ARTICLE

doi:10.1038/nature10011

Controllability of complex networks

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

The ultimate proof of our understanding of natural or technological systems is reflected in our ability to control them. Although control theory offers mathematical tools for steering engineered and natural systems towards a desired state, a framework to control complex self-organized systems is lacking. Here we develop analytical tools to study the controllability of an arbitrary complex directed network, identifying the set of driver nodes with time-dependent control that can guide the system's entire dynamics. We apply these tools to several real networks, finding that the number of driver nodes is determined mainly by the network's degree distribution. We show that sparse inhomogeneous networks, which emerge in many real complex systems, are the most difficult to control, but that dense and homogeneous networks can be controlled using a few driver nodes. Counterintuitively, we find that in both model and real systems the driver nodes tend to avoid the high-degree nodes.

According to control theory, a dynamical system is controllable if, with a suitable choice of inputs, it can be driven from any initial state to any desired final state within finite time[1–3]. This definition agrees with our intuitive notion of control, capturing an ability to guide a system's behaviour towards a desired state through the appropriate manipulation of a few input variables, like a driver prompting a car to move with the desired speed and in the desired direction by manipulating the pedals and the steering wheel. Although control theory is a mathematically highly developed branch of engineering with applications to electric circuits, manufacturing processes, communication systems[4–6], aircraft, spacecraft and robots[2,3], fundamental questions pertaining to the controllability of complex systems emerging in nature and engineering have resisted advances. The difficulty is rooted in the fact that two independent factors contribute to controllability, each with its own layer of unknown: (1) the system's architecture, represented by the network encapsulating which components interact with each other; and (2) the dynamical rules that capture the time-dependent interactions between the components. Thus, progress has been possible only in systems where both layers are well mapped, such as the control of synchronized networks[7–10], small biological circuits[11] and rate control for communication networks[4–6]. Recent advances towards quantifying the topological characteristics of complex networks[12–16] have shed light on factor (1), prompting us to wonder whether some networks are easier to control than others and how network topology affects a system's controllability. Despite some pioneering conceptual work[17–23] (Supplementary Information, section II), we continue to lack general answers to these questions for large weighted and directed networks, which most commonly emerge in complex systems.

Network controllability

Most real systems are driven by nonlinear processes, but the controllability of nonlinear systems is in many aspects structurally similar to that of linear systems[3], prompting us to start our study using the canonical linear, time-invariant dynamics

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) + B\mathbf{u}(t) \quad (1)$$

where the vector $\mathbf{x}(t) = (x_1(t), \ldots, x_N(t))^T$ captures the state of a system of N nodes at time t. For example, $x_i(t)$ can denote the amount of traffic that passes through a node i in a communication network[24] or transcription factor concentration in a gene regulatory network[25]. The $N \times N$ matrix A describes the system's wiring diagram and the interaction strength between the components, for example the traffic on individual communication links or the strength of a regulatory interaction. Finally, B is the $N \times M$ input matrix ($M \leq N$) that identifies the nodes controlled by an outside controller. The system is controlled using the time-dependent input vector $\mathbf{u}(t) = (u_1(t), \ldots, u_M(t))^T$ imposed by the controller (Fig. 1a), where in general the same signal $u_i(t)$ can drive multiple nodes. If we wish to control a system, we first need to identify the set of nodes that, if driven by different signals, can offer full control over the network. We will call these 'driver nodes'. We are particularly interested in identifying the minimum number of driver nodes, denoted by N_D, whose control is sufficient to fully control the system's dynamics.

The system described by equation (1) is said to be controllable if it can be driven from any initial state to any desired final state in finite time, which is possible if and only if the $N \times NM$ controllability matrix

$$C = (B, AB, A^2B, \ldots, A^{N-1}B) \quad (2)$$

has full rank, that is

$$\mathrm{rank}(C) = N \quad (3)$$

This represents the mathematical condition for controllability, and is called Kalman's controllability rank condition[1,2] (Fig. 1a). In practical terms, controllability can be also posed as follows. Identify the minimum number of driver nodes such that equation (3) is satisfied. For example, equation (3) predicts that controlling node x_1 in Fig. 1b with the input signal u_1 offers full control over the system, as the states of nodes x_1, x_2, x_3 and x_4 are uniquely determined by the signal $u_1(t)$ (Fig. 1c). In contrast, controlling the top node in Fig. 1e is not sufficient for full control, as the difference $a_{31}x_2(t) - a_{21}x_3(t)$ (where a_{ij} are the elements of A) is not uniquely determined by $u_1(t)$ (see Fig. 1f and Supplementary Information section III.A). To gain full control, we must simultaneously control node x_1 and any two nodes among $\{x_2, x_3, x_4\}$ (see Fig. 1h, i for a more complex example).

To apply equations (2) and (3) to an arbitrary network, we need to know the weight of each link (that is, the a_{ij}), which for most real

[1]Center for Complex Network Research and Departments of Physics, Computer Science and Biology, Northeastern University, Boston, Massachusetts 02115, USA. [2]Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. [3]Nonlinear Systems Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [4]Department of Mechanical Engineering and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [5]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA.

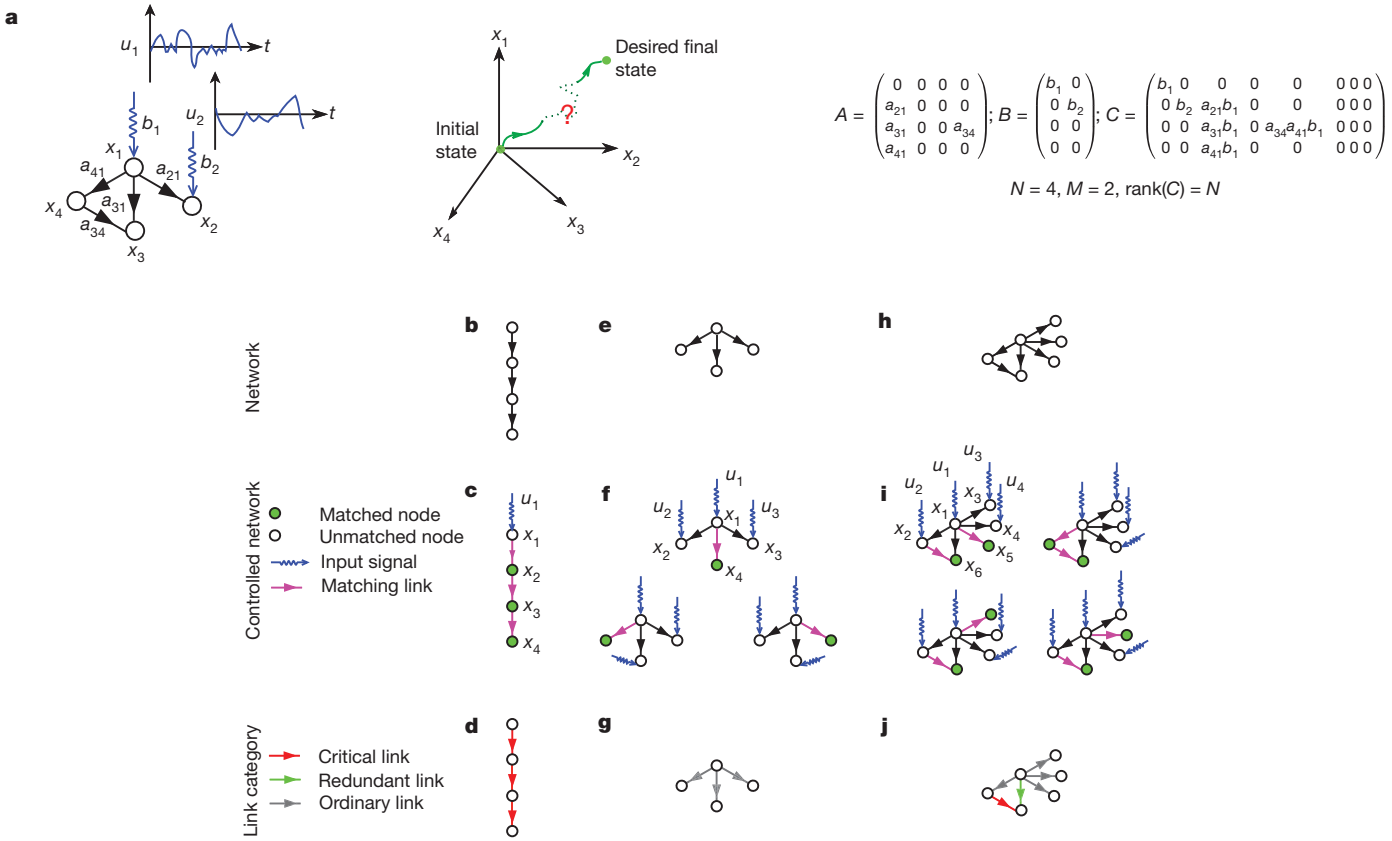

Figure 1 | Controlling a simple network. **a**, The small network can be controlled by an input vector $\mathbf{u} = (u_1(t), u_2(t))^T$ (left), allowing us to move it from its initial state to some desired final state in the state space (right). Equation (2) provides the controllability matrix (C), which in this case has full rank, indicating that the system is controllable. **b**, Simple model network: a directed path. **c**, Maximum matching of the directed path. Matching edges are shown in purple, matched nodes are green and unmatched nodes are white. The unique maximum matching includes all links, as none of them share a common starting or ending node. Only the top node is unmatched, so controlling it yields full control of the directed path ($N_D = 1$). **d**, In the directed path shown in **b**, all links are critical, that is, their removal eliminates our ability to control the network. **e**, Small model network: the directed star. **f**, Maximum matchings of the directed star. Only one link can be part of the maximum matching, which yields three unmatched nodes ($N_D = 3$). The three different maximum matchings indicate that three distinct node configurations can exert full control. **g**, In a directed star, all links are ordinary, that is, their removal can eliminate some control configurations but the network could be controlled in their absence with the same number of driver nodes N_D. **h**, Small example network. **i**, Only two links can be part of a maximum matching for the network in **h**, yielding four unmatched nodes ($N_D = 4$). There are all together four different maximum matchings for this network. **j**, The network has one critical link, one redundant link (which can be removed without affecting any control configuration) and four ordinary links.

