



## OPEN

## Connecting Core Percolation and Controllability of Complex Networks

## SUBJECT AREAS:

APPLIED PHYSICS

COMPLEX NETWORKS

STATISTICAL PHYSICS,  
THERMODYNAMICS AND  
NONLINEAR DYNAMICS

PHYSICS

Tao Jia<sup>1,2,3</sup> & Márton Pósfai<sup>3,4,5</sup>

<sup>1</sup>Social Cognitive Networks Academic Research Center, Rensselaer Polytechnic Institute, Troy, NY, 12180 USA, <sup>2</sup>Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, 12180 USA, <sup>3</sup>Center for Complex Network Research and Department of Physics, Northeastern University, Boston, Massachusetts 02115, USA, <sup>4</sup>Department of Physics of Complex Systems, Eötvös University, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary, <sup>5</sup>Department of Theoretical Physics, Budapest University of Technology and Economics, Budafoki ut 8, Budapest H-1111, Hungary.

Received  
12 February 2014Accepted  
2 June 2014Published  
20 June 2014

Core percolation is a fundamental structural transition in complex networks related to a wide range of important problems. Recent advances have provided us an analytical framework of core percolation in uncorrelated random networks with arbitrary degree distributions. Here we apply the tools in analysis of network controllability. We confirm analytically that the emergence of the bifurcation in control coincides with the formation of the core and the structure of the core determines the control mode of the network. We also derive the analytical expression related to the controllability robustness by extending the deduction in core percolation. These findings help us better understand the interesting interplay between the structural and dynamical properties of complex networks.

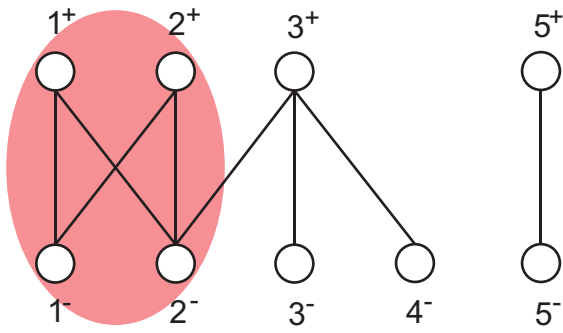
Correspondence and requests for materials should be addressed to T.J. (taojia82@gmail.com)

One fundamental challenge in network science is to understand the impact of the structural property on its functionality. In the last decade considerable advances have been made, particularly on the structural transitions that can bring big impacts on numerous dynamical processes on networks. A lot have been discovered, such as the application of  $k$ -core percolation<sup>1–4</sup> and giant components<sup>5–9</sup> in analysis of epidemic and information spreading on socio-technical systems<sup>10–12</sup>, the use of dominating set in disease outbreak detection, control and social influence propagation<sup>13–17</sup> and the fragility in many real networks caused by their multilayer connections<sup>18–20</sup>. Nevertheless, a lot remains unknown. Core percolation, as one example, is a structural transition in complex networks with a long history. The core represents the remainder of the greedy leaf removal (GLR) procedure in a network<sup>21,22</sup>. While core percolation has applications in several important problems such as conductor-insulator transitions, maximum matching and minimum vertex cover problem<sup>21,23–26</sup>, the physical importance of the core and how the core structure would affect the dynamics of a network are not entirely understood.

Recent advances have brought us an analytical framework of core percolation in uncorrelated random networks with arbitrary degree distribution<sup>27</sup>. The tools introduced not only allow us to predict the emergence of the core but also calculate the expected core size. These findings reveal some interesting interplay between the core and controllability of complex networks. For example, it is observed that the sudden change in controllability robustness and the emergence of the two control modes coincide with the formation of the core<sup>28,29</sup>, suggesting a strong connection between the two topics. Here we analytically explore this connection. The rest of the paper is organized as follows. In the first two sub-sessions, we briefly review the analytical framework of the core percolation and the basic concepts of network controllability. In the third sub-session, we demonstrate the role of the core in control. The core structure determines not only the control mode but also the stability of the control mode under structure perturbations. Finally, we study the controllability robustness in the fourth sub-session, the ability to maintain the control under nodes' failures. By applying the tools in core percolation, we obtain an analytical expression of the fraction of nodes playing different roles in sustaining the controllability of the network.

## Results

**Analytical Framework of Core Percolation.** In core percolation leaf nodes and their neighbors are taken off iteratively from the network according to GLR procedure<sup>21,22</sup>. Specifically, a node with degree one is randomly chosen. This node and its neighbor are removed with all their links. Nodes that becomes isolated are also removed. This procedure is repeated until no node with degree one is left and a core emerges as a compact cluster of nodes left. To systematically study the core percolation, two categories of removable nodes are introduced<sup>27</sup>:  $\alpha$  removable



**Figure 1** | A bipartite graph with a core (highlighted). Node  $1^+$ ,  $2^+$ ,  $1^-$  and  $2^-$  are the core nodes. Node  $3^-$  and  $4^-$  are  $\alpha$  removable because they can become isolated without directly removing themselves (e.g. when node  $3^-$  is chosen in the GLR procedure, node  $3^-$  and  $3^+$  with all their links are removed and node  $4^-$  becomes isolated). Node  $3^+$ ,  $5^+$  and  $5^-$  are  $\beta$  removable as they can become a neighbor of a leaf.

as nodes that can become isolated without directly removing themselves (e.g. node  $3^-$  and  $4^-$  in Fig. 1),  $\beta$  removable as nodes that can become a neighbor of a leaf (e.g. node  $3^+$ ,  $5^+$  and  $5^-$  in Fig. 1). The category of a node  $i$  in an arbitrary graph  $\mathcal{G}$  can be determined by the categories of the neighbor nodes in the subgraph  $\mathcal{G}\setminus i$  where node  $i$  and all its links are excluded: node  $i$  is  $\alpha$  removable in  $\mathcal{G}$  if all its neighbor nodes are  $\beta$  removable in  $\mathcal{G}\setminus i$  and  $\beta$  removable in  $\mathcal{G}$  if at least one neighbor is  $\alpha$  removable in  $\mathcal{G}\setminus i$ . Correspondingly the fraction of  $\alpha$  and  $\beta$  nodes can be expressed as

$$\begin{aligned} n_\alpha &= \sum_k P(k) \tilde{\beta}^k = G(\tilde{\beta}) \\ n_\beta &= 1 - \sum_k P(k) (1 - \tilde{\alpha})^k = 1 - G(1 - \tilde{\alpha}), \end{aligned} \tag{1}$$

where  $P(k)$  is the network's degree distribution,  $G(x) = \sum_k P(k)x^k$  is the generating function of  $P(k)$  and  $\tilde{\alpha}$ ,  $\tilde{\beta}$  are respectively the probability that we find an  $\alpha$  and  $\beta$  nodes at the end of a random chosen link and in the absence of that link (in  $\mathcal{G}\setminus i$ ). These two parameters have been found as

