OPEN

Effect of correlations on network controllability

Márton Pósfai[1,2,3,7], Yang-Yu Liu[1,4], Jean-Jacques Slotine[5] & Albert-László Barabási[1,6,7]

[1]Center for Complex Network Research and Department of Physics, Northeastern University, Boston, [2]Department of Physics of Complex Systems, Eötvös University, Budapest, [3]Department of Theoretical Physics, Budapest University of Technology and Economics, Budapest, [4]Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, [5]Non-linear Systems Laboratory, Department of Mechanical Engineering and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, [6]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, [7]Center for Network Science, Central European University, Budapest.

A dynamical system is controllable if by imposing appropriate external signals on a subset of its nodes, it can be driven from any initial state to any desired state in finite time. Here we study the impact of various network characteristics on the minimal number of driver nodes required to control a network. We find that clustering and modularity have no discernible impact, but the symmetries of the underlying matching problem can produce linear, quadratic or no dependence on degree correlation coefficients, depending on the nature of the underlying correlations. The results are supported by numerical simulations and help narrow the observed gap between the predicted and the observed number of driver nodes in real networks.

While during the past decade significant efforts have been devoted to understanding the structure, evolution and dynamics of complex networks[1–6], only recently has attention turned to an equally important problem: our ability to control them. Given the problem's importance, recent work has extended the concept of pinning control[7–9] and structural controllability[10–13] to complex networks. Here we focus on the latter approach. A networked system is considered controllable if by imposing appropriate external signals on a subset of its components, called driver nodes, the system can be driven from any initial state to any final state in finite time[14–17]. As the control of a system requires a quantitative description of the governing dynamical rules, progress in this area was limited to small engineered systems. Yet, recently Liu et al.[10] showed that the identification of the minimal number of driver nodes required to control a network, N_D, can be derived from the network topology by mapping controllability[16] to the maximum matching in directed networks[18]. The mapping indicated that N_D is mainly determined by the degree distribution $P(k_{in}, k_{out})$. We know, however, that a series of characteristics, from degree correlations[19–21] to local clustering[22] and communities[23–26], cannot be accounted for by $P(k_{in}, k_{out})$ alone, prompting us to ask: which network characteristics affect the system's controllability?

The three most commonly studied deviations from the random network configuration are (i) clustering, manifested as a higher clustering coefficient C than expected based on the degree distribution[27]; (ii) community structure, representing the agglomeration of nodes into distinct communities, captured by the modularity parameter Q[25]; (iii) degree correlations[28]. First, we motivate our work by showing that network characteristics other than the degree distribution also affect network control. Next, we use numerical simulations to identify the network characteristics that affect controllability, finding that only degree correlations have a discernible effect. We then analytically derive $n_D = N_D/N$ for random networks with a given degree distribution and correlation profile. More detailed calculations are provided in the Supplementary Information Sec. III. Finally, we test our predictions on real networks.

Results

Prediction based on the degree distribution. To motivate our study we compared the observed N_D to the prediction based on the degree sequence for several real networks. For this we randomize each network preserving its degree sequence and we calculate N_D^{rand}, the number of driver nodes for the randomized network. Plotting N_D versus N_D^{rand} on log-log scale indicates that the degree sequence correctly predicts the order of magnitude of N_D despite known correlations[19,20] (Fig. 1a). However, by plotting $n_D = N_D/N$ versus $n_D^{rand} = N_D^{rand}/N$ we observe clear deviations from the degree based prediction (Fig. 1b). Our goal is to understand the origin of these deviations, and the degree to which network correlations can explain the observed n_D.
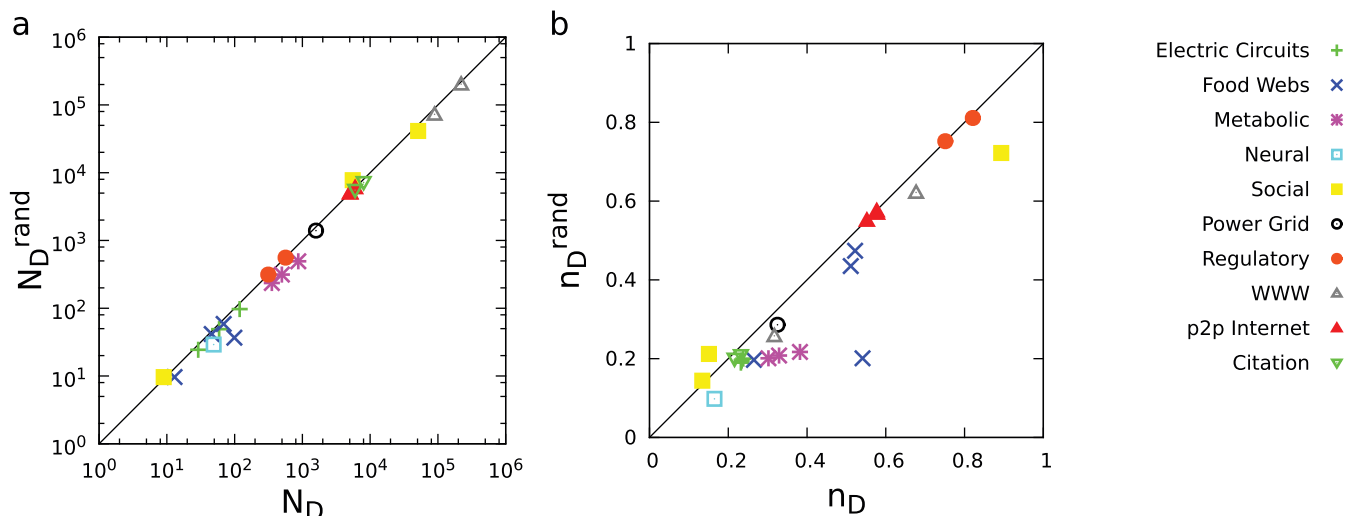
Figure 1 | (a) We compare N_D for real systems to N_D^{rand}, representing the number of driver nodes needed to control their randomized counterparts. Randomization eliminates all local and global correlations, only preserving the degree sequence of the original system. We find that the degree sequence predicts the order of magnitude of N_D correctly, however, small deviations are hidden by the log scale, needed to show the whole span of N_D seen in real systems. (b) These deviations are more obvious if we compare the density of driver nodes $n_D = N_D/N$ and n_D^{rand} in linear scale, finding that for some systems (e.g. regulatory and p2p Internet networks) the degree sequence serves as a good predictor of n_D, while for other systems (e.g. metabolic networks and food webs) n_D deviates from the prediction based solely on the degree sequence.