networks are either unknown (for example regulatory networks) or are known only approximately and are time dependent (for example Internet traffic). Even if all weights are known, a brute-force search requires us to compute the rank of C for $2^N - 1$ distinct combinations, which is a computationally prohibitive task for large networks. To bypass the need to measure the link weights, we note that the system (A, B) is 'structurally controllable'[26] if it is possible to choose the non-zero weights in A and B such that the system satisfies equation (3). A structurally controllable system can be shown to be controllable for almost all weight combinations, except for some pathological cases with zero measure that occur when the system parameters satisfy certain accidental constraints[26,27]. Thus, structural controllability helps us to overcome our inherently incomplete knowledge of the link weights in A. Furthermore, because structural controllability implies controllability of a continuum of linearized systems[28], our results can also provide a sufficient condition for controllability for most nonlinear systems[3] (Supplementary Information, section III.A).

To avoid the brute-force search for driver nodes, we proved that the minimum number of inputs or driver nodes needed to maintain full control of the network is determined by the 'maximum matching' in the network, that is, the maximum set of links that do not share start or end nodes (Fig. 1c, f, i). A node is said to be matched if a link in the maximum matching points at it; otherwise it is unmatched. As we show in the Supplementary Information, the structural controllability problem maps into an equivalent geometrical problem on a network: we can gain full control over a directed network if and only if we directly control each unmatched node and there are directed paths from the input signals to all matched nodes[29]. The possibility of determining N_D, using this mapping, is our first main result. As the maximum matching in directed networks can be identified numerically in at most $O(N^{1/2}L)$ steps[30], where L denotes the number of links, the mapping offers an efficient method to determine the driver nodes for an arbitrary directed network.

Controllability of real networks

We used the tools developed above to explore the controllability of several real networks. The networks were chosen for their diversity: for example, the purpose of the gene regulatory network is to control the dynamics of cellular processes, so it is expected to evolve towards a structure that is efficient from a control perspective, potentially implying a small number of driver nodes (that is, small $n_D \equiv N_D/N$). In contrast, for the World Wide Web or citation networks controllability has no known role, making it difficult even to guess n_D. Finally, it might be argued that social networks, given their perceived neutrality (or even resistance) to control, should have a high n_D, as it is necessary to control most individuals separately to control the whole system. We used the mapping into maximum matching to determine the minimum set of driver nodes (N_D) for the networks in Table 1, the

obtained trend defying our expectations: as a group, gene regulatory networks display high n_D (~0.8), indicating that it is necessary to independently control about 80% of nodes to control them fully. In contrast, several social networks are characterized by some of the smallest n_D values, suggesting that a few individuals could in principle control the whole system.

Given the important role hubs (nodes with high degree) have in maintaining the structural integrity of networks against failures and attacks[31,32], in spreading phenomena[32,33] and in synchronization[8,34], it is natural to expect that control of the hubs is essential to control a network. To test the validity of this hypothesis, we divided the nodes into three groups of equal size according to their degree, k (low, medium and high). As Fig. 2a, b shows for two canonical network models (Erdős–Rényi[35,36] and scale-free[15,37–39]), the fraction of driver nodes is significantly higher among low-k nodes than among the hubs. In Fig. 2c, we plot the mean degree of the driver nodes, $\langle k_D \rangle$, as a function of the mean degree, $\langle k \rangle$, of each network in Table 1 and several network models. In all cases, $\langle k_D \rangle$ is either significantly smaller than or comparable to $\langle k \rangle$, indicating that in both real and model systems the driver nodes tend to avoid the hubs.

To identify the topological features that determine network controllability, we randomized each real network using a full randomization procedure (rand-ER) that turns the network into a directed Erdős–Rényi random network with N and L unchanged. For several networks there is no correlation between the N_D of the original network and the N_D of its randomized counterpart (Fig. 2d), indicating that full randomization eliminates the topological characteristics that influence controllability. We also applied a degree-preserving randomization[40,41] (rand-Degree), which keeps the in-degree, k_{in}, and out-degree, k_{out}, of each node unchanged but selects randomly the nodes that link to each other. We find that this procedure does not alter N_D significantly, despite the observed differences in N_D of six orders of magnitude (Fig. 2e). Thus, a system's controllability is to a great extent encoded by the underlying network's degree distribution, $P(k_{in}, k_{out})$, which is our second and most important finding. It indicates that N_D is determined mainly by the number of incoming and outgoing links each node has and is independent of where those links point.

An analytical approach to controllability

The importance of the degree distribution allows us to determine N_D analytically for a network with an arbitrary $P(k_{in}, k_{out})$. Using the cavity method[42–44], we derived a set of self-consistent equations (Supplementary Information, section IV) whose input is the degree distribution and whose solution is the average n_D (or N_D) over all network realizations compatible with $P(k_{in}, k_{out})$, which is our third key result. As Fig. 2f shows, the analytically predicted N_D agrees perfectly with $N_D^{\text{rand-Degree}}$ (and hence is in good agreement with the exact value, N_D^{real}), offering an effective analytical tool to study

Table 1 | The characteristics of the real networks analysed in the paper

Type	Name	N	L	n_D^{real}	$n_D^{\text{rand-Degree}}$	$n_D^{\text{rand-ER}}$
Regulatory	TRN-Yeast-1	4,441	12,873	0.965	0.965	0.083
	TRN-Yeast-2	688	1,079	0.821	0.811	0.303
	TRN-EC-1	1,550	3,340	0.891	0.891	0.188
	TRN-EC-2	418	519	0.751	0.752	0.380
	Ownership-USCorp	7,253	6,726	0.820	0.815	0.480
Trust	College student	32	96	0.188	0.173	0.082
	Prison inmate	67	182	0.134	0.144	0.103
	Slashdot	82,168	948,464	0.045	0.278	1.7×10^{-5}
	WikiVote	7,115	103,689	0.666	0.666	1.4×10^{-4}
	Epinions	75,888	508,837	0.549	0.606	0.001
Food web	Ythan	135	601	0.511	0.433	0.016
	Little Rock	183	2,494	0.541	0.200	0.005
	Grassland	88	137	0.523	0.477	0.301
	Seagrass	49	226	0.265	0.199	0.203
Power grid	Texas	4,889	5,855	0.325	0.287	0.396
Metabolic	*Escherichia coli*	2,275	5,763	0.382	0.218	0.129
	Saccharomyces cerevisiae	1,511	3,833	0.329	0.207	0.130
	Caenorhabditis elegans	1,173	2,864	0.302	0.201	0.144
Electronic circuits	s838	512	819	0.232	0.194	0.293
	s420	252	399	0.234	0.195	0.298
	s208	122	189	0.238	0.199	0.301
Neuronal	*Caenorhabditis elegans*	297	2,345	0.165	0.098	0.003
Citation	ArXiv-HepTh	27,770	352,807	0.216	0.199	3.6×10^{-5}
	ArXiv-HepPh	34,546	421,578	0.232	0.208	3.0×10^{-5}
World Wide Web	nd.edu	325,729	1,497,134	0.677	0.622	0.012
	stanford.edu	281,903	2,312,497	0.317	0.258	3.0×10^{-4}
	Political blogs	1,224	19,025	0.356	0.285	8.0×10^{-4}
Internet	p2p-1	10,876	39,994	0.552	0.551	0.001
	p2p-2	8,846	31,839	0.578	0.569	0.002
	p2p-3	8,717	31,525	0.577	0.574	0.002
Social communication	UClonline	1,899	20,296	0.323	0.322	0.706
	Email-epoch	3,188	39,256	0.426	0.332	3.0×10^{-4}
	Cellphone	36,595	91,826	0.204	0.212	0.133
Intra-organizational	Freemans-2	34	830	0.029	0.029	0.029
	Freemans-1	34	695	0.029	0.029	0.029
	Manufacturing	77	2,228	0.013	0.013	0.013
	Consulting	46	879	0.043	0.043	0.022

For each network, we show its type and name; number of nodes (N) and edges (L); and density of driver nodes calculated in the real network (n_D^{real}), after degree-preserved randomization ($n_D^{\text{rand-Degree}}$) and after full randomization ($n_D^{\text{rand-ER}}$). For data sources and references, see Supplementary Information, section VI.