$$\begin{aligned} \tilde{\alpha} &= \sum_k Q(k) \tilde{\beta}^{k-1} = A(1 - \tilde{\beta}) \\ \tilde{\beta} &= 1 - \sum_k Q(k) (1 - \tilde{\alpha})^{k-1} = 1 - A(\tilde{\alpha}), \end{aligned} \tag{2}$$

in which  $Q(k) = kP(k)/\langle k \rangle$  is the excess degree distribution,  $\langle k \rangle$  is the average degree and  $A(x) \equiv \sum_k Q(k+1)(1-x)^k$ . On the basis of Eq.(2),  $\tilde{\alpha}$  can be solved as the smallest root of the function  $f(x) \equiv A(A(x)) - x$ , which can be further used to calculate  $\tilde{\beta}$ . The parameters  $\tilde{\alpha}$  and  $\tilde{\beta}$  can then be applied in determining the fraction of  $\alpha$  and  $\beta$  nodes outlined in Eq.(1).

To generalize core percolation to a directed network, different removal procedures are introduced<sup>27,30</sup>. Here we adopt the one that converts the direct network to a bipartite graph, which splits a node  $i$  in directed network into two nodes  $i^+$  (upper) and  $i^-$  (lower) (Fig. 2a, b). A directed link from node  $i$  to  $j$  becomes a connection from node  $i^+$  to node  $j^-$  in the bipartite representation. The out- and in-degree distribution become degree distribution  $P^+(k)$  and  $P^-(k)$  in the bipartite graph, respectively. Correspondingly Eq.(2) becomes

$$\begin{aligned} \tilde{\alpha}^\pm &= A^\pm(1 - \tilde{\beta}^\mp) \\ \tilde{\beta}^\pm &= 1 - A^\pm(\tilde{\alpha}^\mp), \end{aligned} \tag{3}$$

and  $\tilde{\alpha}^\pm$  is the smallest root of function  $f^\pm(x) \equiv A^\pm(A^\mp(x)) - x$  where  $A^\pm(x) \equiv \sum_k Q^\pm(k+1)(1-x)^k$ . The fraction of  $\alpha$  and  $\beta$  nodes in + and - set is

$$\begin{aligned} n_\alpha^\pm &= \sum_k P^\pm(k) (\tilde{\beta}^\mp)^k = G^\pm(\tilde{\beta}^\mp) \\ n_\beta^\pm &= 1 - \sum_k P^\pm(k) (1 - \tilde{\alpha}^\mp)^k = 1 - G^\pm(1 - \tilde{\alpha}^\mp), \end{aligned} \tag{4}$$

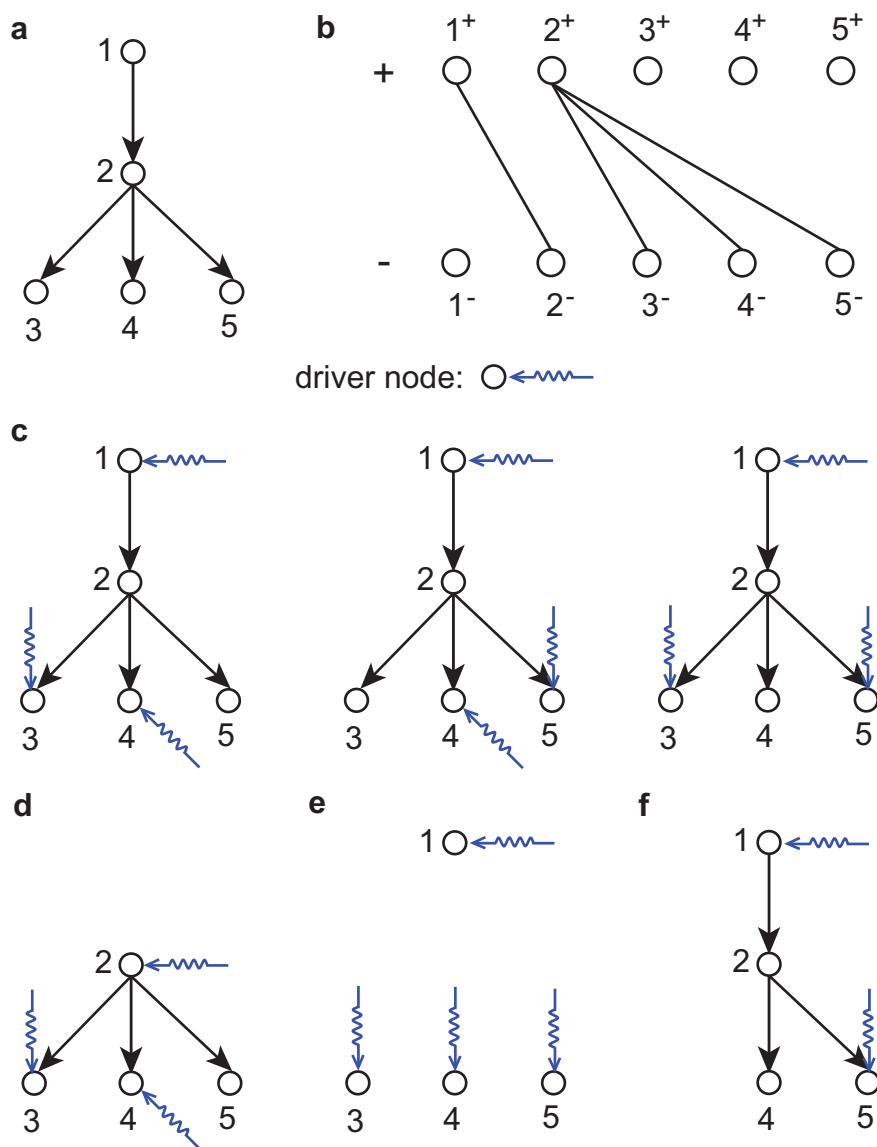
It is noteworthy that in sparse networks without cores, all nodes are either  $\alpha$  or  $\beta$  removable therefore  $n_\alpha^\pm + n_\beta^\pm = 1$ . As the network becomes denser, a core emerges after the average degree exceeds a critical value. At the formation of the core, different categories of removable nodes and the unremovable core nodes appear, making  $n_\alpha^\pm + n_\beta^\pm < 1$ . Nevertheless, parameters  $\tilde{\alpha}^\pm$  and  $\tilde{\beta}^\pm$  used to find  $n_\alpha^\pm$  and  $n_\beta^\pm$  can also be used to find the fraction of core nodes in the + and - set as

$$n_{\text{core}}^\pm = \sum_{k=0}^\infty P^\pm(k) \sum_{s=2}^k \binom{k}{s} (\tilde{\beta}^\mp)^{k-s} (1 - \tilde{\beta}^\mp - \tilde{\alpha}^\mp)^s. \tag{5}$$

