lambda = e^{\frac{\pi q \sqrt{-1}}{N+1}} + e^{-\frac{\pi q \sqrt{-1}}{N+1}} = 2\cos\frac{q\pi}{N+1}, \tag{S17}$$

so the maximum algebra multiplicity is $\delta(\lambda_q) = 1$ $(q = 1, 2, \cdots, N)$.

### Ring network

The adjacency matrix of a directed ring of $N$ nodes [58, 59] is

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{S18}$$

which has the property $C^{\mathrm{T}} = C^{-1}$. The characteristic polynomial of $C$ is $p_C(\zeta) = \zeta^N - 1$ [59], so the eigenvalues of $C$ are $\zeta_q = e^{\frac{2(q-1)\pi i}{N}}$ $(q = 1, 2, \cdots, N)$, where $i$ is the imaginary unit.

The coupling matrix of an undirected ring network of $N$ nodes is $A = C + C^{\mathrm{T}} = C + C^{-1}$. Hence, the eigenvalues of $A$ are $\lambda_q = \zeta_q + \zeta_q^{-1} = 2\cos\frac{2(q-1)\pi}{N}$ with $\lambda_q = \lambda_{N-q+2}$, which gives $\delta(\lambda_q) = 2$ $(q = 1, 2, \cdots, N)$ for $N > 4$.

### Star graph

The coupling matrix of a star graph of $N$ nodes is [58, 59]:

$$A = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{S19}$$

The characteristic polynomial of $A$ is

$$p_A(\lambda) = \mid \lambda I_N - A \mid = [\lambda^2 - (N-1)](\lambda + 1)^{N-2}. \tag{S20}$$

The eigenvalues, which are roots of $p_A(\lambda)$, and the corresponding algebraic multiplicities are: $\lambda_1 = \sqrt{N-1}$, $\delta(\lambda_1) = 1$; $\lambda_2 = -\sqrt{N-1}$, $\delta(\lambda_2) = 1$; and $\lambda_3 = -1$, $\delta(\lambda_3) = N - 2$.

### Fully connected network

The coupling matrix is [58, 59]

$$A = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix}. \tag{S21}$$

The characteristic polynomial of $A$ is

$$p_A(\lambda) = \mid \lambda I_N - A \mid = [\lambda - (N-1)](\lambda + 1)^{N-1}. \tag{S22}$$

The eigenvalues and the respective algebraic multiplicities are: $\lambda_1 = N - 1$, $\delta(\lambda_1) = 1$ and $\lambda_2 = -1$, $\delta(\lambda_2) = N - 1$.

## Supplementary Note 4: Interplay between maximum multiplicity and network structure

We aim to reveal the interplay between our theory and network structure by the aid of the maximum matching algorithm. Let's recall the following general formula:

$$N_\mathrm{D} = \max_i\{N - \mathrm{rank}(\lambda_i I_N - A)\}, \tag{S23}$$

where $\lambda_i$ is the eigenvalue of the network coupling matrix $A$. If $A$ is diagnolizable, e.g., as for an undirected network, we have

$$N_\mathrm{D} = \max_i\{\delta(\lambda_i)\}, \tag{S24}$$

which is the maximum number of identical eigenvalues. For an arbitrary network, it is difficult to make statement about the maximum multiplicity and network structure. However, for a sparse network (weighted or unweighted), the zero eigenvalue dominates the eigenvalue spectrum. In this case, we have

$$N_\mathrm{D} = \max\{1, N - \mathrm{rank}(A)\}, \tag{S25}$$

revealing the underlying relationship between structure and exact controllability in that $N_\mathrm{D}$ is exclusively determined by the rank of $A$. The task then becomes that to understand the interdependence between the matrix rank and its structure. We can demonstrate that for sparse networks, the rank of the coupling matrix is approximately equal to the maximum matching [34], based on which we can uncover the relationship between the exact controllability and the network structure by means of the maximum matching and estimate the exact controllability through the cavity method [36,37]. The key to explaining the approximate equality of rank and maximum matching lies in understanding the connection between our maximum multiplicity theory and the structural controllability.

First, we have proved that under the condition of zero value domination, e.g., for sparse networks, the exact controllability is rigorously determined by $N - \mathrm{rank}(A)$.