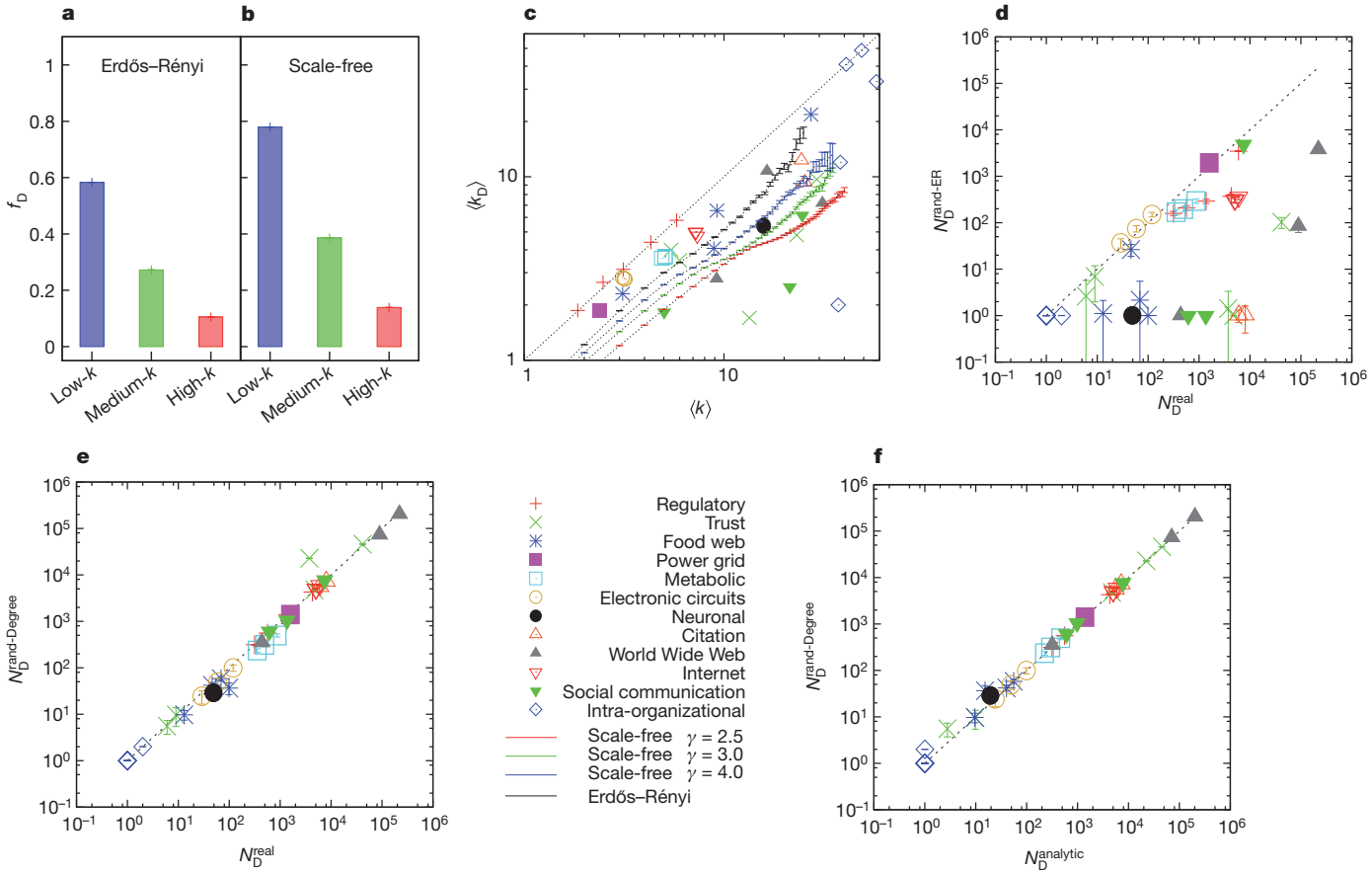

Figure 2 | Characterizing and predicting the driver nodes (N_D). **a**, **b**, Role of the hubs in model networks. The bars show the fractions of driver nodes, f_D, among the low-, medium- and high-degree nodes in two network models, Erdős–Rényi (**a**) and scale-free (**b**), with $N = 10^4$ and $\langle k \rangle = 3$ ($\gamma = 3$), indicating that the driver nodes tend to avoid the hubs. Both the Erdős–Rényi and the scale-free networks are generated from the static model[38] and the results are averaged over 100 realizations. The error bars (s.e.m.), shown in the figure, are smaller than the symbols. **c**, Mean degree of driver nodes compared with the mean degree of all nodes in real and model networks, indicating that in real systems the hubs are avoided by the driver nodes. **d**, Number of driver nodes, $N_D^{rand\text{-}ER}$, obtained for the fully randomized version of the networks listed in Table 1, compared with the exact value, N_D^{real}. **e**, Number of driver nodes, $N_D^{rand\text{-}Degree}$, obtained for the degree-preserving randomized version of the networks shown in Table 1, compared with N_D^{real}. **f**, The analytically predicated $N_D^{analytic}$ calculated using the cavity method, compared with $N_D^{rand\text{-}Degree}$. In **d**–**f**, data points and error bars (s.e.m.) were determined from 1,000 realizations of the randomized networks.

the impact of various network parameters on N_D. Although the cavity method does not offer a closed-form solution, we can derive the dependence of n_D on key network parameters in the thermodynamic limit ($N \to \infty$). We find, for example, that for a directed Erdős–Rényi network n_D decays as

$$n_D \approx e^{-\langle k \rangle/2} \tag{4}$$

for large $\langle k \rangle$. For scale-free networks with degree exponent $\gamma_{in} = \gamma_{out} = \gamma$ in the large-$\langle k \rangle$ limit[38], we have

$$n_D \approx \exp\left[-\frac{1}{2}\left(1-\frac{1}{\gamma-1}\right)\langle k \rangle\right] \tag{5}$$

which has the same $\langle k \rangle$ dependence as equation (4) in the $\gamma \to \infty$ limit. Equation (5) predicts that $\gamma_c = 2$ is a critical exponent for the controllability of an infinite scale-free network, as only for $\gamma > \gamma_c$ can we control the full system through a finite subset of nodes (that is, $n_D < 1$). For $\gamma \leq \gamma_c$ in the thermodynamic limit, all nodes must be individually controlled (that is, $n_D = 1$). We note that γ_c is different from $\gamma = 3$, which is the critical exponent for a number of network phenomena driven by the divergence of $\langle k^2 \rangle$, from network robustness to epidemic spreading[31–33,45]. To check the validity of the analytical predictions, we determined the $\langle k \rangle$ dependence of n_D numerically for both Erdős–Rényi and scale-free networks, confirming the asymptotic exponential dependence of n_D on $\langle k \rangle$, as predicted by equations (4) and (5). Furthermore, the predicted n_D value is in excellent agreement with the numerical results for $\gamma > 3$ (Fig. 3d, e). Near $\gamma = 2$, however, n_D as predicted by the cavity method deviates from the exact n_D value owing to degree correlations that are prominent at $\gamma_c = 2$ and can be eliminated by imposing a degree cut-off in constructing the scale-free networks[39,46] (Supplementary Information, section IV.B).

Equation (5) also shows that n_D decreases as γ increases (for fixed $\langle k \rangle$), indicating that n_D is affected by degree heterogeneity, representing the spread between the less connected and the more connected nodes. We defined the degree heterogeneity as $H = \Delta/\langle k \rangle$, where $\Delta = \sum_i \sum_j |k_i - k_j| P(k_i) P(k_j)$ is the average absolute degree difference of all pairs of nodes (i and j) drawn from the degree distribution $P(k)$. The degree heterogeneity is zero ($H = 0$) for networks in which all nodes have the same degree, such as the random regular digraph (Fig. 3a), in which the in- and out-degrees of the nodes are fixed to $\langle k \rangle/2$ but the nodes are connected randomly. For $\langle k \rangle \geq 2$, this graph always has a perfect matching[47], which means that a single driver node can control the whole system (Supplementary Information, section IV.B1). The degree heterogeneity increases as we move from the random regular digraph to an Erdős–Rényi network (Fig. 3b) and eventually to scale-free networks with decreasing γ (Fig. 3c). Overall, the fraction of driver nodes, n_D, increases monotonically with H, whether we keep γ (Fig. 3f) or $\langle k \rangle$ (Fig. 3g) constant.

Taking these results together, we find that the denser a network is, the fewer driver nodes are needed to control it, and that small changes in the average degree induce orders-of-magnitude variations in n_D.