Numerical simulations. We start from a directed network with Poisson[29,30] or scale-free degree distribution[31,32]. The scale-free network is generated by the static model described in the Methods section. We use simulated annealing to add various network characteristics by link rewiring, while leaving the in- and out-degrees unchanged, tuning each measure to a desired value, for details see the Methods section. We computed n_D using the Hopcroft-Karp algorithm[33].

Clustering. We use the global clustering coefficient[27] defined for directed networks as

$$C = \frac{3 \cdot \text{ number of triangles}}{2 \cdot \text{ number of adjacent edge pairs}}. \quad (1)$$

The simulations indicate that changes in C only slightly alter n_D and that the effect is not systematic (Fig. 2a). Hence we conclude that C plays a negligible role in determining n_D.

Modularity. We quantify the community structure using[25,26]:

$$Q = \frac{1}{E} \sum_{vw} \left[A_{vw} - \frac{k_v^{(in)} k_w^{(out)}}{E} \right] \delta_{c_v, c_w}, \quad (2)$$

where A_{vw} is the adjacency matrix, c_v and c_w are the communities the v and w nodes belong to, respectively. Specifying Q still leaves a great amount of freedom in the number and size of the communities. We therefore choose to randomly divide the nodes into N_C equally sized groups, and increase the edge density within these groups, elevating Q to the desired value.

The simulations indicate that this community structure has no effect on n_D (Fig. 2b). While adding communities to networks can be achieved in many different ways, and the effect of modularity can be explored in more detail (e.g. hierarchical organization of communities[23,34,35], overlapping community structure[24,36], etc), we have failed to detect systematic, modularity induced changes in n_D, prompting us to conclude that Q does not play a leading role in n_D.

Degree correlations. In directed networks each node has an in-degree (k_i) and an out-degree (k_o), thus we can define four correlation coefficients: correlations between the source node's in- and out-degree, and the target node's in- and out-degree (Figs. 3, 4)[28]. We use the Pearson coefficient to quantify each correlation with a single parameter:
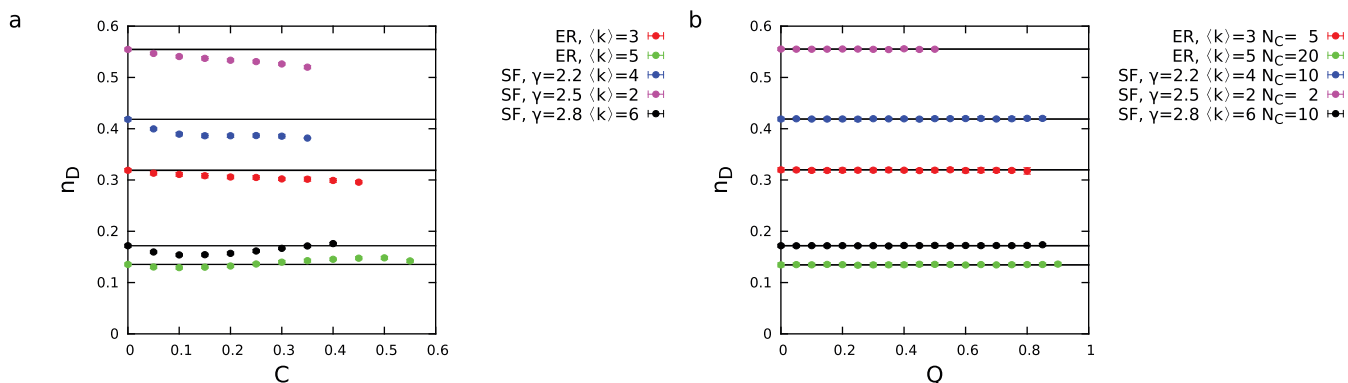