**Controllability of Complex Networks.** The controllability of complex systems is a fundamental challenge of contemporary science that draws considerable interests in multidisciplinary fields<sup>28,29,31-50</sup>. According to control theory<sup>51,52</sup>, the dynamic process of controlling a linear time-invariant system can be described by the equation  $dx(t)/dt = Ax(t) + Bu(t)$ , where the state vector  $\mathbf{x}(t) = (x_1(t), \dots, x_N(t))^T$  captures the state of a system of  $N$  components at time  $t$ . The  $N \times N$  state matrix  $A$  corresponds to the internal interactions of the system. The input matrix  $B$  is an  $N \times M$  matrix indicating how the  $M$  external signals  $\mathbf{u}(t) = (u_1(t), \dots, u_M(t))^T$  are exerted to the system to drive it from any initial state to any desired final state within finite time. Recently an efficient methodology has been introduced to identify the minimum driver node set (MDS), the smallest set of nodes whose time-dependent control yields the control over the whole system<sup>28</sup>. The procedure is to convert a directed network into a bipartite graph (Fig. 2a, b) and find the maximum matching of bipartite graph<sup>53</sup>. The minimum driver nodes are the unmatched nodes in the - set. If a perfect matching exists and all nodes in the - set are matched, one input signal would be sufficient to drive the system. In this case the number of driver node is one.

The methodology proposed indicates the existence of multiple MDSs, hence a node does not necessarily participate in all MDSs (Fig. 2c)<sup>32</sup>. Accordingly a node can be categorized by its participation in control: *critical* if it participates in all MDSs, *redundant* if it is not included in any MDSs and *intermittent* if it is in some MDSs but not all. The fraction of critical nodes  $n_c$  is purely determined by the in-degree distribution as  $n_c = P_{\text{in}}(0)$ . However, the fraction of redundant nodes  $n_r$  displays a bifurcation after some critical average degree  $\langle k_c \rangle$ : networks with identical degree distribution and average degree can have very high or low value of  $n_r$  (Fig. 3, 4). Such a bimodality feature leads to the two distinct control modes, with significant difference in the total number of MDS choices. While the two control modes coexist with equal probability in dense networks with identical in- and out-degree distributions, networks with different in- and out-degree distribution may have one mode dominate or follow only control mode<sup>29</sup>.

**Core Percolation and the Bimodality in Control.** The emergence and the absence of bimodality can be best explained using the knowledge of core percolation. Indeed,  $\beta$  nodes, as the neighbor of a leaf node, are always matched in all possible maximum matching configurations and  $\alpha$  nodes are not. Therefore  $\beta$  nodes in the - set are redundant nodes and  $\alpha$  nodes in the - set are not redundant.



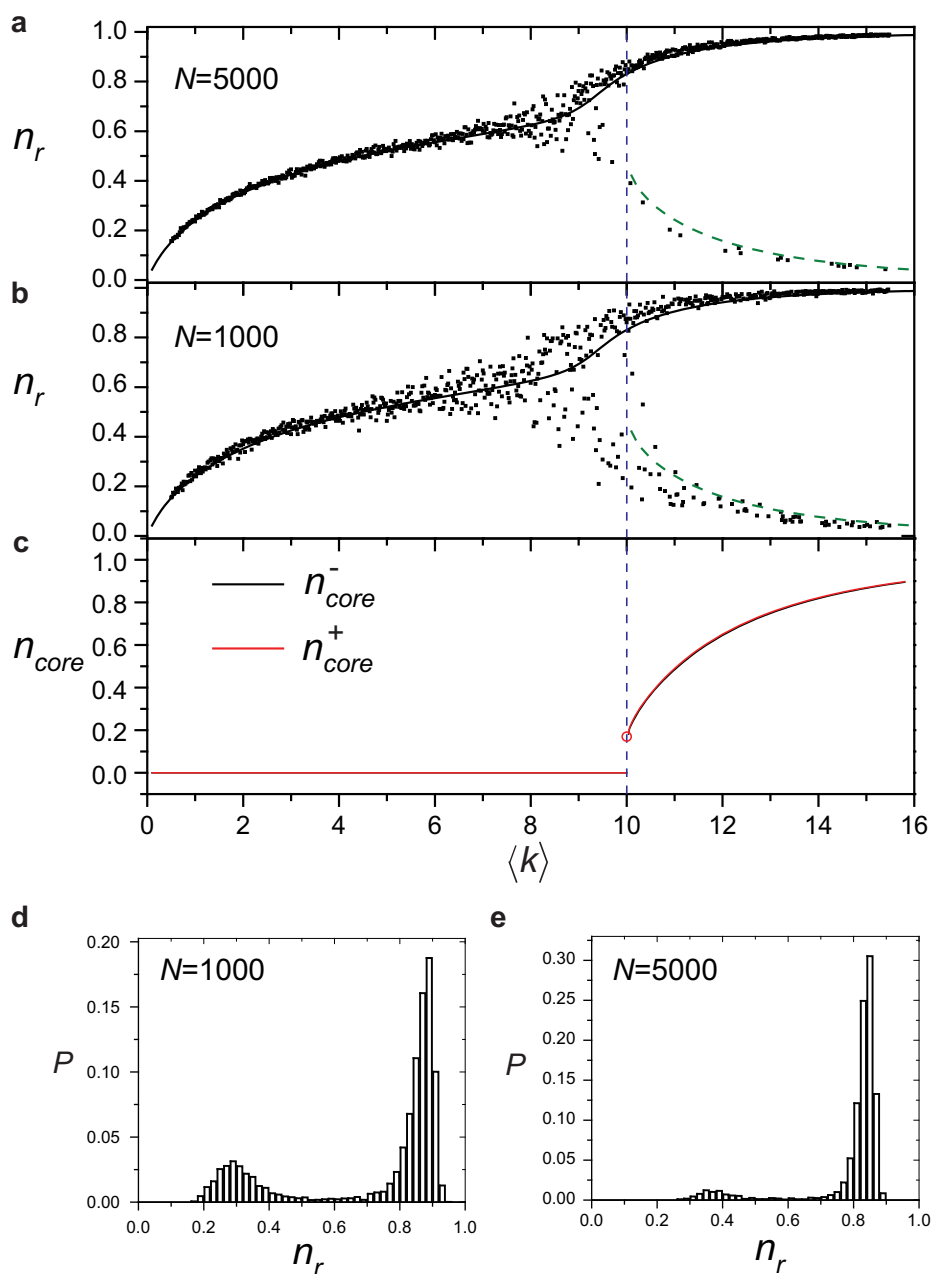
**Figure 2** | (a) An example of a directed network with five nodes. (b) The bipartite representation of the directed network in (a) where a node  $i$  in the directed network is split into two nodes  $i^+$  (upper) and  $i^-$  (lower). A directed link from node 1 to node 2 in (a) corresponds to a connection between node  $1^+$  and node  $2^-$ . (c) Three different control configurations to control the network in (a), indicating different participations of nodes in control: node 1 is *critical* as it is a driver node in all cases, node 2 is *redundant* because it does not participate in any of the driver node sets and nodes 3, 4 and 5 are *intermittent* as they are driver nodes in some situations but not all. (d–f) The control of the network in (a) after removing one node and all its links. (e) Node 1 is *structurally redundant* as its removal does not change the number of driver node. (d) Node 2 is *structurally critical* because more driver nodes are needed in its absence. (f) The number of driver nodes decreases by 1 without node 3, therefore node 3 is *structurally ordinary*.