Second, as we have demonstrated in the main text, for sparse networks with identical weights, the structural controllability is quite close to the exact controllability with negligible small difference. This can be explained in terms of existence of linear constrains resulting from the linear correlations among nodes. In the state space of a linear network system, there may exist some linear constrains that prevent the trajectory of the system state from reaching an arbitrary state in the phase space. Implementing controllers can eliminate all constrains so that control toward any point in the state space is possible. All the constraints can be classified to two categories: in terms of topology and link weights. Structural controllability only considers the topological constraints by assuming that constraints resulting from link weight are of zero measure. For exact controllability, both constraints are important, but sparse networks represent an exceptional case, because of the low probability of generating weight constraints. Consequently, as illustrated in Figs. 2 and 3 in the main text, in this case structural controllability is approximately equivalent to the exact controllability.

Based on the above argument, we have, for sparse networks, $N_{\mathrm{D}} = N - \mathrm{rank}(A) \approx N - N_{\mathrm{m}}(A)$, where $N_{\mathrm{m}}(A)$ is the number of nodes in the maximum matching set of $A$. This thus yields

$$\mathrm{rank}(A) \approx N_{\mathrm{m}}(A). \tag{S26}$$

As presented in Ref. [34], the cavity method can be used to estimate $N_{\mathrm{m}}$ [36,37].

For dense networks with unit weights, we have

$$N_{\mathrm{D}} = \max\{1, N - \mathrm{rank}(I_N + A)\}. \tag{S27}$$

The main feature is then how to reveal the relationship between the rank of matrix $I_N + A$ and the network topology. However, in this situation, we have observed that the exact controllability is quite different from the structural controllability due to the very high likelihood of the occurrence of linear constraints resulting from link weights. Hence, the maximum matching algorithm and the cavity method cannot be directly used to uncover the relationship. In order to overcome this difficulty, we consider the complement graph of matrix $I_N + A$, which is given by $\mathcal{J}_N - I_N - A$, where $\mathcal{J}_N$ is the matrix whose elements are all one. We then have

$$\mathrm{rank}(\mathcal{J}_N - I_N - A) \leq \mathrm{rank}(\mathcal{J}_N) + \mathrm{rank}(I_N + A), \tag{S28}$$

where the equality holds if one of the ranks in the right hand side is zero. Note that $\mathrm{rank}(\mathcal{J}_N) = 1$, which is quite close to zero and far from $N$ if the network size is large enough. We thus have, approximately,

$$\mathrm{rank}(I_N + A) \approx \mathrm{rank}(\mathcal{J}_N - I_N - A). \tag{S29}$$

The complement graph of the original network is sparse and can be related to the maximum matching as

$$\mathrm{rank}(I_N + A) \approx N_{\mathrm{m}}(\mathcal{J}_N - I_N - A). \tag{S30}$$

The maximum matching of the complement graph on the right-hand side can be found by the cavity method due to the sparsity of the complement graph. We thus have the following relationship between maximum multiplicity and the maximum matching:

$$N_{\mathrm{D}} = \max_i\{\mu(\lambda_i)\} \approx N - \mathrm{rank}(I_N + A) \approx N - N_{\mathrm{m}}(\mathcal{J}_N - I_N - A). \tag{S31}$$

In Supplemental Materials of Ref. [34], the cavity method for maximum matching of directed networks is detailed. Specifically, for a directed network with similar in- and out-degree distribution $P(k)$, the density of driver nodes is given by

$$n_{\mathrm{D}} = G(w_2) + G(1 - w_1) - 1 + \langle k \rangle w_1(1 - w_2), \tag{S32}$$

where $G(x)$ is the generating function

$$G(x) = \sum_{k=0}^{\infty} P(k) x^k. \tag{S33}$$

The quantities $w_1$ and $w_2$ in Eq. (S32) can be obtained by the following self-consistent equations:

$$\begin{aligned} w_1 &= H[1 - H(1 - w_1)], \\ w_2 &= 1 - H[1 - H(w_2)], \end{aligned} \tag{S34}$$

where

$$H(x) = \sum_0^\infty Q(k+1)x^k \tag{S35}$$

is a generating function and $Q(k) = kP(k)/\langle k\rangle$. Equation (S34) is valid for general networks in the absence of degree-degree correlations. For ER random networks, $P(k)$ follows the Poisson distribution $e^{-\langle k\rangle}\langle k\rangle^k/k!$. We thus have $G(x) = H(x) = \exp[-\langle k\rangle(1-x)]$, and