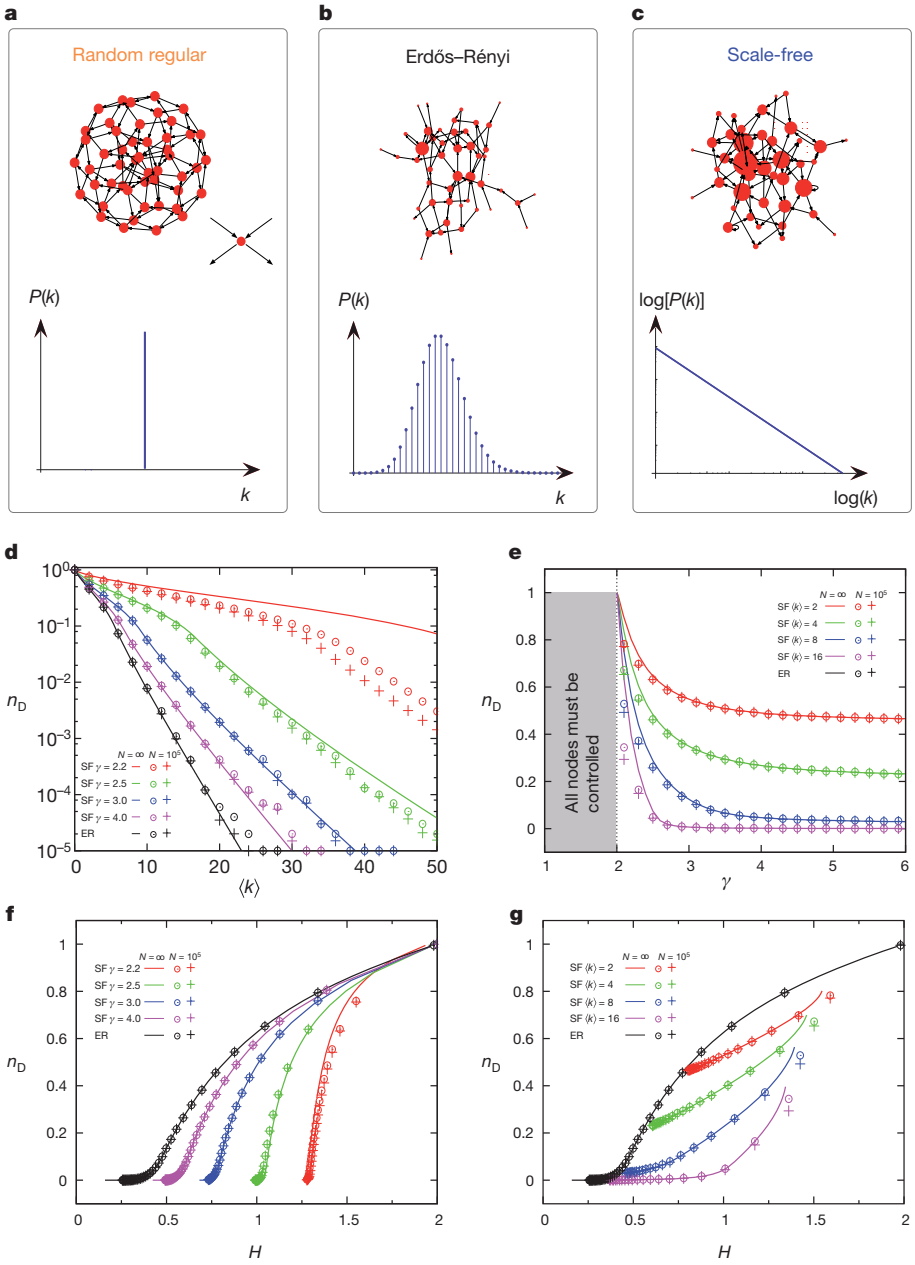

Figure 3 | The impact of network structure on the number of driver nodes. **a**–**c**, Characteristics of the explored model networks. A random-regular digraph (**a**), shown here for $\langle k\rangle = 4$, is the most degree-homogeneous network as $k_{in} = k_{out} = \langle k\rangle/2$ for all nodes. Erdős–Rényi networks (**b**) have Poisson degree distributions and their degree heterogeneities are determined by $\langle k\rangle$. Scale-free networks (**c**) have power-law degree distributions, yielding large degree heterogeneities. **d**, Driver node density, n_D, as a function of $\langle k\rangle$ for Erdős–Rényi (ER) and scale-free (SF) networks with different values of γ. Both the Erdős–Rényi and the scale-free networks are generated from the static model[38] with $N = 10^5$. Lines are analytical results calculated by the cavity method using the expected degree distribution in the $N \to \infty$ limit. Symbols are calculated for the constructed discrete network: open circles indicate exact results calculated from the maximum matching algorithm, and plus symbols indicate the analytical results of the cavity method using the exact degree sequence of the constructed network. For large $\langle k\rangle$, n_D approaches its lower bound, N^{-1}, that is, a single driver node ($N_D = 1$) in a network of size N. **e**, n_D as a function of γ for scale-free networks with fixed $\langle k\rangle$. For infinite scale-free networks, $n_D \to 1$ as $\gamma \to \gamma_c = 2$, that is, it is necessary to control almost all nodes to control the network fully. For finite scale-free networks, n_D reaches its maximum as γ approaches γ_c (Supplementary Information). **f**, n_D as a function of degree heterogeneity, H, for Erdős–Rényi and scale-free networks with fixed γ and variable $\langle k\rangle$. **g**, n_D as a function of H for Erdős–Rényi and scale-free networks for fixed $\langle k\rangle$ and variable γ. As γ increases, the curves converge to the Erdős–Rényi result (black) at the corresponding $\langle k\rangle$ value.

Furthermore, the larger are the differences between node degrees, the more driver nodes are needed to control the system. Overall, networks that are sparse and heterogeneous, which are precisely the characteristics often seen in complex systems like the cell or the Internet[13–16], require the most driver nodes, underscoring that such systems are difficult to control.

Robustness of control

To see how robust is our ability to control a network under unavoidable link failure, we classify each link into one of the following three categories (Fig. 1d, g, j): 'critical' if in its absence we need to increase the number of driver nodes to maintain full control; 'redundant' if it can be removed without affecting the current set of driver nodes; or 'ordinary' if it is neither critical nor redundant. Figure 4 shows the densities of critical ($l_c = L_c/L$), redundant ($l_r = L_r/L$) and ordinary ($l_o = L_o/L$) links for each real network, and indicates that most networks have few or no critical links. Most links are ordinary, meaning that they have a role in some control configurations but that the network can still be controlled in their absence.

To understand the factors that determine l_c, l_r and l_o, in Fig. 5a, c, f we show their $\langle k\rangle$ dependence for model systems. The behaviour of l_c is the easiest to understand: for small $\langle k\rangle$, all links are essential for control ($l_c \approx 1$). As $\langle k\rangle$ increases, the network's redundancy increases, decreasing l_c. The increasing redundancy suggests that the density of redundant links, l_r, should always increase with $\langle k\rangle$, but it does not: it reaches a maximum at a critical value of $\langle k\rangle$, $\langle k\rangle_c$, after which it decays. This non-monotonic behaviour results from the competition of two topologically distinct regions of a network, the core and leaves[43]. The core represents a compact cluster of nodes left in the network after applying a greedy leaf removal procedure[48], and leaves are nodes with $k_{in} = 1$ or $k_{out} = 1$ before or during leaf removal. The core emerges through a percolation transition (Fig. 5b, d): for $k < \langle k\rangle_c$, $n_{core} = N_{core}/N = 0$, so the system consists of leaves only (Fig. 5e). At $\langle k\rangle = \langle k\rangle_c$, a small core emerges, decreasing the number of leaves. For Erdős–Rényi networks, we predict that $\langle k\rangle_c = 2e \approx 5.436564$ in agreement with the numerical result (Fig. 5a, b), a value that coincides with $\langle k\rangle$ where l_r reaches its maximum. Indeed, l_r starts decaying at $\langle k\rangle_c$ because for $\langle k\rangle > \langle k\rangle_c$ the number of distinct maximum

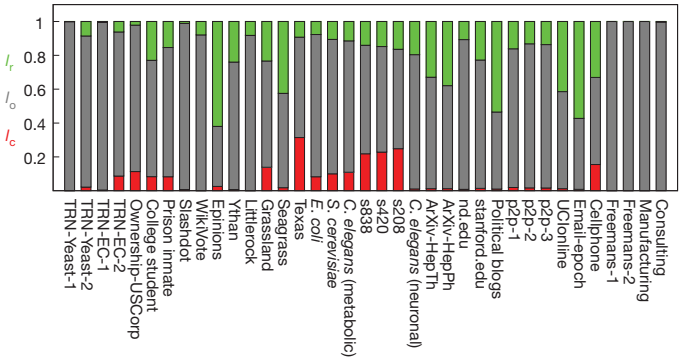

Figure 4 | Link categories for robust control. The fractions of critical (red, l_c), redundant (green, l_r) and ordinary (grey, l_o) links for the real networks named in Table 1. To make controllability robust to link failures, it is sufficient to double only the critical links, formally making each of these links redundant and therefore ensuring that there are no critical links in the system.

matchings increases exponentially (Supplementary Information, section IV.C) and, as a result, the chance that a link does not participate in any control configuration decreases. For scale-free networks, we observe the same behaviour, with the caveat that $\langle k\rangle_c$ decreases with γ (Fig. 5c, d).

Discussion and conclusions

Control is a central issue in most complex systems, but because a general theory to explore it in a quantitative fashion has been lacking, little is known about how we can control a weighted, directed network—the configuration most often encountered in real systems. Indeed, applying Kalman's controllability rank condition (equation (3)) to large networks is computationally prohibitive, limiting previous work to a few dozen nodes at most[17–19]. Here we have developed the tools to address controllability for arbitrary network topologies and sizes. Our key finding, that N_D is determined mainly by the degree