Figure 2 | **Effect of the clustering coefficient C and modularity Q on the density of driver nodes, n_D.** Network size is $N = 10,000$. Each data point is an average over 50 independent runs; the error bars, typically smaller than the symbol size, represent the standard deviation of the measurements.
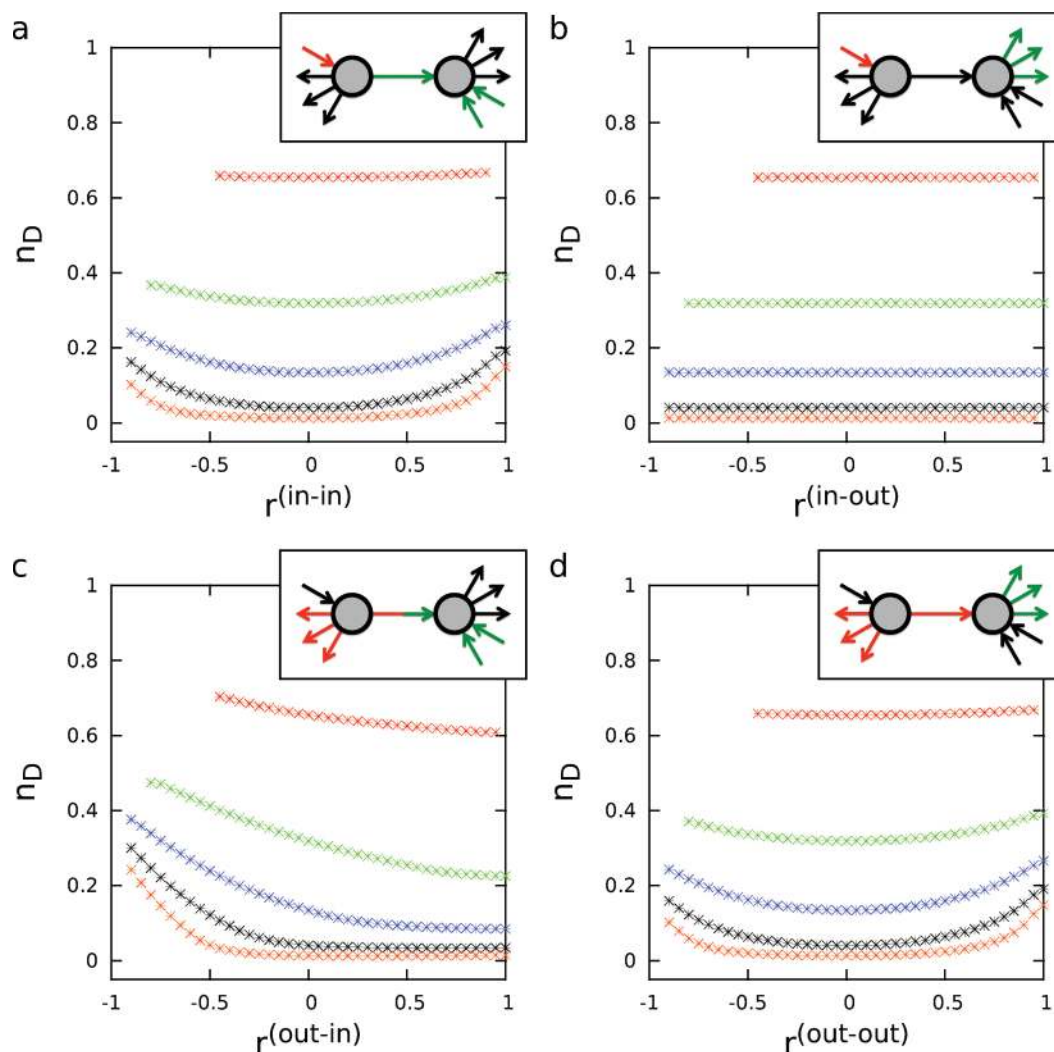
Figure 3 | The impact of degree-degree correlations on the density of driver nodes (n_D) for the Erdős-Rényi model ($N = 10, 000$) for average degrees $\langle k \rangle$ = 1 (red), $\langle k \rangle$ = 3 (green), $\langle k \rangle$ = 5 (blue), $\langle k \rangle$ = 7 (black) and $\langle k \rangle$ = 9 (orange). The results are similar for the scale-free model (see Fig. 4). Each data point is an average of 100 independent runs.

$$r^{(\alpha - \beta)} = \frac{\frac{1}{E} \sum_e \left(k_e^{(\alpha)} - \overline{k^{(\alpha)}} \right) \left(j_e^{(\beta)} - \overline{j^{(\beta)}} \right)}{\sigma^{(\alpha)} \sigma^{(\beta)}}, \qquad (3)$$

where $\sum_e \cdot$ sums over all edges, $\alpha, \beta \in \{\text{in, out}\}$ is the degree type, $k^{(\alpha)}$ is the degree of the source node, $j^{(\beta)}$ is the degree of the target node. And $\overline{j_\alpha} = \frac{1}{E} \sum_e j_e^\alpha$ is the average degree of the nodes at the beginning of each link, $\sigma_\alpha^2 = \frac{1}{E} \sum_e \left(k_e^{(\alpha)} - \overline{k^{(\alpha)}} \right)^2$ is the variance; $\overline{k^{(\beta)}}$ and $\sigma^{(\beta)}$ are defined similarly.

Simulations shown in Figs. 3 and 4 indicate that degree correlations systematically affect n_D. We observe three distinct types of behavior:

(i) n_D depends monotonically on $r^{(\text{out-in})}$, so that low (negative) correlations increase n_D and high (positive) correlations lower n_D (Figs. 3c, 4c);

(ii) Both $r^{(\text{in-in})}$ and $r^{(\text{out-out})}$ increase n_D, independent of the sign of the correlations (Figs. 3a, 3d, 4a, 4d);

(iii) $r^{(\text{in-out})}$ has no effect on n_D (Figs. 3c, 4c).

The behavior is qualitatively the same for Erdős-Rényi (Fig. 3) and scale-free (Fig. 4) networks.

The diversity of these numerical results require a deeper explanation. Therefore in the remaining of the paper we focus on understanding analytically the role of degree correlations, which, by systematically altering n_D, affect the system's controllability.

Analytical framework. The task of identifying the driver nodes can be mapped to the problem of finding a maximum matching of the network[10]. A matching is a subset of links that do not share start or end points. We call a node matched if a link in the matching points at it and we gain full control over a network if we control the unmatched nodes. The cavity method has been successfully used to calculate the size of the maximum matching for undirected[37] and directed[10] network ensembles with given degree distribution. Here we study network ensembles with a given degree correlation profile.