$$n_{\rm D} = w_1 - w_2 + \langle k\rangle w_1(1 - w_2), \tag{S36}$$

where $\langle k\rangle$ is the average in- or out-degree in the network, $w_1 = H(w_2) = \exp[-\langle k\rangle(1-w_2)]$ and $w_2 = 1 - H(1-w_1) = 1 - \exp[-\langle k\rangle w_1]$. For $k \gg 1$, we have

$$n_{\rm D} \approx \exp[-\langle k\rangle]. \tag{S37}$$

For scale-free networks, we have

$$n_{\rm D} \approx \exp\left[-\langle k\rangle\left(1 - \frac{1}{\gamma - 1}\right)\right]. \tag{S38}$$

Supplementary Fig. S2 shows the measure of the controllability $n_{\rm D}$ calculated from both the maximum geometric multiplicity and the cavity method for undirected and directed Erdös-Rényi (ER) [43] and Barabási-Albert (BA) [12] networks with unit link weights. In all cases, there is good agreement between the results from both methods. In particular, for large network size, e.g., $N = 10000$, the exact controllability obtained by the maximum multiplicity theory matches that predicted by the cavity method in the thermodynamic limit. In this regard, the relationship between the maximum multiplicity and the network structure is bridged by the maximum matching algorithm for both sparse and dense networks.

## Supplementary Note 5: Energy of control

The exact-controllability framework allows us to study the issue of control energy in, for example, unweighted networks given the minimum number of driver nodes. The control-energy problem in unweighted networks has been explored by Yan et al. [39] with respect to relatively large average node degree. In this case, according to the exact-controllability framework, single controller can ensure full control of the network with convergent control energy. However, except the case of requiring a single controller, one has to rely on the framework to identify the minimum number of controllers to address the important issue of control energy.

Similar to the approach in Ref. [39], we begin from the definition of energy in canonical control theory to show how to make use of the transformed matrix $Q$ to simplify the calculation and derive the upper and lower bounds of energy. For system (S1), without loss of generality, we consider driving the network from an initial state $\mathbf{x}_0$ at time 0, to the destination at time $t_1$. According to the linear control theory [45], the controller is $\mathbf{u} = -B^{\mathrm{T}}e^{-A^{\mathrm{T}}t}W^{-1}e^{At}\mathbf{x}_0$, so that the minimum total energy $E = \int_0^{t_1} \parallel \mathbf{u}\|^2 \, dt$ can be evaluated as

$$E = \mathbf{x}_0^{\mathrm{T}} W^{-1} \mathbf{x}_0, \tag{S39}$$

where

$$W = \int_0^{t_1} e^{-At} B B^{\mathrm{T}} e^{-A^{\mathrm{T}}t} dt \tag{S40}$$

is the controllability Gramian matrix. When system (S1) is controllable, $W$ is positive definite (as well as the matrix $S$ defined below) [45]. We take undirected networks with $A^{\mathrm{T}} = A$ as an example. System (S1) can be transformed to

$$\dot{\mathbf{y}} = \Lambda \mathbf{y} + Q\mathbf{u}, \tag{S41}$$

where $\mathbf{y} = P^{\mathrm{T}}\mathbf{x}$, $Q = P^{\mathrm{T}}B$, $P^{\mathrm{T}}P = PP^{\mathrm{T}} = I_N$ and $P$ is the matrix whose columns are an orthonormal basis of eigenvectors of $A$. For an undirected network, $W$ can be calculated as

$$W = P \int_0^{t_1} e^{-\Lambda t} Q Q^{\mathrm{T}} e^{-\Lambda t} dt P^{\mathrm{T}}.$$

Defining

$$S = \int_0^{t_1} e^{-\Lambda t} Q Q^{\mathrm{T}} e^{-\Lambda t} dt, \tag{S42}$$

we have

$$E = \mathbf{x}_0^{\mathrm{T}} P S^{-1} P^{\mathrm{T}} \mathbf{x}_0 = \mathbf{y}_0^{\mathrm{T}} S^{-1} \mathbf{y}_0 \tag{S43}$$

with $\mathbf{y}_0 = P^{\mathrm{T}}\mathbf{x}_0$. For convenience, we consider the normalized energy $\epsilon$ with the inequality [55]