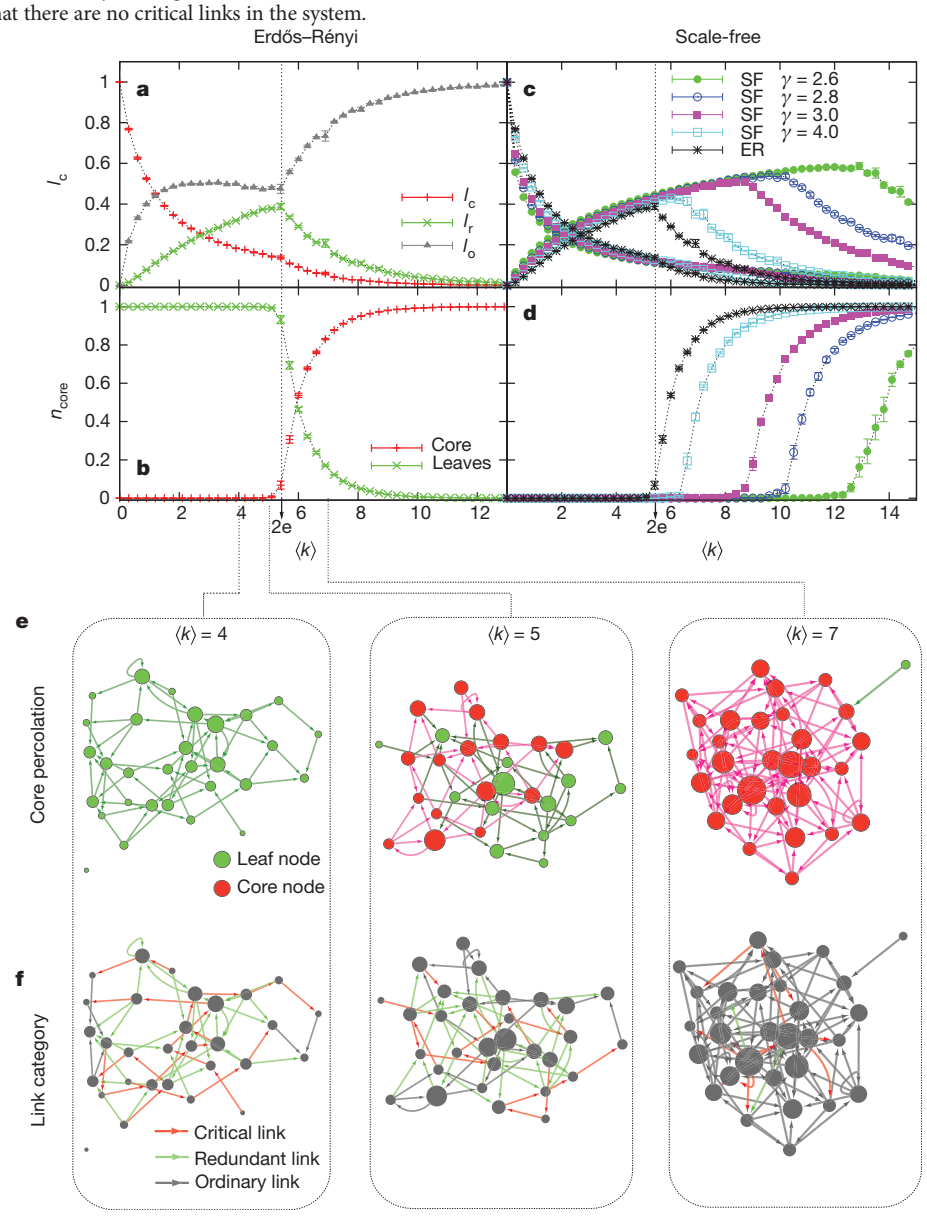

Figure 5 | Control robustness. **a**, Dependence on $\langle k\rangle$ of the fraction of critical (red, l_c), redundant (green, l_r) and ordinary (grey, l_o) links for an Erdős–Rényi network: l_r peaks at $\langle k\rangle = \langle k\rangle_c = 2e$ and the derivative of l_c is discontinuous at $\langle k\rangle = \langle k\rangle_c$. **b**, Core percolation for Erdős–Rényi network occurs at $k = \langle k\rangle_c = 2e$, which explains the l_r peak. **c**, **d**, Same as in **a** and **b** but for scale-free networks. The Erdős–Rényi and scale-free networks[38] have $N = 10^4$ and the results are averaged over ten realizations with error bars defined as s.e.m. Dotted lines are only a guide to the eye. **e**, The core (red) and leaves (green) for small Erdős–Rényi networks ($N = 30$) at different $\langle k\rangle$ values ($\langle k\rangle = 4$, 5, 7). Node sizes are proportional to node degrees. **f**, The critical (red), redundant (green) and ordinary (grey) links for the above Erdős–Rényi networks at the corresponding $\langle k\rangle$ values.

distribution, allows us to use the tools of statistical physics to predict N_D from $P(k_{in}, k_{out})$ analytically, offering a general formalism with which to explore the impact of network topology on controllability.

The framework presented here raises a number of questions, answers to which could further deepen our understanding of control in complex environments. For example, although our analytical work focused on uncorrelated networks, the algorithmic method we developed can identify N_D for arbitrary networks, providing a framework in which to address the role of correlations systematically[40,49,50]. Taken together, our results indicate that many aspects of controllability can be explored exactly and analytically for arbitrary networks if we combine the tools of network science and control theory, opening new avenues to deepening our understanding of complex systems.

Received 18 November 2010; accepted 16 March 2011.

1. Kalman, R. E. Mathematical description of linear dynamical systems. *J. Soc. Indus. Appl. Math. Ser. A* **1,** 152–192 (1963).
2. Luenberger, D. G. *Introduction to Dynamic Systems: Theory, Models, & Applications* (Wiley, 1979).
3. Slotine, J.-J. & Li, W. *Applied Nonlinear Control* (Prentice-Hall, 1991).
4. Kelly, F. P., Maulloo, A. K. & Tan, D. K. H. Rate control for communication networks: shadow prices, proportional fairness and stability. *J. Oper. Res. Soc.* **49,** 237–252 (1998).
5. Srikant, R. *The Mathematics of Internet Congestion Control* (Birkhäuser, 2004).
6. Chiang, M., Low, S. H., Calderbank, A. R. & Doyle, J. C. Layering as optimization decomposition: a mathematical theory of network architectures. *Proc. IEEE* **95,** 255–312 (2007).
7. Wang, X. F. & Chen, G. Pinning control of scale-free dynamical networks. *Physica A* **310,** 521–531 (2002).
8. Wang, W. & Slotine, J.-J. E. On partial contraction analysis for coupled nonlinear oscillators. *Biol. Cybern.* **92,** 38–53 (2005).
9. Sorrentino, F., di Bernardo, M., Garofalo, F. & Chen, G. Controllability of complex networks via pinning. *Phys. Rev. E* **75,** 046103 (2007).
10. Yu, W., Chen, G. & Lü, J. On pinning synchronization of complex dynamical networks. *Automatica* **45,** 429–435 (2009).
11. Marucci, L. *et al.* How to turn a genetic circuit into a synthetic tunable oscillator, or a bistable switch. *PLoS ONE* **4,** e8083 (2009).
12. Strogatz, S. H. Exploring complex networks. *Nature* **410,** 268–276 (2001).
13. Dorogovtsev, S. N. & Mendes, J. F. F. *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford Univ. Press, 2003).
14. Newman, M., Barabási, A.-L. & Watts, D. J. *The Structure and Dynamics of Networks* (Princeton Univ. Press, 2006).
15. Caldarelli, G. *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford Univ. Press, 2007).
16. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74,** 47–97 (2002).
17. Tanner, H. G. in *Proc. 43rd IEEE Conf. Decision Contr.* Vol. 3, 2467–2472 (2004).
18. Lombardi, A. & Hörnquist, M. Controllability analysis of networks. *Phys. Rev. E* **75,** 56110 (2007).
19. Liu, B., Chu, T., Wang, L. & Xie, G. Controllability of a leader–follower dynamic network with switching topology. *IEEE Trans. Automat. Contr.* **53,** 1009–1013 (2008).
20. Rahmani, A., Ji, M., Mesbahi, M. & Egerstedt, M. Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM J. Contr. Optim.* **48,** 162–186 (2009).
21. Kim, D.-H. & Motter, A. E. Slave nodes and the controllability of metabolic networks. *N. J. Phys.* **11,** 113047 (2009).
22. Mesbahi, M. & Egerstedt, M. *Graph Theoretic Methods in Multiagent Networks* (Princeton Univ. Press, 2010).
23. Motter, A. E., Gulbahce, N., Almaas, E. & Barabási, A.-L. Predicting synthetic rescues in metabolic networks. *Mol. Syst. Biol.* **4,** 168 (2008).
24. Pastor-Satorras, R. & Vespignani, A. *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge Univ. Press, 2004).
25. Lezon, T. R., Banavar, J. R., Cieplak, M., Maritan, A. & Fedoroff, N. V. Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. *Proc. Natl Acad. Sci. USA* **103,** 19033–19038 (2006).
26. Lin, C.-T. Structural controllability. *IEEE Trans. Automat. Contr.* **19,** 201–208 (1974).
27. Shields, R. W. & Pearson, J. B. Structural controllability of multi-input linear systems. *IEEE Trans. Automat. Contr.* **21,** 203–212 (1976).
28. Lohmiller, W. & Slotine, J.-J. E. On contraction analysis for nonlinear systems. *Automatica* **34,** 683–696 (1998).
29. Yu, W., Chen, G., Cao, M. & Kurths, J. Second-order consensus for multiagent systems with directed topologies and nonlinear dynamics. *IEEE Trans. Syst. Man Cybern. B* **40,** 881–891 (2010).
30. Hopcroft, J. E. & Karp, R. M. An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. *SIAM J. Comput.* **2,** 225–231 (1973).
31. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature* **406,** 378–382 (2000).
32. Cohen, R., Erez, K., Ben-Avraham, D. & Havlin, S. Resilience of the Internet to random breakdowns. *Phys. Rev. Lett.* **85,** 4626–4628 (2000).
33. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86,** 3200–3203 (2001).
34. Nishikawa, T., Motter, A. E., Lai, Y.-C. & Hoppensteadt, F. C. Heterogeneity in oscillator networks: are smaller worlds easier to synchronize? *Phys. Rev. Lett.* **91,** 014101 (2003).
35. Erdős, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5,** 17–60 (1960).
36. Bollobás, B. *Random Graphs* (Cambridge Univ. Press, 2001).
37. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286,** 509–512 (1999).
38. Goh, K.-I., Kahng, B. & Kim, D. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.* **87,** 278701 (2001).
39. Chung, F. & Lu, L. Connected component in random graphs with given expected degree sequences. *Ann. Combin.* **6,** 125–145 (2002).
40. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296,** 910–913 (2002).
41. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298,** 824–827 (2002).
42. Mézard, M. & Parisi, G. The Bethe lattice spin glass revisited. *Eur. Phys. J. B* **20,** 217–233 (2001).
43. Zdeborová, L. & Mézard, M. The number of matchings in random graphs. *J. Stat. Mech.* **05,** 05003 (2006).
44. Zhou, H. & Ou-Yang, Z.-c. Maximum matching on random graphs. Preprint at ⟨http://arxiv.org/abs/cond-mat/0309348⟩ (2003).
45. 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).
46. Boguñá, M., Pastor-Satorras, R. & Vespignani, A. Cut-offs and finite size effects in scale-free networks. *Eur. Phys. J. B* **38,** 205–209 (2004).
47. Lovász, L. & Plummer, M. D. *Matching Theory* (American Mathematical Society, 2009).
48. Bauer, M. & Golinelli, O. Core percolation in random graphs: a critical phenomena analysis. *Eur. Phys. J. B* **24,** 339–352 (2001).
49. Newman, M. E. J. Assortative mixing in networks. *Phys. Rev. Lett.* **89,** 208701 (2002).
50. Pastor-Satorras, R., Vázquez, A. & Vespignani, A. Dynamical and correlation properties of the Internet. *Phys. Rev. Lett.* **87,** 258701 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Song, G. Bianconi, H. Zhou, L. Vepstas, N. Gulbahce, H. Jeong, Y.-Y. Ahn, B. Barzel, N. Blumm, D. Wang, Z. Qu and Y. Li for discussions. This work was supported by the Network Science Collaborative Technology Alliance sponsored by the US Army Research Laboratory under Agreement Number W911NF-09-2-0053; the Office of Naval Research under Agreement Number N000141010968; the Defense Threat Reduction Agency awards WMD BRBAA07-J-2-0035 and BRBAA08-Per4-C-2-0033; and the James S. McDonnell Foundation 21st Century Initiative in Studying Complex Systems.