We calculate n_D analytically for a given $P(k_{\text{in}}, k_{\text{out}})$ and selected degree-degree correlation $e(j_{\text{in}}, j_{\text{out}}; k_{\text{in}}, k_{\text{out}})$, representing the probability of a directed link pointing from a node with degrees j_{in} and j_{out} to a node with degrees k_{in} and k_{out}. In the absence of degree correlations (neutral case)

$$e^{(0)}(j_i, j_o; k_i, k_o) = P^{(\text{in})}(j_i) Q^{(\text{out})}(j_o) Q^{(\text{in})}(k_i) P^{(\text{out})}(k_o), \qquad (4)$$

where $Q^{(\text{out})}(j_o) = \frac{2j_o}{\langle k \rangle} P^{(\text{out})}(j_o)$, $Q^{(\text{in})}(k_i) = \frac{2k_i}{\langle k \rangle} P^{(\text{in})}(k_i)$ and $\langle k \rangle$ is the average degree. To ensure analytical tractability we chose[21]
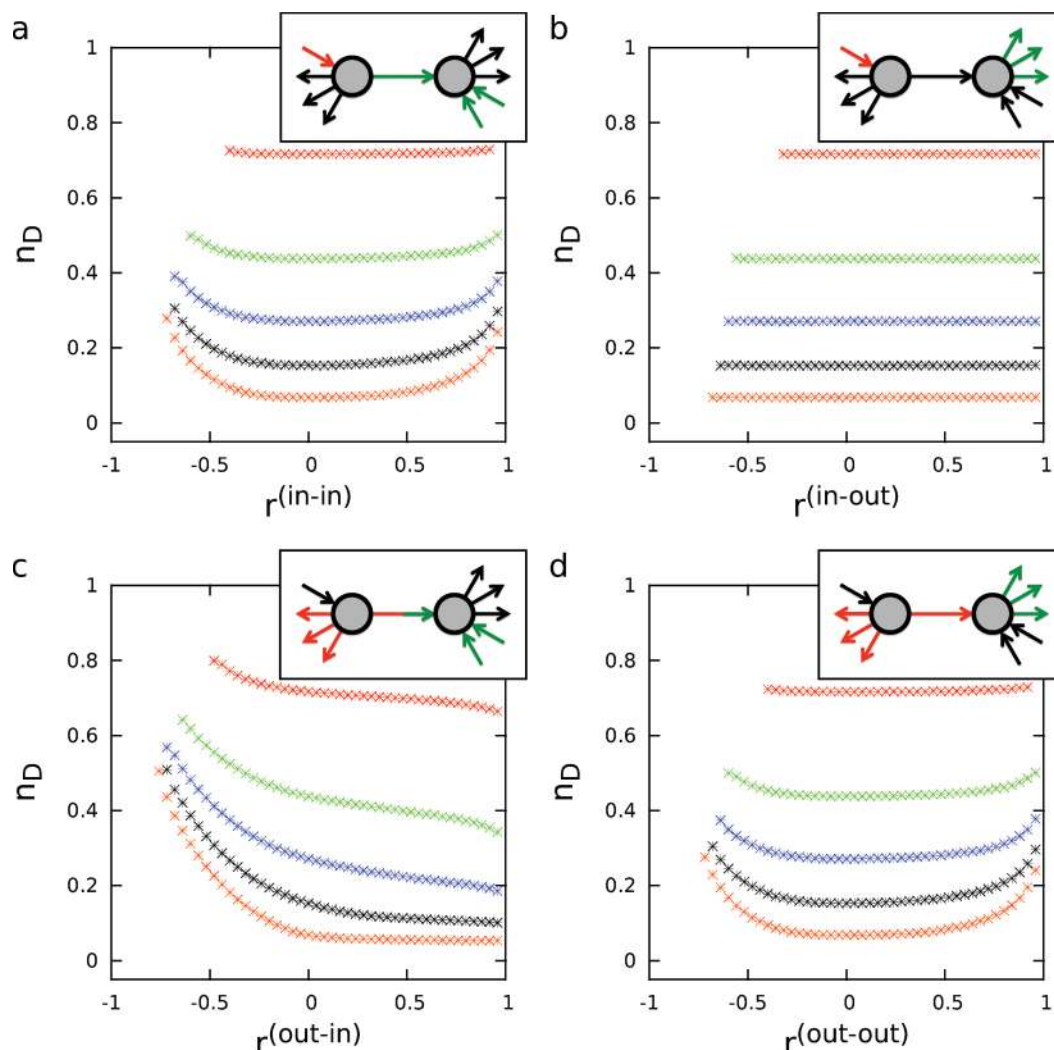
Figure 4 | The impact of degree-degree correlations on the density of driver nodes (n_D) for the scale-free model ($N = 10, 000$, $\gamma = 2.5$) for average degrees $\langle k \rangle = 1$ (red), $\langle k \rangle = 3$ (green), $\langle k \rangle = 5$ (blue), $\langle k \rangle = 7$ (black) and $\langle k \rangle = 9$ (orange). The results are similar for the Erdős-Rényi model (see Fig. 3). Each data point is an average of 100 independent runs.

$$e^{(\text{in-in})}(j_i, j_o; k_i, k_o) = Q^{(\text{out})}(j_o) P^{(\text{out})}(k_o)$$
$$\left[P^{(\text{in})}(j_i) Q^{(\text{in})}(k_i) + r^{(\text{in-in})} m^{(\text{in-in})}(j_i, k_i) \right], \quad (5a)$$

$$e^{(\text{in-out})}(j_i, j_o; k_i, k_o) = Q^{(\text{out})}(j_o) Q^{(\text{in})}(k_i)$$
$$\left[P^{(\text{in})}(j_i) P^{(\text{out})}(k_o) + r^{(\text{in-out})} m^{(\text{in-out})}(j_i, k_o) \right], \quad (5b)$$

$$e^{(\text{out-in})}(j_i, j_o; k_i, k_o) = P^{(\text{in})}(j_i) P^{(\text{out})}(k_o)$$
$$\left[Q^{(\text{out})}(j_o) Q^{(\text{in})}(k_i) + r^{(\text{out-in})} m^{(\text{out-in})}(j_o, k_i) \right], \quad (5c)$$

$$e^{(\text{out-out})}(j_i, j_o; k_i, k_o) = Q^{(\text{in})}(j_i) P^{(\text{in})}(k_i)$$
$$\left[P^{(\text{out})}(j_o) Q^{(\text{out})}(k_o) + r^{(\text{out-out})} m^{(\text{out-out})}(j_o, k_o) \right]. \quad (5d)$$