Before the formation of the core, a node is either  $\alpha$  or  $\beta$  node, hence the fraction of redundant nodes  $n_r = n_\beta^- = 1 - n_\alpha^-$  where  $n_\alpha^-$  and  $n_\beta^-$  can be found by Eqs.(3) and (4).

When core emerges, however,  $\alpha$  and  $\beta$  nodes are not the only nodes in the network ( $n_\alpha^\pm + n_\beta^\pm < 1$ ) and  $n_r$  depends on the shape of the core. When there are fewer core nodes in the  $+$  set than in the  $-$  set ( $N_{\text{core}}^+ < N_{\text{core}}^-$ ), all the core nodes in the  $-$  set are not always matched. This corresponds to the lower branch of the bifurcation with  $n_r = n_\beta^-$  as only the  $\beta$  nodes in the  $-$  set are redundant (Fig. 4b, d). When  $N_{\text{core}}^+ \geq N_{\text{core}}^-$ , however, all the core nodes in the  $-$  set are always matched, giving rise to a large number of always matched nodes. In this case, the network is on the upper branch of the bifurcation curve and  $n_r = 1 - n_\alpha^-$  as only  $\alpha$  nodes in the  $-$  set are not redundant (Fig. 4a, c). In summary, depending on the core structure,  $n_r$  can be high or low as

$$n_r = \begin{cases} 1 - n_\alpha^-, & \text{if } N_{\text{core}}^+ \geq N_{\text{core}}^- \\ n_\beta^-, & \text{if } N_{\text{core}}^+ < N_{\text{core}}^- \end{cases} \quad (6)$$

Eq.(6) confirms the emergence of bifurcation at the formation of the core (Fig. 3). More importantly, it explains the condition for the coexistence of the two control modes. When the in- and out-degree distribution are the same, the expected  $n_{\text{core}}^+$  and  $n_{\text{core}}^-$  are identical, giving the same probability of  $N_{\text{core}}^+ \geq N_{\text{core}}^-$  and  $N_{\text{core}}^+ < N_{\text{core}}^-$  (the probability that  $N_{\text{core}}^+ = N_{\text{core}}^-$  is negligible in large systems). Hence the network is equally likely to be on either of the control mode. When the in- and out-degrees are asymmetric, however, Eq.(5) gives different  $n_{\text{core}}^+$  and  $n_{\text{core}}^-$  values<sup>27</sup>, implying that in sufficiently large networks where the mean-field equation Eq.(5) applies, only one type of the core structure is allowed. Consequently any tiny differences between the in- and out-degree distribution eliminates the

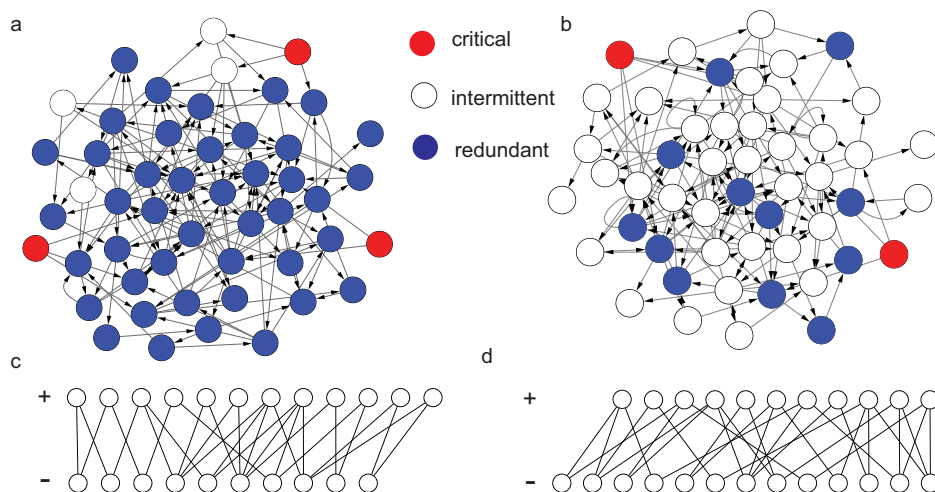


**Figure 3** | (a–c) Dependency between the core and the control modes for networks with  $P_{out,in} \sim k^{-\gamma_{out,in}}$ ,  $\gamma_{out} = 3$ ,  $\gamma_{in} = 2.85$  and different sizes. The bifurcation emerges at the formation of the core. The solid lines in (a) and (b) correspond to Eq.(6) with the expected core structure  $N_{core}^+ \geq N_{core}^-$ . The dashed lines are based on Eq.(6) when  $N_{core}^+ < N_{core}^-$ , as the minority of network realizations different from expectation. (c) The expected fraction of core nodes on the + and - set ( $n_{core}^+$  and  $n_{core}^-$ ) based on Eq.(5). The difference between  $n_{core}^+$  and  $n_{core}^-$  is very small ( $\approx 0.005$ ) and the two curves of  $n_{core}^+$  and  $n_{core}^-$  almost overlap. (d,e) The distribution of  $n_r$  of an ensemble of network realizations with  $P_{out,in} \sim k^{-\gamma_{out,in}}$ ,  $\gamma_{out} = 3$ ,  $\gamma_{in} = 2.85$ ,  $\langle k \rangle = 10$  and different sizes. As the networks size increases, the core structure is closer to the expectation, eliminating the lower branch of the bifurcation curve.

bifurcation of  $n_r$  and forces the large network to follow one or the other control mode<sup>29</sup>. Yet, in small systems the mean-field equation may not hold and the core structure different from the expectation can emerge due to random fluctuations. As an example, networks with  $P_{out,in} \sim k^{-\gamma_{out,in}}$ ,  $\gamma_{out} = 3$  and  $\gamma_{in} = 2.85$  have very small difference between  $n_{core}^+$  and  $n_{core}^-$  ( $n_{core}^+ - n_{core}^- \approx 0.005$  Fig. 3c). While  $N_{core}^+ > N_{core}^-$  is expected, networks with size  $N = 1000$  also generate cores with  $N_{core}^+ < N_{core}^-$  and the two modes coexists with one mode that dominates in probability (Fig. 3b, d). As the network size increases (e.g.  $N = 5000$ ), the expected difference between  $N_{core}^+$  and  $N_{core}^-$  is more significant and networks that fall into the other

branch become very rare (Fig. 3a, e). Eventually the gap between  $N_{core}^+$  and  $N_{core}^-$  exceeds fluctuations and one branch will vanish as the system size increases, effectively forcing the system into only one control mode.