Author Contributions All authors designed and did the research. Y.-Y.L. analysed the empirical data and did the analytical and numerical calculations. A.-L.B. was the lead writer of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.-L.B. (alb@neu.edu).

SUPPLEMENTARY INFORMATION

doi:10.1038/nature10011

Contents

I. Introduction 2

II. Previous Work and Relation of Controllability to Other Problems 2

A. Controllability of Undirected Network 2

B. Network Synchronizability 4

C. Congestion Control 4

III. Structural Control Theory 5

A. Why Linear Dynamics? 6

B. Simple Examples 8

C. Structural Controllability Theorem 10

D. Minimum Inputs Theorem 13

IV. Maximum Matching 16

A. Statistical physics description 16

B. Internal energy 17

1. $r-$regular random digraph 21

2. Poisson-distributed digraph 23

3. Exponentially distributed digraph 24

4. Power-law distributed digraphs 25

5. Static model 27

6. Chung-Lu model 29

C. Entropy 32

V. Control Robustness 37

VI. Network Datasets 38

References 42

I. INTRODUCTION

This Supplementary Information is organized as follows. In Sec. II, we clarify the fundamental differences between our work and previous research on network controllability. In Sec. III, we give a short introduction to the structural control theory, which can be simply applied to directed networks. The reasons why we focus on linear dynamics are given in Sec. III A. Some simple examples in Sec. III B demonstrate the difference between controllability, structural controllability and strong structural controllability. The sufficient and necessary conditions for a linear system to be structurally controllable are given by Lin's structural controllability theorem (SCT), which is discussed in Sec. III C. Based on SCT, we derived the minimum input theorem in Sec. III D, which gives the minimum number of inputs that we need to fully control a directed network. This theorem also enables us to find the driver nodes to which the external inputs should be injected, based on a deep relation between structural controllability and maximum matching. In Sec. IV, we analytically derived the average size and number of the maximum matchings for a random directed network ensemble with a prescribed degree distribution, using the cavity method developed in statistical physics. In Sec. V, we show the results on control robustness against node failure, compared to the results on link failure shown in the main text. The real-world networks analyzed in this work are listed and briefly described in Sec. VI.

II. PREVIOUS WORK AND RELATION OF CONTROLLABILITY TO OTHER PROBLEMS

A. Controllability of Undirected Network

Network controllability is a vast area of research with a long history[1–5]. Here, we clarify the relationship and differences between our work and earlier research in this direction. In particular, we want to emphasize that:

1. The classical concept and condition of controllability was usually applied to *undirected* networks. While the control of undirected networks is a problem with its own intellectual challenges, it has so far found little applications to complex systems. The reason is that to best of our knowledge, most, if not all, real-world complex networks where control is

expected to have an impact are *directed*. This dramatically limits the applicability of the previous results.

2. From a numerical point of view, the previous approaches are limited to small networks due to two levels of difficulties: (1) To identify the minimum set of driver nodes, the number of configurations to be tested increase like $\mathcal{O}(2^N)$, prohibitive for networks with thousands to millions of nodes. (2) Kalman's rank condition is difficult to test for networks with millions of nodes. For these reasons, the largest network size explored in pervious work is very limited, a few dozen nodes at most[1–3], which pale compared to any known network of interest for network science or complex systems.

3. Previous research implicitly assumes the interaction strengths or link weights can be exactly measured. Many of them even assume that all links have identical weights. Consequently, the controllability problem is significantly simplified, naturally translating into a spectral graph theoretic problem, e.g. the spectrum of the Laplacian matrix of the network[4, 5]. Despite some rigorous work relating network controllability to its symmetry structure, we argue that the assumption the link weights can be exactly measured is not realistic in most real-world networks, because measurement error and uncertainties cannot be avoided in real systems. For example, in regulatory networks, which are directed and weighted, we do not have the tools yet to estimate the link weights.

   Due to the undirected and unweighted assumptions, the obtained conclusions are misleading in some sense. For example, some authors arrived to the conclusion that a complete graph is uncontrollable with a single leader (driver node in our language). Mathematically, this conclusion is correct, and it follows rigorously from the formalism used by the authors[1, 4]. But the result heavily depends on the assumption that all the link weights are exactly known and the same. As long as there are some uncertainties introduced in the link weights, this conclusion needs to be revisited. Actually we know that for the linear system

$$\dot{\mathbf{X}}(t) = \mathbf{A} \cdot \mathbf{X}(t) + \mathbf{B} \cdot \mathbf{u}(t), \tag{S1}$$

   the set of all controllable pairs is open and dense in the space of all pairs $(\mathbf{A}, \mathbf{B})$ with standard metric[6]. In other words, if a pair $(\mathbf{A}_0, \mathbf{B}_0)$ is not controllable, then for every $\epsilon > 0$, there exists a completely controllable pair $(\mathbf{A}, \mathbf{B})$ with $||\mathbf{A} - \mathbf{A}_0|| < \epsilon$ and $||\mathbf{B} - \mathbf{B}_0|| < \epsilon$ where $||\cdot||$ denotes matrix norm. That is, an uncontrollable system will become controllable if we slightly change the weights of some links. This is the key reason why some conclusions of

previous works need to be rechecked whenever there are parameter uncertainties. In fact, our work based on structural controllability predicts that a complete graph is structurally controllable with any single node chosen as a driver.

In contrast with these limitations, our results offer the analytic tools to explore directed networks (hence overcoming the limitation (1)), of arbitrary size (overcoming limitation (2)) and of arbitrary weights (bypassing limitation (3)).

B. Network Synchronizability

Network synchronizability[7–12] is often also called "controllability" or "pinning controllability". We emphasize that the controllability we study here is fundamentally different from synchronizability. We ask whether the system can fully explore its state space, while synchronizability addresses whether the system can exhibit specific spatio-temporal symmetries (for instance, whether all its nodes can follow a common time-varying trajectory) and the stability of such symmetries. For example, we show that driver nodes tend to avoid hubs, while on the contrary hubs can facilitate synchrony by providing common inputs to downstream nodes[12, 13].

The consensus or agreement problem in multi-agent systems[5, 14] can be viewed as a particular case of the synchronization problem. Generally speaking, the consensus problem studies processes by which a collection of interacting agents converge to a common value or achieve a common goal. This is fundamentally different from the controllability problem we address here.

C. Congestion Control

Congestion control [15–17] is a problem of similar difficulty and importance with the controllability problem, and occasionally the two are mentioned together. Yet, they are fundamentally different.

Roughly speaking, congestion control is a technique to monitor network utilization and manipulate data transmission rate so that network traffic jam or congestion is avoided. Obviously, this is of particular interest for communication networks, e.g. the Internet. For example, in the Transmission Control Protocol (TCP), the congestion control mechanism has been implemented to facilitate reliable and fast data transfer. Congestion control has been formalized as a distributed optimization problem[15]: The objective is to maximize the sum of source (or user) utilities $U(x_i)$ as functions of rates $\{x_i\}$, the constraints are linear flow constraints, and optimization variables

are source rates $\{x_i\}$. The utility $U(x_i)$ measures how much benefit a source i obtains by transmitting at rate x_i. This optimization problem is often called network utility maximization (NUM) problem[17]. In fact, each variant of congestion control protocol, e.g. TCP, can be viewed as a distributed algorithm solving a specified basic NUM problem with a particular type of utility function.

From a mathematical point of view, congestion control is a mechanism for resource allocation[15, 16], which is unique to transport systems, most notably, the Internet. Controllability, however, is a far more general problem, with applications from communication systems to the cell.