By fixing $m^{(\alpha-\beta)}(j, k)$ ($\alpha, \beta \in \{\text{in, out}\}$) we obtain a one parameter network ensemble characterized by $r^{(\alpha-\beta)}$, where $m^{(\alpha-\beta)}(j, k)$ satisfies the constraints

$$\sum_{j=0}^{\infty} m^{(\alpha-\beta)}(j, k) = \sum_{k=0}^{\infty} m^{(\alpha-\beta)}(j, k) = 0, \quad (6)$$

$$\sigma^{(\alpha)} \sigma^{(\beta)} \sum_{j,k=0}^{\infty} jk \cdot m^{(\alpha-\beta)}(j, k) = 1, \quad (7)$$

and all elements of $e^{(\alpha-\beta)}(j, k)$ are between 0 and 1.

Our goal is to understand the relation between n_D and the degree correlation coefficient $r^{(\alpha-\beta)}$. Assuming that $r^{(\alpha-\beta)}$ is small we treat the correlations as perturbations to the neutral case, discussing the impact of the four $r^{(\alpha-\beta)}$ correlations separately.

Out-in correlations. Using equation (5c) and keeping the first non-zero correction we obtain (Supplementary Information Sec. III.):

$$\overline{n_D}^{(\text{out-in})} = \overline{n_D}^{(0)} - r^{(\text{out-in})} \frac{\langle k \rangle}{4} [M_1(\hat{w}_2, 1 - w_1) + M_1(1 - \hat{w}_1, w_2)], (8)$$

where $\overline{n_D}^{(0)}$ is the fraction of driver nodes of the uncorrelated network; w_i and $\hat{w}_i$ only depend on $P(k_{\text{in}}, k_{\text{out}})$[10], and

$$M_1(x, y) = \sum_{j,k=1}^{\infty} m^{(\text{out-in})}(j, k) x^{j-1} y^{k-1}. \quad (9)$$

Equation (8) predicts that $\overline{n_D}$ depends linearly on $r^{(\text{out-in})}$, a prediction supported by simulations for small $r^{(\text{out-in})}$ (Figs. 3c and 4c). This behavior is also revealed by the equivalent problem of finding the
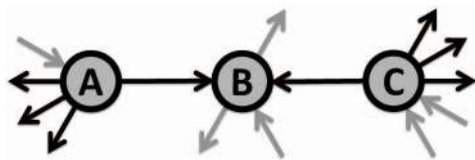
Figure 5 | One-step out-out correlations induce positive two-step correlation. Positive (negative) correlation between neighboring nodes means that if node A has high out-degree, then node B is likely to have high (low) out-degree, and hence C will likely have high out-degree.

maximum matching of graphs[10]. For a node A with out-degree k_0, by definition only one edge can be in the matching. If the remainder $k_0 - 1$ edges point to nodes with degree 1 (disassortative case), A inhibits them from being matched, so we have to control each of them individually, increasing n_D. If the remainder $k_0 - 1$ edges point to hubs (assortative case), these hubs are likely to be matched through another incoming edge, decreasing n_D.

Out-out correlations. The cavity method indicates that for out-out correlations the first nonzero correction is of order $(r^{(out-out)})^2$:

$$\overline{n_D}^{(out-out)} = \overline{n_D}^{(0)} + r^{(out-out)2} \frac{\langle k \rangle}{8} \left[H^{(in)\prime}(1-w_1)M_2(\hat{w}_2) + H^{(out)\prime}(w_2)M_2(1-\hat{w}_1) \right], \quad (10)$$

where $H^{(\alpha)}(x) = \sum_{k=1}^{\infty} Q^{(\alpha)}(k)x^{k-1}$ ($\alpha \in \{in, out\}$) only depends on $P(k_{in}, k_{out})$ and

$$M_2(x) = \sum_{j,k=1,l=0}^{\infty} \frac{m^{(out-out)}(l,j)m^{(out-out)}(l,k)}{P^{(out)}(l)}x^{j-1}x^{k-1}. \quad (11)$$

Equation (10) predicts that $\overline{n_D}^{(out-out)}$ does not depend on the out-out correlation of the directly connected nodes, but only on the correlation between the second neighbors, hence its dependence is quadratic in $r^{(out-out)}$, a prediction supported by numerical simulations (Figs. 3d and 4d). Indeed, positive (negative) $r^{(out-out)}$ correlation between the immediate neighbors means that if node A has high out-degree, then node B is expected to have high (low) out-degree, and therefore C is likely to have high out-degree (Fig. 5). That is, both positive and negative one-step out-out correlations induce positive

two-step correlations, accounting for the symmetry of the effect observed in simulations (Figs. 3d and 4d).

In-in correlations. Switching the direction of each link does not change the matching, but turns out-out correlations into in-in correlations. So $\overline{n_D}^{in-in}$ can be obtained by exchanging $P^{(in)}(k_{in})$ and $P^{(out)}(k_{out})$ in equation (10), predicting again a quadratic dependence on $r^{(in-in)}$, supported by the numerical simulations (Figs. 3a and 4a).

In-out correlations. The equations for $\overline{n_D}$ do not depend on the in-degree of the source and the out-degree of the target of a link, hence we predict that $r^{(in-out)}$ does not play a role in network controllability, a prediction supported by the simulations (see Figs. 3b and 4b).