The above results reveal the importance of core as a fundamental structure that controls the two control modes. This feature allows us to switch the control modes by changing the core balance. The most intuitive way to induce such a switch is to reverse the direction of all links in the network. A network originally with  $N_{core}^+ > N_{core}^-$  that is on the upper branch of the bifurcation will have  $N_{core}^+ < N_{core}^-$  in the transpose network staying on the lower branch and vice versa. While



**Figure 4** | (a), (b) Two networks on different control modes, with high (a) and low (b) fraction of redundant nodes. (c), (d) The core of the directed network in (a) and (b), respectively. Network with a high fraction of redundant nodes (a) yields a core with  $N_{\text{core}}^+ > N_{\text{core}}^-$  (c), and vice versa.

the change of all links' directions is too drastic, sometimes the switch of the control mode can be induced by only local changes, occasionally as little as flipping the direction of a single well chosen link<sup>29</sup>. Our chance of finding such a link depends on the core structure: if  $N_{\text{core}}^+$  and  $N_{\text{core}}^-$  is close, the network is sensitive to link changes. Indeed in a sample of 1000 realizations of Erdős-Rényi networks<sup>5</sup>, all networks that are able to switch the control modes via a link's reversal have very close core size (Fig. 5a, b). As the core size depends on the network size and degree distribution, the control mode is more stable in large networks and networks with asymmetric in- and out-degrees, in which more structural changes are required to change the control mode. Small networks with identical in- and out-degree distribution, however, are more likely to have close  $N_{\text{core}}^+$  and  $N_{\text{core}}^-$  therefore sensitive to structural perturbations (Fig. 5c, d).

**Core Percolation and the Robustness of Control.** The control mode can be sensitive to structural changes, but the network controllability is relatively robust. Removing a node or link can only change the number of driver node  $N_D$  by 1<sup>41,42,54</sup>. To measure the importance of nodes in sustaining the controllability over the network, a different set of node category has been introduced<sup>28</sup>. A node is *structurally critical* if the number of driver nodes has to be increased to maintain full control in its absence ( $N_D$  increases by 1), a node is *structurally redundant* if it can be removed without affecting the current set of driver nodes ( $N_D$  does not change) and a node is *structurally ordinary* if it is neither structurally critical or structurally redundant ( $N_D$  decreases by 1) (Fig. 2d). While the fraction of nodes in each category can be numerically studied, the analytical approach is missing.

Here we use the tools in core percolation to derive the analytical expression of the fraction of nodes in each category. In the controllability problem, a node has dual roles. On one hand, a node's dynamics is controlled via internal or external channels pointed to it. On the other hand, a node serves as means to control other neighboring nodes it points to. Such dual roles can be best seen in a network's bipartite representation. The nodes in the + set can be considered as "superiors" that influence others internally and the nodes in the - set are "subordinates" that need to be controlled. Accordingly the consequence of a node's removal relies on the node's role in both + and - set. If a node in the - set is always matched, its removal will not change the number of unmatched nodes in the - set. Otherwise, the number of unmatched nodes in the - set will decrease by 1. Similarly if a node in the + set is always matched, it matches a node in the - set in all matchings. Removing this node

increases the number of unmatched nodes in the - set by 1. Otherwise, the number of unmatched nodes in the - set will not change as there exists alternative configurations matching the same number of nodes. The impact of node  $i$ 's removal (node  $i^+$  and node  $i^-$  in the bipartite graph) is summarized in Table 1. As the number of driver nodes equals the number of unmatched nodes in the - set, we readily have the relationship between a node's structural role in control and its matching status. A node  $i$  in a directed network is structurally critical if in its bipartite representation both nodes  $i^-$  and  $i^+$  are always matched. A node  $i$  is structurally ordinary if neither node  $i^-$  nor  $i^+$  is always matched. Otherwise node  $i$  is structurally redundant.

The expression of always matched nodes in the - set ( $n_r$ ) is obtained in Eq.(6). With the symmetry between the + and - sets, we can find the expression of structurally critical, ordinary and redundant nodes as

$$n_{\text{sc}} = \begin{cases} n_{\beta}^+ (1 - n_{\alpha}^-), & \text{if } N_{\text{core}}^+ \geq N_{\text{core}}^- \\ n_{\beta}^- (1 - n_{\alpha}^+), & \text{if } N_{\text{core}}^+ < N_{\text{core}}^- \end{cases} \quad (7)$$

$$n_{\text{so}} = \begin{cases} (1 - n_{\beta}^+) n_{\alpha}^-, & \text{if } N_{\text{core}}^+ \geq N_{\text{core}}^- \\ (1 - n_{\beta}^-) n_{\alpha}^+, & \text{if } N_{\text{core}}^+ < N_{\text{core}}^- \end{cases} \quad (8)$$