III. STRUCTURAL CONTROL THEORY

We introduce a linear time-invariant (LTI) dynamics (Eq. 1) on a directed network $G(\mathbf{A})$. In control theory, $\mathbf{X} \in \mathbb{R}^N$ is called the state vector, $\mathbf{A} \in \mathbb{R}^{N\times N}$ is the state matrix, $\mathbf{B} \in \mathbb{R}^{N\times M}$ is the input matrix, and $\mathbf{u} \in \mathbb{R}^M$ is the input or control vector. Note that in the state matrix $\mathbf{A} := (a_{ij})_{N\times N}$, the element a_{ij} is 0 if $(j \to i)$ is not a link in $G(\mathbf{A})$, i.e. node-j does not affect node-i. Otherwise, a_{ij} gives the strength or weight that node-j can affect node-i. Positive (or negative) value of a_{ij} means the link $(j \to i)$ is excitatory (or inhibitory). Note that if all the links are excitatory with unit strength, then the state matrix $\mathbf{A}$ is just the transpose of the adjacency matrix of the directed network.

The whole system, denoted as $(\mathbf{A}, \mathbf{B})$, can be represented by a digraph $G(\mathbf{A}, \mathbf{B}) = (V, E)$ with $V = V_A \cup V_B$ the vertex set and $E = E_A \cup E_B$ the edge set. Here, $V_A = \{x_1, \cdots, x_N\} := \{v_1, \cdots, v_N\}$ is the set of *state vertices*, corresponding to the N nodes in the network (see Fig. S2a). $V_B = \{u_1, \cdots, u_M\} := \{v_{N+1}, \cdots, v_{N+M}\}$ is the set of *input vertices*, corresponding to the M inputs (see Fig. S2b). $E_A = \{(x_j, x_i)|a_{ij} \neq 0\}$ is the set of edges between state vertices, i.e. the links in the network. $E_B = \{(u_j, x_i)|b_{ij} \neq 0\}$ is the set of edges between input vertices and state vertices. The M input vertices are also called the *origins* of the digraph $G(\mathbf{A}, \mathbf{B})$. The state vertices connected to the origins are called *controlled nodes*, e.g. x_1, x_2, x_3, x_4 and x_5 in Fig. S2b. Denote the number of controlled nodes as M', one has $M' \geq M$ because one input vertex can be connected to multiple state vertices. Those controlled nodes which do not share input vertices are called *driver nodes*, e.g. x_2, x_4 and x_5 in Fig. S2b. Obviously, the number of driver nodes equals M, which is the number of inputs. Any system is fully controllable if we control each node individually, i.e. $M = N$. The question is therefore how to identify the minimum number of inputs (denoted as N_{I}) or equivalently the minimum number of driver nodes (denoted as N_{D}). By

definition, $N_{\rm D} = N_{\rm I} = M_{\rm min}$.

The concept of *structural controllability* was first introduced by Lin in 1970s [6]. The basic idea follows. In the LTI system (Eq. S1), the matrices $\mathbf{A}$ and $\mathbf{B}$ are considered to be structured ones, i.e. their elements are either fixed zeros or independent free parameters. This reflects the fact that in reality the system parameters are often not known precisely except the zeros that mark the absence of connections between components of the system. The system $(\mathbf{A}, \mathbf{B})$ is called *structurally controllable* if it is possible to fix the free parameters in $\mathbf{A}, \mathbf{B}$ to certain values so that the obtained system $(\mathbf{A}, \mathbf{B})$ is controllable in the usual sense ($\text{rank}\,\mathbf{C} = N$). Note that if a system is structurally controllable then it is controllable for almost all parameter values except for those in some *proper algebraic variety* in the parameter space. Since a proper algebraic variety has Lebesgue measure zero [6, 18], structural controllability is a generic property of the system. Roughly speaking, a structurally controllable system is either controllable or becomes controllable after slightly changing the weights of certain interconnections, and it remains controllable for possibly large changes of the link weights. As far as physical systems are concerned, structural controllability in practice implies controllability, since a possible loss of controllability of a structurally controllable system can occur only in pathological cases when there are accidental constraints of the system parameters. When the system is controllable for any values (other than zero) of the indeterminate parameters of the system may take, it is called *strongly structurally controllable* [19].

Note that imposing structural controllability for complex networks is also akin in spirit to the notion of using strong robustness as a means to select dynamic models for biological systems [20].

A. Why Linear Dynamics?

The linear dynamics (Eq.S1) may be suitable for the consensus or agreement dynamics in multi-agent networks[1, 3–5]. But we have to admit that by using linear dynamics we did not intend to model any particular type of dynamical processes on any particular networks. The dynamical rules on real-world networks are so diverse that writing a general dynamical equation that would capture them all is plainly impossible. Moreover, for many networks, especially biological networks, we do not even know the dynamical rules. We expect that real systems are driven by some nonlinear dynamical rules. Yet, the linear dynamics (Eq.S1) is still essential (and unavoidable) in the path towards addressing the controllability of these systems, for reasons explained below:

Most importantly, studying linear controllability offers a framework to explain the controllability of nonlinear systems. Indeed, a basic starting point for exploring the controllability of any nonlinear

system is the study of controllability of the linearized version of the nonlinear dynamical system, e.g. Eq.S1.

Moreover, as we focus on the structural version of controllability, the fact that the controllability matrix of the linearized system has full rank at all points is *sufficient* for most systems to yield controllability of the actual nonlinear system [21]. In other words, because structural controllability leads to controllability of a continuum of linearized systems, our results can provide sufficient controllability conditions in the full nonlinear case. This is not particularly surprising, since knowing a nonlinear dynamics at just one point and its linearizations everywhere (or more generally, almost everywhere in a convex connected region) is equivalent to knowing the dynamics everywhere (or everywhere in the region), a remark which also forms the basis of modern stability and convergence analysis tools such as nonlinear contraction theory [22]. Indeed, for arbitrary P and N, and any smooth function $\mathbf{f} : \mathbb{R}^P \times \mathbb{R}^+ \to \mathbb{R}^N$, one has

$$\mathbf{f}(\mathbf{r}_2, t) = \mathbf{f}(\mathbf{r}_1, t) + \left(\int_{\lambda=0}^{\lambda=1} \frac{\partial}{\partial \mathbf{r}} \mathbf{f}(\mathbf{r}_1 + \lambda(\mathbf{r}_2 - \mathbf{r}_1), t) \, \mathrm{d}\lambda \right) \cdot (\mathbf{r}_2 - \mathbf{r}_1). \tag{S2}$$

Intuitively, around each linearization point the system remains controllable in a finite ball, by continuity, so that a trajectory between any two points can be covered by a finite number of such controllability balls. In general, nonlinearity may actually enhance a system's controllability by helping the system explore the full state space [21, 23, 24]. Fully exploiting these nonlinear effects is the subject of future research.

Finally, the non-trivial topology of real-world networks brings an intrinsic layer of complexity to the controllability problem. Before we explore the fully nonlinear dynamical setting, we have to understand the effects of topological characteristics on linear controllability, which naturally serves as a prerequisite of nonlinear controllability problem. From the advances towards understanding complex networks accumulated in the last decade, we know that network topology fundamentally affects the dynamical processes on it, from epidemic spreading to synchronization phenomenon. It is fair to expect that the network topology would definitely affect controllability as well. To approach this problem in a systematic fashion, we have to proceed from the simplest dynamics that offers a mathematically correct description of controllability, helping us avoid any entanglement due to nonlinear effects.

Notice that the framework we developed applies similarly to discrete-time system[25, 26].

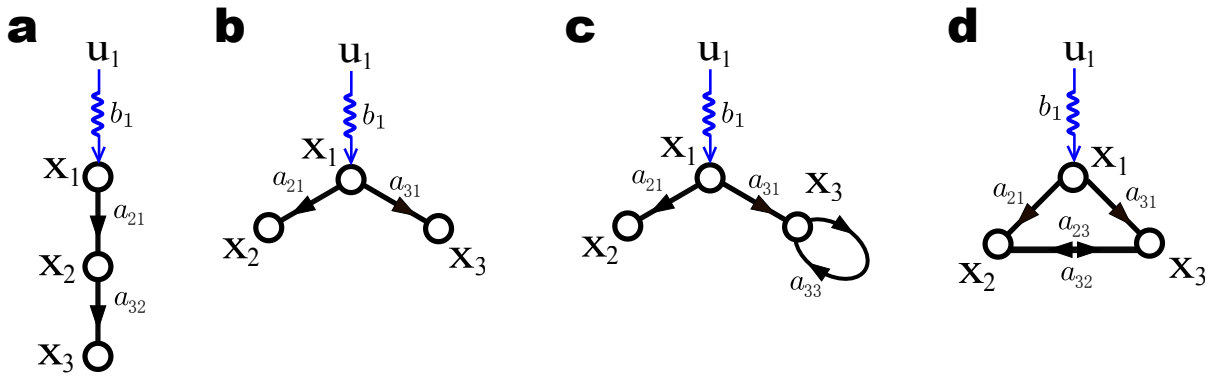

Figure S1: **Control a simple network. a**, A directed path can be completely controlled by controlling the starting node only. The controllability is independent of the detailed (non-zero) values of b_1, a_{21}, and a_{32}, so the system is strongly structurally controllable. **b**, A directed star can never be completely controlled by controlling the central hub (node x_1) only. **c**, This example network, generated by adding a self-edge to the star shown in **b**, can be completely controlled by controlling node x_1 only. The controllability is independent of the detailed (non-zero) values of b_1, a_{21}, a_{31}, and a_{33}, so the system is strongly structurally controllable. **d**, This network is completely controllable for almost all weights combinations. It will be uncontrollable only in some pathological cases, e.g. those weights satisfy the constraint $a_{32}a_{21}^2 = a_{23}a_{31}^2$ exactly, so the system is structurally controllable.