Taken together, we predict that the functional dependence of $\overline{n_D}$ on degree correlations defines three classes of behaviors, depending on the matching problem's underlying symmetries: $\overline{n_D}$ has no dependence on $r^{(in-out)}$, linear dependence on $r^{(out-in)}$ and quadratic dependence on $r^{(in-in)}$ and $r^{(out-out)}$. These predictions are fully supported by numerical simulations (Figs. 3 and 4): for small r we see no dependence on $r^{(in-out)}$, an asymmetric, monotonic dependence on $r^{(out-in)}$, and a symmetric on $r^{(in-in)}$ and $r^{(out-out)}$.

To directly compare the analytical predictions to simulations we need to know the complete $e(j_i, j_o; k_i, k_o)$ distribution, which is not explicitly set in our simulations. So to test the results we use a rewiring method that sets the $e(j_i, j_o; k_i, k_o)$ distribution, not only the r correlation coefficient[21]. This method is not as robust as our original algorithm and the range of accessible r values is more restricted. However, since our results are based on perturbation scheme we only expect them to be correct for small r values. Indeed, we find that the predictions quantitatively reproduce the numerical results in a fair interval of $r^{(\alpha-\beta)}$ (Fig. 6).

Real networks. We test the predictions provided by the developed analytical and numerical tools on a set of publicly available network datasets. When complex systems are mapped to networks, the links connecting the nodes represent interactions between them. In this context self-loops represent self-interactions, with a strong, well understood impact on controllability[10,38]. While in some systems self-loops are obviously present (e.g. neural networks), in others they are manifestly absent (e.g. electric circuits[39]). Our purpose here is to test the effect of correlations, hence we rely on datasets that capture the wiring diagram of various complex systems with different correlation properties. Therefore, even if in a few of these
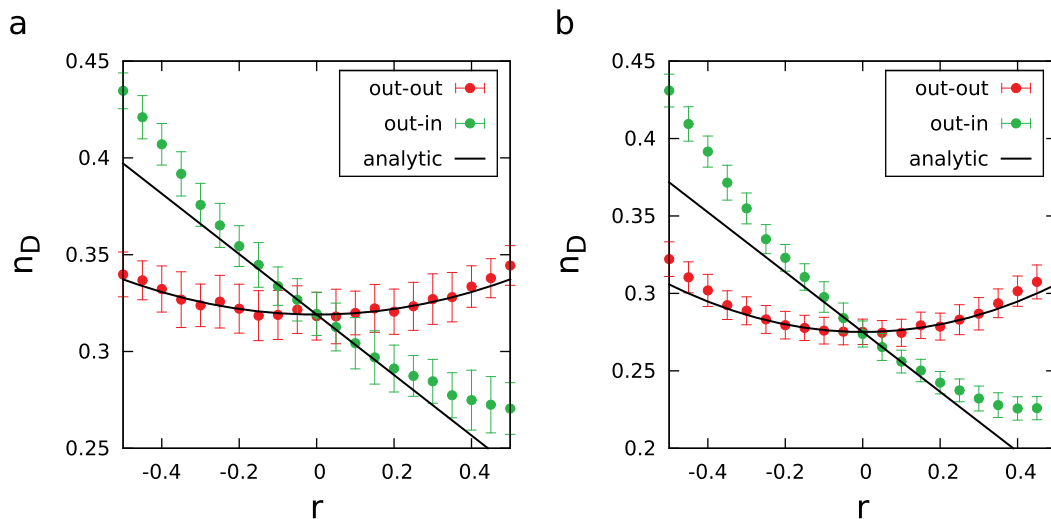

Figure 6 | The analytic formulas are tested with simulations on an (a) Erdős-Rényi model and on a (b) scale-free model. We used the algorithm proposed in[21] to set $e^{(\alpha-\beta)}(j_i, j_o; k_i, k_o)$. For (a) network we choose $N = 1,000$ and $\langle k \rangle = 3$; for (b) $N = 1,000$, $\gamma = 2.5$ and $\langle k \rangle = 4$. Each data point is an average over 100 independent runs; the errors represent by the standard deviation of the measurements.
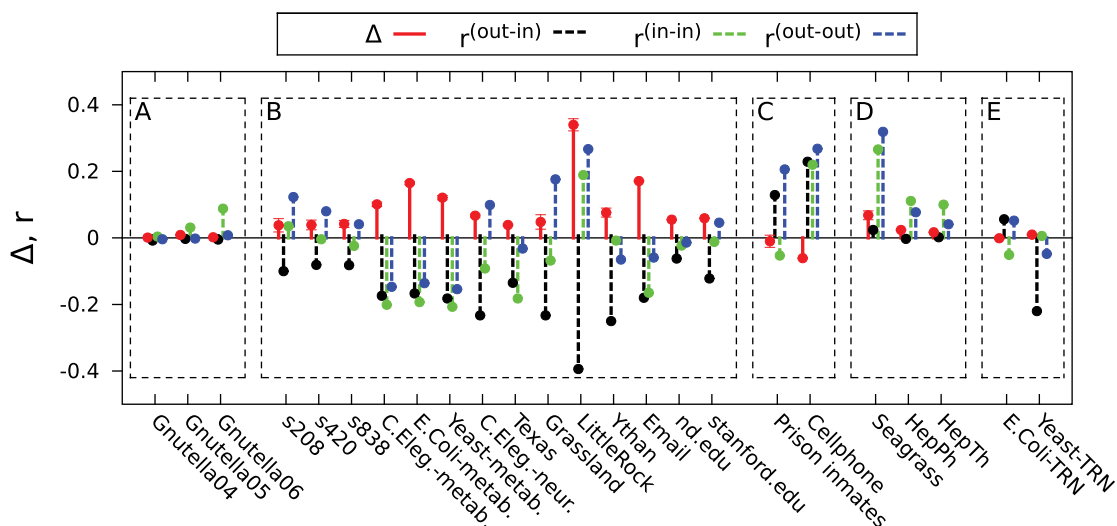
Figure 7 | The observed and predicted deviation between N_D and N_D^{rand}. Red line: $\Delta = (N_D - N_D^{rand})/N$, the prediction error based on the degree sequence. Dashed lines: correlations relevant to controllability. For each network Δ is calculated by averaging over 50 independent configurations.

maps self-loops are missing, it is beyond the scope of this work to complete these networks. However, when studying controllability of a particular system, careful thought has to be put into whether self-loops are present or not. We present a systematic study on the effect of self-loops in the Supplementary Information Sec. II.B.