$$\frac{1}{\lambda_{\text{max}}(S)} \leq \epsilon = \frac{\mathbf{x}_0^{\mathrm{T}} W^{-1} \mathbf{x}_0}{\mathbf{x}_0^{\mathrm{T}} \mathbf{x}_0} = \frac{\mathbf{y}_0^{\mathrm{T}} S^{-1} \mathbf{y}_0}{\mathbf{y}_0^{\mathrm{T}} \mathbf{y}_0} \leq \frac{1}{\lambda_{\text{min}}(S)} \tag{S44}$$

where $\lambda_{\text{max}}(S)$ and $\lambda_{\text{min}}(S)$ stand for the maximum and minimum eigenvalue of $S$, respectively. Therefore, for a given transformed control matrix $Q$ (or equivalently $B = PQ$) constructed by our exact-controllability theory, we can compute the total energy as well as the normalized energy's lower bound $1/\lambda_{\text{max}}(S)$ and upper bound $1/\lambda_{\text{min}}(S)$ from Eq. (S44). The two bounds are in agreement with that in Ref. [39].

When the network can be fully controlled by only one controller, $S$ is well defined. We stress that, for single controller, there may be more than one nodes controlled simultaneously to achieve full control. However, for any configuration of driver nodes associated with a single controller, the transformed control matrix $Q$ must have the form $Q = P^{\mathrm{T}}B = (\alpha_1, \alpha_2, ..., \alpha_N)^{\mathrm{T}}, \alpha_i \neq 0$ [45]. Consequently, $S$ can be derived as

$$\begin{aligned} S &= \int_0^{t_1} e^{-\Lambda t} Q Q^{\mathrm{T}} e^{-\Lambda t} dt \\ &= \int_0^{t_1} \begin{bmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \\ \vdots \\ \alpha_N e^{\lambda_N t} \end{bmatrix} \begin{bmatrix} \alpha_1 e^{\lambda_1 t} & \alpha_2 e^{\lambda_2 t} & \cdots & \alpha_N e^{\lambda_N t} \end{bmatrix} dt \end{aligned} \tag{S45}$$

with $i, j = 1, 2, \cdots, N$, so that

$$\begin{aligned} S_{ij} &= \alpha_i \alpha_j \int_0^{t_1} e^{-(\lambda_i+\lambda_j)t} dt \\ &= \begin{cases} \alpha_i \alpha_j t_1 & \lambda_i + \lambda_j = 0, \\ \alpha_i \alpha_j \frac{1-e^{-(\lambda_i+\lambda_j)t_1}}{\lambda_i+\lambda_j} & \lambda_i + \lambda_j \neq 0. \end{cases} \end{aligned} \tag{S46}$$

Equation (S46) gives the elements in the matrix $S$. However, to derive the control energy, we need the inverse of $S$, which at present cannot be resolved analytically. In this sense, the transformed matrix $Q$ can be used to simplify the calculation of energy rather than offering analytical results. Yet, how to theoretically predict control energy of complex networks is extremely challenging and remains an outstanding problem.

## Supplementary Note 6: Observability of complex networks

The observability of a complex system is defined as the minimum number of observers required to probe the states of nodes so as to reconstruct the dynamics of the whole system. Here we prove that the observability of an arbitrary network is equivalent to its exact controllability in the framework of PBH rank condition.

Consider a complex network system in the presence of observers [17], which is described by

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u}, \\ \mathbf{z} &= C_2 \mathbf{x}, \end{aligned} \tag{S47}$$

where $C_2$ is the $r \times N$ observe matrix. There exists a PBH rank condition akin to exact controllability, which can be described as

$$\text{rank} \begin{pmatrix} \lambda I_N - A \\ C_2 \end{pmatrix} = N \tag{S48}$$

for all $\lambda \in \sigma(A)$. This equation is equivalent to

$$\text{rank}[\lambda I_N - A^{\mathrm{T}}, C_2^{\mathrm{T}}] = N. \tag{S49}$$

In general $A$ and $A^{\mathrm{T}}$ have the same eigenvalues and the same algebraic and geometric multiplicities [55]. Consequently, the observability of a complex network is equivalent to the controllability of the same network, i.e.,

$$N_{\mathrm{O}} = N_{\mathrm{D}} = \max_i \{\mu(\lambda_i)\}, \tag{S50}$$

where the number of observers $N_{\mathrm{O}}$ is equal to the rank of $C_2$. Analogous to controllability, with respect to undirected networks or diagonlizable matrix $A$, we have

$$N_{\mathrm{O}} = N_{\mathrm{D}} = \max_i \{\delta(\lambda_i)\}, \tag{S51}$$

where $\delta(\lambda_i)$ is the algebraic multiplicity of $\lambda_i$.