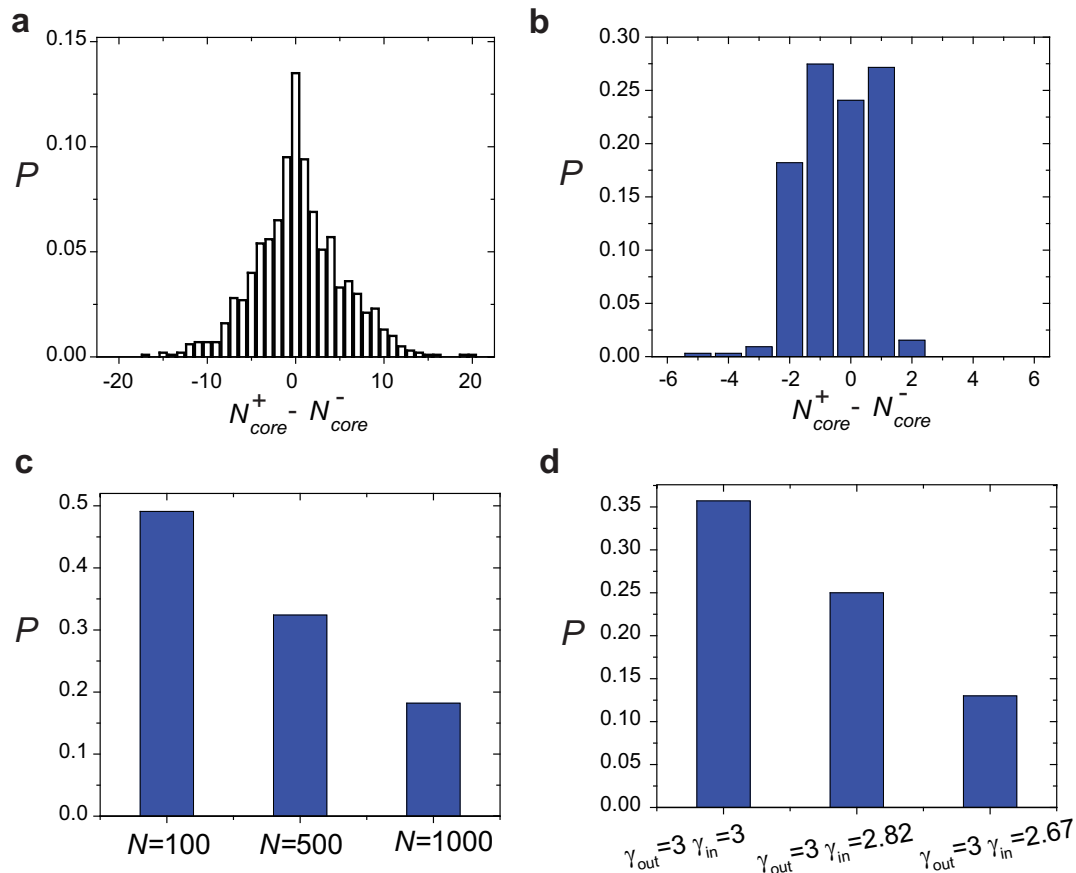
and

$$n_{\text{sr}} = 1 - n_{\text{sc}} - n_{\text{so}}, \quad (9)$$

where  $n_{\text{sc}}$ ,  $n_{\text{so}}$  and  $n_{\text{sr}}$  are fractions of structurally critical, structurally ordinary and structurally redundant nodes.

The results of Eqs.(7-9) are numerically tested (Fig. 6a). There is a sudden change of  $n_{\text{sc}}$ ,  $n_{\text{so}}$  and  $n_{\text{sr}}$  on the formation of the core ( $\langle k_c \rangle = 2e$  in the Erdős-Rényi network<sup>22,30</sup>) accompanied by the change of fraction of  $\alpha$  and  $\beta$  nodes. Note that whether a node is a driver node depends on the matching in the - set only. The always matched nodes in the - set can be high or low depending on the two structures of the core. This generates the bifurcation feature where nodes' participations in control differ dramatically. When a node's role in controllability robustness is concerned, however, it depends on matching in both the - and + sets. Therefore  $n_{\text{sc}}$ ,  $n_{\text{so}}$  or  $n_{\text{sr}}$  do not show a bifurcation feature. Indeed, when the in- and out-degree distribution are the same,  $N_{\text{core}}^+ \geq N_{\text{core}}^-$  and  $N_{\text{core}}^+ < N_{\text{core}}^-$  are equally likely, but  $\beta^+ = \beta^-$  and  $\alpha^+ = \alpha^-$ , giving rise to a single value of  $n_{\text{sc}}$ ,  $n_{\text{so}}$  and  $n_{\text{sr}}$  regardless of the shape of the core. When the in- and





**Figure 5** | (a) Distribution of  $N_{core}^+ - N_{core}^-$  in 1000 realizations of Erdős-Rényi networks with size  $N = 500$  and  $\langle k \rangle = 6$ . (b) Among these samples in (a), the distribution of  $N_{core}^+ - N_{core}^-$  conditioning on the existence of a single link whose reversal causes a switch of control mode. The networks sensitive to structural perturbations are typically with very close number of core nodes on each set. (c) The chance to switch the control mode by flipping one link's direction decreases with network size. The statistics are based on 1000 realizations of Erdős-Rényi networks with  $\langle k \rangle = 6$  and different size  $N$ . (d) The chance to switch the control mode by flipping one link's direction decreases with degree asymmetry. The statistics are based on 1000 realizations of scale-free networks with  $P_{out,in} \sim k^{-\gamma_{out,in}}$ ,  $\langle k \rangle = 12$  and  $N = 500$ .

out-degree distribution are different, only one condition ( $N_{core}^+ \geq N_{core}^-$  or  $N_{core}^+ > N_{core}^-$ ) is allowed in infinite networks, generating only one value of  $n_{sc}$ ,  $n_{so}$  and  $n_{sr}$ . Small networks with different in- and out-degree distribution do have cores with  $N_{core}^+ \geq N_{core}^-$  or  $N_{core}^+ > N_{core}^-$  (Fig. 3). But the difference in the expected value of  $n_{sc}$ ,  $n_{so}$  and  $n_{sr}$  in the two cases are typically small (Fig. 6b). As the result no obvious bifurcation curve can be observed with the random noise due to the small system size.

## Discussion

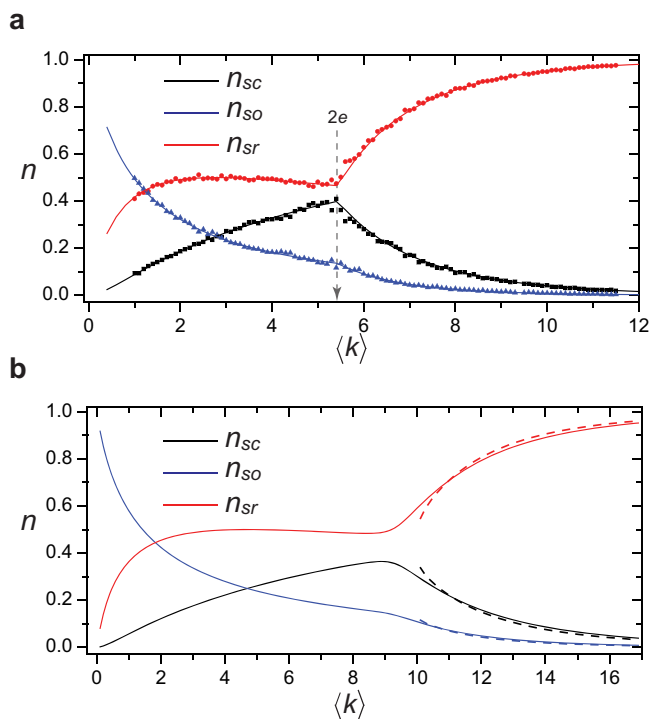
In summary, we apply the recent advances in core percolation to the controllability of complex networks. Historically, core percolation has been discovered to be related to a wide range of important problems in complex networks. Here we add a new connection to the network controllability. In particular, we reveal the importance of core as a fundamental structure that generates the two control modes. The core structure determines the control mode that is related to the participation of nodes in control under the minimum driver nodes. The stability of the control mode under structural perturbations also relies on the balance of the core. Moreover, we derive the analytical expression of the fraction of nodes with different controllability robustness. The expression obtained demonstrates dependency on the structure of the core.