To test the impact of our predictions on real networks we calculate

$$\Delta = \frac{N_D - N_D^{rand}}{N}, \qquad (12)$$

where N_D^{rand} represent the number of driver nodes for the degree-preserved randomized version of the original network. Hence if $\Delta = 0$ then $P(k_{in}, k_{out})$ accurately determines N_D; if $\Delta \neq 0$ then the structural properties not captured by the degree sequence influence its controllability. We measure the correlations in several real networks and based on our numerical and analytical results we predict the sign of Δ (Fig. 7). We grouped the networks according to our predictions. We provide the details of each network dataset in the Supplementary Information Table SI.

Group A. The networks of p2p Internet (Gnutella filesharing clients) do not have strong correlations, therefore we expect n_D to be correctly approximated by the prediction based on $P(k_{in}, k_{out})$ (i.e. $\Delta \approx 0$), in line with the empirical observations.

Group B. As in most networks the three relevant correlations coexist to some degree (Fig. 7), it is impossible to isolate their individual role. Yet, the networks in this group (electric circuits, metabolic networks, neural networks, power grids and food webs with exception of the Seagrass network) all have negative out-in and nonzero in-in and out-out correlations, each of which individually increase n_D as we showed above. Therefore we predict $\Delta > 0$, in line with the empirical observations.

Group C. Only the prison social-trust and the cell phone network feature significant positive out-in correlations. These networks also display nonzero in-in and out-out correlation, leading to the coexistence of two competing effects: out-in correlations decrease n_D and the out-out and in-in correlations increase n_D. Since the out-in correlation is a first order effect (equation (8)), while out-out and in-in correlations are only of second order (equation (10)), we expect a decrease in n_D (i.e. $\Delta < 0$), consistent with the empirical results.

Group D. The Seagrass food web and citation networks do not feature significant out-in correlations, only the secondary in-in and out-out

correlations, hence we expect n_D to increase ($\Delta > 0$), consistent with the observations.

Group E. Only the transcriptional regulatory networks are somewhat puzzling in that they show degree correlations, yet the degree sequence still correctly gives n_D. However, the simulations indicated that the effect of correlations is negligible for high n_D. And our analytical results showed that the value of the correction depends on details of $e(j_i, j_o; k_i, k_o)$, not captured by the Pearson coefficient r. These observations highlight that even though in most cases our qualitative predictions based on r are valid, in some cases further investigation is required.

Discussion

The goal of our paper was to clarify the higher order network characteristics that influence controllability. We studied the effect of three topological characteristics: clustering, modularity and degree correlations. We used numerical simulations to identify the role of the relevant characteristics, finding that changes in the clustering coefficient and the community structure have no systematic effect on the the minimum number of driver nodes n_D. In contrast degree correlations showed a robust effect, whose magnitude and direction depends on the type of correlation. Using the cavity method we derived n_D for networks with given degree distribution and correlation profiles, finding results that are consistent with our numerical simulations. For real networks these numerical and analytic results enabled us to qualitatively explain the deviation of the observed n_D from the prediction based only on $P(k_{in}, k_{out})$.

Our results not only offer a new perspective on the role of topological properties on network controllability, but also raise several questions. Future research directions include determining the optimal network structure to minimize the number of necessary driver nodes, and studying how different network characteristics influence the robustness of the control configuration.

Methods

Generating a scale-free network. We use the static model to generate directed scale-free networks[40]. We start from N disconnected nodes and assign a weight $w_i = (i + i_0)^{-\alpha}$ to each node i ($i = 1 \ldots N$). We randomly select two nodes i and j with probability proportional to w_i and w_j respectively and if they are yet not connected, we connect them. We allow self-loops, but avoid multi-edges. We repeat the process until L links have been placed. The resulting network has average degree $\langle k \rangle = 2L/N$, and $P^{(in/out)}(k) \sim k^{-\gamma}$ for large k, where $\gamma = 1 + \frac{1}{\alpha}$, and maximum degree $k_{max} \sim i_0^{-\alpha}$.

To systematically study correlations, the starting network has to be uncorrelated. However, the presence of hubs may induce unwanted degree correlations[41], and may

also considerably limit the maximum and minimum correlations accessible via rewiring[42]. We overcome these difficulties by introducing a structural cutoff in the degrees, choosing i_0 to ensure $k_{max} < (\langle k \rangle N)^{1/2}$ [43]. Note, that in the static model of Goh et al. $i_0 = 0$ [40].

As both in- and out-degree of node i is proportional to w_i, the above procedure results in correlations between the in- and out-degrees of node i. To eliminate the correlations, we randomize the in-degree sequence while keeping the out-degree sequence unchanged.

Rewiring algorithm. We use degree preserving rewiring[20] to add each network characteristic. Suppose that the chosen network characteristic is quantified by a metric X. To set its value to X^*, we define the $E(X) = |X - X^*|$ energy, so $E(X^*)$ is a global minimum. We minimize this energy by simulated annealing[44]: (1) choose two links at random with uniform probability; (2) rewire the two links and calculate the energy $E(X)$ of the resulted network; (3) accept the new configuration with probability

$$p = \begin{cases} 1, & \text{if } \Delta E \le 0 \\ e^{-\beta \Delta E}, & \text{if } \Delta E > 0, \end{cases} \quad (13)$$

where the β parameter is the inverse temperature; (4) repeat from step one and gradually increase β. Stop if $|E(X) - E(X^*)|$ is smaller than a predefined value.