## Supplementary Note 7: Network data sets

The details of the real-world networks studied in this paper are presented in Table S1, including category of networks, index, name, number $N$ of nodes, number $L$ of directed or undirected links, types of networks (weighted or unweighted), and the description.

## Supplementary References


58. Van Mieghem, P. *Performance Analysis of Communications Systems and Networks.* (Cambridge Univ. Press, Cambridge, UK, 2006).

59. Brouwer, A. E. & Haemers, W.H. *Spectra of graphs.* (Springer, 2012).

60. Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M. & Alon, U. Superfamilies of evoved and designed networks. *Science* **303**, 1538-1542 (2004).

61. Van Duijn, M. A. J., Zeggelink, E. P. H., Huisman, M., Stokman, F. N. & Wasseur, F. W. Evolution of sociology freshmen into a friendship network. *J. Math. Sociol.* **27**, 153-191 (2003).

62. Leskovec, J., Lang, K. J., Dasgupta, A. & Mahoney, M. W. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*. **6**, 29-123 (2009).

63. Baird, D., Luczkovich, J. & Christian, R. R. Assessment of spatial and temporal variability in ecosystem attributes of the St Marks National Wildlife Refuge, Apalachee Bay, Florida. *Estuarine, Coastal, and Shelf Science* **47**, 329-349 (1998).

64. Christian, R. R. & Luczkovich, J. J., Organizing and understanding a winter's seagrass food web network through effective trophic levels. *Ecol. Model.* **117**, 99-124 (1999).

65. Dunne, J. A., Williams, R. J. & Martinez, N. D. Food-web structure and network theory: the role of connectance and size. *Proc. Natl. Acad. Sci. USA* **99**, 12917-12922 (2002).

66. Ulanowicz, R. E. & DeAngelis, D. L. Network Analysis of Trophic Dynamics in South Florida Ecosystems. *US Geological Survey Program on the South Florida Ecosystem*. 114, (2005).

67. http://pil.phys.uniroma1.it/∼gcalda/cosinsite/extra/data/food- webs/WEB.html.

68. Martinez, N. D. Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. *Ecol. Monogr.* **61**, 367-392 (1991).

69. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. Network motifs: simple building blocks of complex networks. *Science Signaling* **298**, 824-827 (2002).

70. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440-442 (1998).

71. http://vlado.fmf.uni-lj.si/pub/networks/data/cite/default.htm.

72. http://www.cise.ufl.edu/research/sparse/matrices/Pajek/Koho- nen.html.

73. Adamic, L. A. & Glance, N. The Political Blogosphere and the 2004 US Election: divided they blog. *Proceeding LinkKDD '05 Proceedings of the 3rd international workshop on Link discovery*, 36-43 (2005).

74. Leskovec, J., Kleinberg, J. & Faloutsos, C. Graph evolution: Densification and shrinking diameters. ACM Transactions on Knowledge Discovery from Data(TKDD), **1**, 2 (2007).

75. Freeman, S. C. & Freeman, L. C. The networkers network: A study of the impact of a new communications medium on sociometric structure. *Social Science Research Reports*, **46**. University of California, Irvine, CA. (1979).

76. Cross, R. L. & Parker, A. *The hidden power of social networks: Understanding how work really gets done in organizations.* (Harvard Business School Press, Boston, MA, 2004).

77. http://www.clairlib.org/index.php/Corpora.

78. Colizza, V., Pastor-Satorras, R. & Vespignani, A. Reaction-Diffusion processes and meta-population models in Heterogeneous Networks. *Nat. Phys.* **3**, 276-282 (2007).

79. Newman, M. E. J. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* **74**, 036104 (2006).

80. Opsahl, T. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Soc. Netw.* **35**, 159-167 (2013)

81. Opsahl, T. & Panzarasa, P. Clustering in weighted networks. *Soc. Netw.* **31**, 155-163 (2009).

82. Duch, J. & Arenas, A. Community detection in complex networks using extremal optimization. *Phys. Rev. E* **72**, 027104 (2005).