The results presented raise several intriguing questions awaiting answers. For example, it is found that switching the balance of the core is crucial in changing the control modes. However, an efficient

algorithm to identify a series of structure variations needed to change the core size is still missing. We lack the method to change the control mode in arbitrary networks. The calculation in controllability robustness is based on uncorrelated in- and out-degree distribution. The effects of higher order correlations require further investigations. Finally, the analytical framework of core percolation is limited to model networks. In many real systems, core is quite different from the analytical expectations. For example, many real networks have multiple pieces of cores which is not observed in model networks, many real networks containing cores are not dense enough to yield cores in theory and for those dense networks, the analytical prediction of the core size can be off. The theoretical work in this paper can not be simply generalized to real networks without proper modifications. Such studies are left for future work.

**Table 1** | Impact of node  $i$ 's removal on the number of unmatched nodes in the  $-$  set. Node  $i$  in the directed network corresponds to node  $i^+$  and  $i^-$  in its bipartite representation. Number 0 means no change and  $\pm 1$  means the number of unmatched nodes in the  $-$  set increases and decreases by 1 respectively

Node	Always Matched	Not Always Matched
$i^-$	0	-1
$i^+$	+1	0



**Figure 6** | (a)  $n_{sc}$ ,  $n_{so}$  and  $n_{sr}$  in Erdős-Rényi networks with  $N = 10000$  and different average degree  $\langle k \rangle$ . The solid lines correspond to Eqs.(7–9). There is a sudden change at  $\langle k \rangle = 2e$  when the core emerges. (b) Analytical results of  $n_{sc}$ ,  $n_{so}$  and  $n_{sr}$  based on Eqs.(7–9) in networks with  $P_{out,in} \sim k^{-\gamma_{out,in}}$ ,  $\gamma_{out} = 3$ ,  $\gamma_{in} = 2.85$  and different  $\langle k \rangle$ . The solid lines correspond to the expected core structures ( $N_{core}^+ \geq N_{core}^-$ ) and the dashed lines are the results for core structures different from the expectation ( $N_{core}^+ < N_{core}^-$ ). The difference of the results between the two core structures are small. Therefore even though in small systems both core structures are possible, there is no bifurcation observed.

## Methods

**Generating a scale free network.** The scale-free networks<sup>55</sup> analyzed are generated via the static model<sup>56</sup>. We start from  $N$  disconnected nodes indexed by integer number  $i$  ( $i = 1, \dots, N$ ). The weight  $w_i^\pm = i^{-\alpha^\pm}$  is assigned to each node in the  $out$  and the  $in$  set, with  $\alpha^\pm$  a real number in the range  $[0, 1)$ . Randomly selected two nodes  $i$  and  $j$  respectively from the  $out$  set and the  $in$  set, with probability proportional to  $w_i^+$  and  $w_j^-$ . Connect node  $i$  and  $j$  if there is no connection between them, corresponding to a directed link from node  $i$  to node  $j$  in the digraph. Otherwise randomly choose another pair. Repeated the procedure until  $\langle k_{in} \rangle = \langle k_{out} \rangle = \langle k \rangle / 2$  links are created. The degree distribution under this construction is

$$P^\pm(k) = \frac{[(k)(1 - \alpha^\pm)/2]^{1/\alpha^\pm} \Gamma(k - 1/\alpha^\pm, (k)(1 - \alpha^\pm)/2)}{\alpha^\pm \Gamma(k+1)}$$

with  $\Gamma(s)$  the gamma function and  $\Gamma(s, x)$  the upper incomplete gamma function. In the large  $k$  limit, the distribution becomes  $P^\pm(k) \sim k^{-(1+\alpha^\pm)} = k^{-\gamma^\pm}$ .

**Eliminating correlations between in- and out-degree distribution.** The calculation of controllability robustness is based on the assumption that a node's role in the  $+$  set is independent of its role in the  $-$  set, which requires independent in- and out-degree distribution. The scale free network directly generated by static model has degree correlations. For example, node 1 in both the  $+$  and  $-$  set have the largest expected degree. To eliminate the degree correlation, we randomize the sequence of  $w^\pm$ .

- Pittel, B., Spencer, J. & Wormald, N. Sudden emergence of a giant  $k$ -core in a random graph. *J. Comb. Theory A* **67**, 111–151 (1996).
- Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F.  $k$ -core organization of complex networks. *Phys. Rev. Lett.* **96**, 040601 (2006).
- Goltsev, A. V., Dorogovtsev, S. N. & Mendes, J.  $k$ -core (bootstrap) percolation on complex networks: Critical phenomena and nonlocal effects. *Phys. Rev. E* **73**, 056101 (2006).
- Zhao, J.-H., Zhou, H.-J. & Liu, Y.-Y. Inducing effect on the percolation transition in complex networks. *Nat. Commun.* **4**, 2412 (2013).