Note, that keeping the degree sequence bounds the possible values of X that can be reached by rewiring. In all cases we study the full interval of accessible X values.

1. Ben-Naim, E., Frauenfelder, H. & Toroczkai, Z. (Eds.) *Complex Networks* (Springer, Berlin, 2004).
2. Newman, M., Barabási, A.-L. & Watts, D. J. *The Structure and Dynamics of Networks* (Princeton University Press, Princeton, 2006).
3. Caldarelli, G. *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, Oxford, 2007).
4. Barrat, A., Barthélemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2009).
5. Cohen, R. & Havlin, S. *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, Cambridge, 2010).
6. Chen, G., Wang, X. & Li, X. *Introduction to Complex Networks: Models, Structures and Dynamics* (Higher Education Press, Beijing, 2012).
7. Wang, X. F. & Chen, G. Pinning control of scale-free dynamical networks. *Physica A* **310**, 521–531 (2002).
8. Sorrentino, F., di Bernardo, M., Garofalo, F. & Chen, G. Controllability of complex networks via pinning. *Phys. Rev. E* **75**, 1–6 (2007).
9. Gutiérrez, R., Sendiña-Nadal, I., Zanin, M., Papo, D. & Boccaletti, S. Targeting the dynamics of complex networks. *Sci. Rep.* **2** (2007).
10. Liu, Y.-Y., Slotine, J.-J. & Barabasi, A.-L. Controllability of complex networks. *Nature* **473**, 167–173 (2011).
11. 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).
12. 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).
13. Nepusz, T. & Vicsek, T. Controlling edge dynamics in complex networks. *Nat. Phys.* **8**, 568–573 (2012). 10.1038/nphys2327.
14. Slotine, J.-J. & Li, W. *Applied Nonlinear Control* (Prentice-Hall, 1991).
15. Luenberger, D. G. *Introduction to Dynamic Systems: Theory, Models, & Applications* (John Wiley & Sons, New York, 1979).
16. Kalman, R. E. Mathematical description of linear dynamical systems. *J. Soc. Indus. and Appl. Math. Ser. A* **1**, 152 (1963).
17. Chui, C. K. & Chen, G. *Linear Systems and Optimal Control* (Springer-Verlag, New York, 1989).
18. Lovász, L. & Plummer, M. D. *Matching Theory* (American Mathematical Society, Rhode Island, 2009).
19. Pastor-Satorras, R., Vazquez, A. & Vespignani, A. Dynamical and correlation properties of the internet. *Phys. Rev. Lett.* **87**, 258701 (2001).
20. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).
21. Newman, M. E. J. Mixing patterns in networks. *Phys. Rev. E* **67**, 026126 (2003).
22. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
23. Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barabási, A.-L. Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551–1555 (2002).
24. Palla, G., Derényi, I., Farkas, I. & Vicsek, T. Uncovering the overlapping community structure of complex networks in nature and society. *Nature* **435**, 814–818 (2005).
25. Leicht, E. A. & Newman, M. E. J. Community structure in directed networks. *Phys. Rev. Lett.* **100**, 118703 (2008).
26. Fortunato, S. Community detection in graphs. *Physics Reports* **486**, 75–174 (2010).
27. Barrat, A. & Weigt, M. On the properties of small-world network models. *Eur. Phys. J. B* **13**, 547–560 (2000).
28. Foster, J. G., Foster, D. V., Grassberger, P. & Paczuski, M. Edge direction and the structure of networks. *Proc. Natl. Acad. Sci.* **107**, 10815 (2010).
29. Erdös, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–60 (1960).
30. Bollobás, B. *Random Graphs* (Cambridge University Press, Cambridge, 2001).
31. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509 (1999).
32. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
33. Hopcroft, J. E. & Karp, R. M. An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. *SIAM J. Comput.* **2**, 225 (1973).
34. Ravasz, E. & Barabási, A.-L. Hierarchical organization in complex networks. *Phys. Rev. E* **67**, 026112 (2003).
35. Mones, E., Vicsek, L. & Vicsek, T. Hierarchy measure for complex networks. *PLoS ONE* **7**(3): e33799 (2012).
36. Ahn, Y.-Y., Bagrow, J. P. & Lehmann, S. Link communities reveal multiscale complexity in networks. *Nature* **466**, 761–764 (2010).
37. Zdeborová, L. & Mézard, M. The number of matchings in random graphs. *J. Stat. Mech.* **05**, P05003 (2006).
38. Cowan, N. J., Chastain, E., Vilhena, D. A. & Bergstrom, C. T. Controllability of Real Networks. *arXiv:1106.2573v3* (2011).
39. Lin, C.-T. Structural Controllability. *IEEE Trans. Auto. Contr.* **19**, 201 (1974).
40. Goh, K.-I., Kahng, B. & Kim, D. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.* **87**, 278701 (2001).
41. Boguñá, M., Pastor-Satorras, R. & Vespignani, A. Cut-offs and finite size effects in scale-free networks. *The European Physical Journal B - Condensed Matter and Complex Systems* **38**, 205–209 (2004).
42. Menche, J., Valleriani, A. & Lipowsky, R. Asymptotic properties of degree-correlated scale-free networks. *Phys. Rev. E* **81**, 046103 (2010).
43. Chung, F. & Lu, L. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics* **6**, 125–145 (2002).
44. Press, W., Flannery, B., Teukolsky, S. & Vetterling, W. *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, 1992).

Acknowledgements

Author contributions

All authors designed and did the research. M.P. analysed the empirical data and did the analytical and numerical calculations. A.-L.B. was the lead writer.

Additional information