- Erdős, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–60 (1960).
- Bollobás, B. *Random Graphs* (Cambridge University Press, Cambridge, 2001).
- Callaway, D. S., Newman, M. E. J., Strogatz, S. H. & Watts, D. J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.* **85**, 5468–5471 (2000).
- Newman, M. E. J., Strogatz, S. H. & Watts, D. J. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118 (2001).
- Newman, M., Barabási, A.-L. & Watts, D. J. *The Structure and Dynamics of Networks* (Princeton University Press, Princeton, 2006).
- Pastor-Satorras, R., Vázquez, A. & Vespignani, A. Dynamical and correlation properties of the internet. *Phys. Rev. Lett.* **87**, 258701 (2001).
- Kitsak, M. *et al.* Identification of influential spreaders in complex networks. *Nat. Phys.* **6**, 888–893 (2010).
- Vespignani, A. Modelling dynamical processes in complex socio-technical systems. *Nat. Phys.* **8**, 32–39 (2011).
- Eubank, S., Kumar, V., Marathe, M. V., Srinivasan, A. & Wang, N. Structural and algorithmic aspects of massive social networks. In *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*, 718–727 (2004).
- Kelleher, L. L. & Cozzens, M. B. Dominating sets in social network graphs. *Math. Soc. Sci.* **16**, 267–279 (1988).
- Wang, F. *et al.* On positive influence dominating sets in social networks. *Theor. Comput. Sci.* **412**, 265–269 (2011).
- Molnár, F., Sreenivasan, S., Szymanski, B. K. & Korniss, G. Minimum dominating sets in scale-free network ensembles. *Sci. Rep.* **3**, 1736 (2013).
- Nacher, J. C. & Akutsu, T. Analysis on critical nodes in controlling complex networks using dominating sets. In *Signal-Image Technology & Internet-Based Systems (SITIS), 2013 International Conference on*, 649–654 (IEEE, 2013).
- Gao, J., Buldyrev, S. V., Stanley, H. E. & Havlin, S. Networks formed from interdependent networks. *Nat. Phys.* **8**, 40–48 (2011).
- Li, W., Bashan, A., Buldyrev, S. V., Stanley, H. E. & Havlin, S. Cascading failures in interdependent lattice networks: The critical role of the length of dependency links. *Phys. Rev. Lett.* **108**, 228702 (2012).
- De Domenico, M. *et al.* Mathematical formulation of multilayer networks. *Phys. Rev. X* **3**, 041022 (2013).
- Karp, R. M. & Sipser, M. Maximum matchings in sparse random graphs. *Proc. 22nd Ann. IEEE Symp. Found. Comp.* 364–375 (1981).
- Bauer, M. & Golinelli, O. Core percolation in random graphs: a critical phenomena analysis. *Eur. Phys. J. B* **24**, 339–352 (2001).
- Weigt, M. & Hartmann, A. K. Number of guards needed by a museum: a phase transition in vertex covering of random graphs. *Phys. Rev. Lett.* **84**, 6118 (2000).
- Bauer, M. & Golinelli, O. Exactly solvable model with two conductor-insulator transitions driven by impurities. *Phys. Rev. Lett.* **86**, 2621 (2001).
- Zhou, H. Vertex cover problem studied by cavity method: Analytics and population dynamics. *Eur. Phys. J. B* **32**, 265–270 (2003).
- Zdeborová, L. & Mézard, M. The number of matchings in random graphs. *J. Stat. Mech. Theor. Exp.* **2006**, P05003 (2006).
- Liu, Y.-Y., Csóka, E., Zhou, H. & Pósfai, M. Core percolation on complex networks. *Phys. Rev. Lett.* **109**, 205703 (2012).
- Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. *Nature* **473**, 167–173 (2011).
- Jia, T. *et al.* Emergence of bimodality in controlling complex networks. *Nat. Commun.* **4**, 2002 (2013).
- Azimi-Tafreshi, N., Dorogovtsev, S. N. & Mendes, J. F. Core organization of directed complex networks. *Phys. Rev. E* **87**, 032815 (2013).
- Sun, J. & Motter, A. E. Controllability transition and nonlocality in network control. *Phys. Rev. Lett.* **110**, 208701 (2013).
- Jia, T. & Barabási, A.-L. Control capacity and a random sampling method in exploring controllability of complex networks. *Sci. Rep.* **3**, 2354 (2013).
- Pósfai, M., Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Effect of correlations on network controllability. *Sci. Rep.* **3**, 1067 (2013).
- Nepusz, T. & Vicsek, T. Controlling edge dynamics in complex networks. *Nat. Phys.* **8**, 568–573 (2012).
- 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).
- Kim, K. H. & Sauro, H. M. Adjusting phenotypes by noise control. *PLoS Comput. Biol.* **8**, e1002344 (2012).
- Pósfai, M. & Hövel, P. Phase transition in the controllability of temporal networks. *arXiv* 1312.7595 (2013).
- Yuan, Z., Zhao, C., Di, Z., Wang, W.-X. & Lai, Y.-C. Exact controllability of complex networks. *Nat. Commun.* **4**, 2447 (2013).
- Delpini, D. *et al.* Evolution of controllability in interbank networks. *Sci. Rep.* **3**, 1626 (2013).
- Wang, X. F. & Chen, G. Pinning control of scale-free dynamical networks. *Physica A* **310**, 521–531 (2002).
- Pu, C.-L., Pei, W.-J. & Michaelson, A. Robustness analysis of network controllability. *Physica A* **391**, 4420–4425 (2012).
- 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).
- Komareji, M. & Bouffanais, R. Resilience and controllability of dynamic collective behaviors. *PLoS ONE* **8**, e82578 (2013).



44. Nacher, J. C. & Akutsu, T. Structural controllability of unidirectional bipartite networks. *Sci. Rep.* **3**, 1647 (2013).
45. Galbiati, M., Delpini, D. & Battiston, S. The power to control. *Nat. Phys.* **9**, 126–128 (2013).
46. Pan, Y. & Li, X. Structural Controllability and Controlling Centrality of Temporal Networks. *PLoS ONE* **9**, e94998 (2014).
47. Zhang, J., Yuan, Z. & Zhou, T. Physical limits of feedback noise-suppression in biological networks. *Phys. Biol.* **6**, 046009 (2009).
48. Rodríguez-Caso, C., Corominas-Murtra, B. & Solé, R. V. On the basic computational structure of gene regulatory networks. *Molecular BioSystems* **5**, 1617–1629 (2009).
49. Ruths, J. & Ruths, D. Control profiles of complex networks. *Science* **343**, 1373–1376 (2014).
50. Onnela, J.-P. Flow of control in networks. *Science* **343**, 1325–1326 (2014).
51. Kalman, R. E. Mathematical description of linear dynamical systems. *J. Soc. Indust. and Appl. Math. Ser. A* **1**, 152–192 (1963).
52. Slotine, J.-J. & Li, W. *Applied Nonlinear Control* (Prentice-Hall, 1991).
53. Hopcroft, J. E. & Karp, R. M. An  $n^{5/2}$  algorithm for maximum matchings in bipartite graphs. *SIAM J. Comput.* **2**, 225–231 (1973).
54. Nie, S., Wang, X., Zhang, H., Li, Q. & Wang, B. Robustness of controllability for networks based on edge-attack. *PLoS ONE* **9**, e89066 (2014).
55. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
56. Goh, K.-I., Kahng, B. & Kim, D. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.* **87**, 278701 (2001).

## Acknowledgments

We thank Y.-Y. Liu, E. Csóka, C. Song, D. Wang, A.-L. Barabási, B.K. Szymanski and G. Korniss for discussions. This work was supported in part by the Army Research Laboratory

under Cooperative Agreement Number W911NF-09-2-0053; the Defense Advanced Research Projects Agency under Agreement Number 11645021 and the Defense Threat Reduction Agency award WMD BRBAA07-J-2-0035. M. Pósfai has received funding from the European Union Seventh Framework Programme (FP7/20072013) under grant agreement No. 270833. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

## Author contributions

T.J. and M.P. designed the research. T.J. and M.P. performed numerical simulation and analytical deduction. T.J. prepared the initial draft of the manuscript.

## Additional information

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

**How to cite this article:** Jia, T. & Pósfai, M. Connecting Core Percolation and Controllability of Complex Networks. *Sci. Rep.* **4**, 5379; DOI:10.1038/srep05379 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>