B. Simple Examples

Here, we show simple examples to illustrate the difference between controllability, structural controllability, and strong structural controllability.

The linear dynamics shown in Fig. S1(a) can be written as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ 0 \\ 0 \end{bmatrix} u(t).$$

The controllability matrix is given by

$$\mathbf{C} = [\mathbf{B}, \mathbf{A}\cdot\mathbf{B}, \mathbf{A}^2\cdot\mathbf{B}] = b_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & a_{21} & 0 \\ 0 & 0 & a_{32}a_{21} \end{bmatrix}.$$

Since $\operatorname{rank}\mathbf{C} = 3 = N$, the system is controllable. Note that the system is always controllable as long as those weights, i.e. a_{21}, a_{32} and b_1, are non-zero. In other words, its controllability is independent of the detailed values of a_{21}, a_{32} and b_1. This is an example of the so called *strong structural controllability* [19].

The linear dynamics shown in Fig. S1(b) can be written as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ 0 \\ 0 \end{bmatrix} u(t).$$

The controllability matrix is given by

$$\mathbf{C} = [\mathbf{B}, \mathbf{A} \cdot \mathbf{B}, \mathbf{A}^2 \cdot \mathbf{B}] = b_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & a_{21} & 0 \\ 0 & a_{31} & 0 \end{bmatrix}.$$

Since $\operatorname{rank} \mathbf{C} = 2 < N$, the system is uncontrollable. Note that this is independent of the detailed values of a_{21}, a_{31}, and b_1. No matter how we tune them, the system is uncontrollable. Indeed, the dynamics equation suggests that the system will get stuck in the plane $a_{31}x_2(t) = a_{21}x_3(t)$ in the state space.

The linear dynamics shown in Fig. S1(c) can be written as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & 0 & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ 0 \\ 0 \end{bmatrix} u(t).$$

The controllability matrix is given by

$$\mathbf{C} = [\mathbf{B}, \mathbf{A} \cdot \mathbf{B}, \mathbf{A}^2 \cdot \mathbf{B}] = b_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & a_{21} & 0 \\ 0 & a_{31} & a_{33}a_{31} \end{bmatrix}.$$

Since $\operatorname{rank} \mathbf{C} = 3 = N$, the system is controllable. In this example, the controllability is independent of the detailed values of a_{21}, a_{31} , a_{33}, and b_1, as long as they are non-zero. This is another example of strong structural controllability. Notice the small difference between the networks shown in Fig. S1(b) and Fig. S1(c). The presence of a self-edge fundamentally changes the controllability of the system.

The linear dynamics shown in Fig. S1(d) can be written as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ 0 \\ 0 \end{bmatrix} u(t).$$

The controllability matrix can be easily calculated

$$\mathbf{C} = [\mathbf{B}, \mathbf{A}\cdot\mathbf{B}, \mathbf{A}^2\cdot\mathbf{B}] = b_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & a_{21} & a_{23}a_{31} \\ 0 & a_{31} & a_{32}a_{21} \end{bmatrix}.$$

In most cases, we have rank $\mathbf{C} = 3 = N$, so the system will be controllable. In case $\binom{a_{21}}{a_{31}} \propto \binom{a_{23}a_{31}}{a_{32}a_{21}}$, e.g. $a_{32}a_{21}^2 = a_{23}a_{31}^2$, then rank $(\mathbf{C}) = 2 < N$ and the system will be uncontrollable. However, this case is pathological. In fact, slightly change any link's weight will break the constraint and make the system controllable. Thus, we call this system structurally controllable. (Apparently, it is not strongly structurally controllable.)

C. Structural Controllability Theorem

The structural controllability of LTI systems has been well studied after Lin's seminal work [6, 18, 27–31]. Both sufficient and necessary conditions for a system to be structurally controllable are given by Lin's *structural controllability theorem*, which has clear graph-theoretic interpretations. Before we state Lin's theorem, we introduce several definitions.

For a general graph, it is said to be *covered* or *spanned* by a subgraph if the subgraph and the graph have the same vertex set. For a digraph, a sequence of oriented edges $\{(v_1 \to v_2), (v_2 \to v_3), \cdots, (v_{k-1} \to v_k)\}$ where all vertices $\{v_1, v_2, \cdots, v_k\}$ are distinct is called an *elementary path*. When v_k coincides with v_1, it is called an *elementary cycle*.

Now we consider the particular digraph $G(\mathbf{A}, \mathbf{B})$ describing the controlled network (see Fig. S2b).

Definition 1 (inaccessibility)**.** *A state vertex x_i in the digraph $G(\mathbf{A}, \mathbf{B})$ is called* inaccessible *iff there are no directed paths reaching x_i from the input vertices (origins).*

Definition 2 (dilation)**.** *The digraph $G(\mathbf{A}, \mathbf{B})$ contains a* dilation *iff there is a subset $S \subset V_A$ such that $|T(S)| < |S|$. Here, the neighborhood set $T(S)$ of a set S is defined to be the set of all vertices v_j that there exists an oriented edge from v_j to a vertex in S, i.e. $T(S) = \{v_j | (v_j \to v_i) \in E(G), v_i \in S\}$. The origins are not allowed to belong to S but may belong to $T(S)$. $|S|$ or $|T(S)|$ is the cardinality of set S or $T(S)$, respectively.*

Remark The definition of inaccessibility can not be absorbed into that of dilation because an inaccessible node could have a self-edge, which causes no dilation. In case all nodes are isolated and each of them has a self-edge, then there are no dilations, but all nodes are inaccessible!

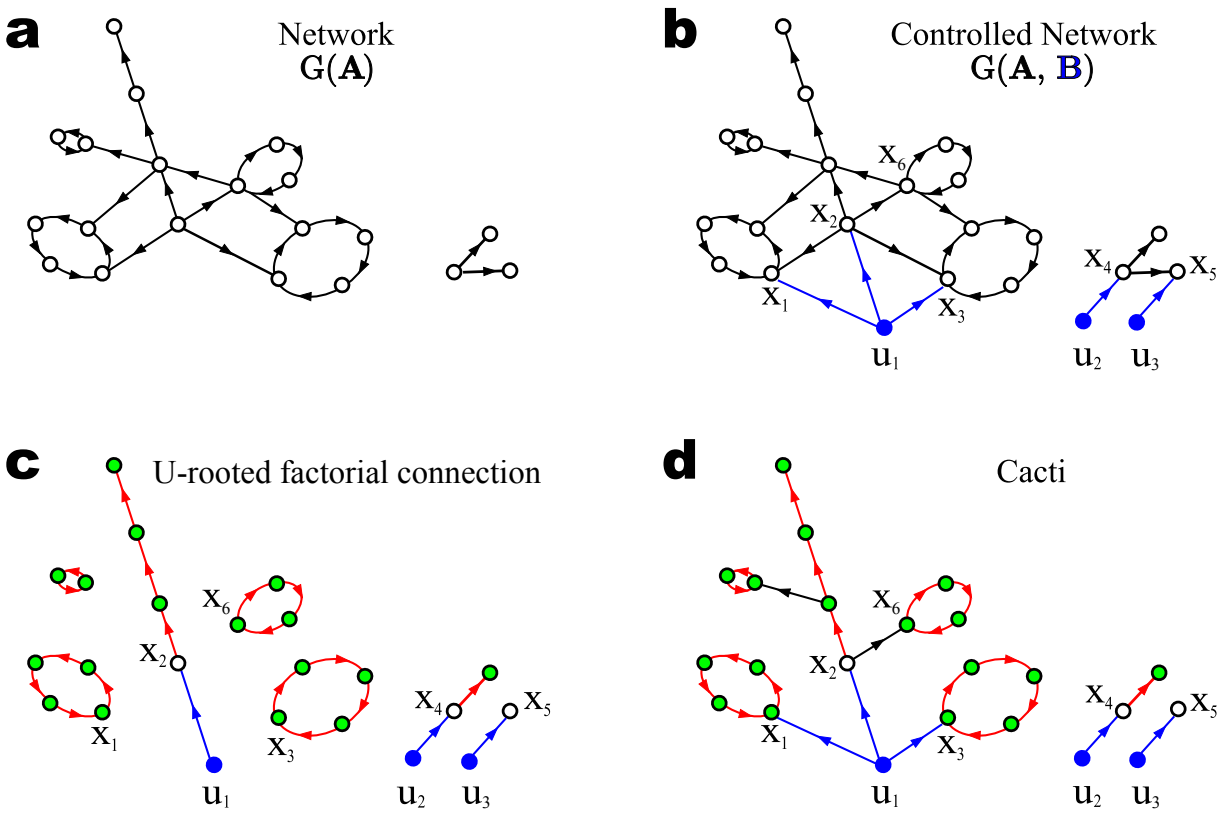

Figure S2: **A schematic digram shows the control of a directed network.** **a**, A directed network $G(\mathbf{A})$. All nodes in the network are called state vertices. **b**, The network is controlled by three input vertices (origins), which are marked in blue. The controlled network is described by a digraph $G(\mathbf{A}, \mathbf{B})$. The state vertices pointed by the inputs are called controlled nodes, e.g. x_1, x_2, x_3, x_4 and x_5. **c**, The U-rooted factorial connection of the digraph $G(\mathbf{A}, \mathbf{B})$ is composed of vertex-disjoint stems and cycles. Matching edges are marked in red. Matched nodes are marked in green. The input vertices are connected to unmatched nodes (marked in white), forming stems. The unmatched nodes, e.g. x_2, x_4 and x_5, are called the independent controlled nodes or driver nodes, which do not share the input signals. **d**, The cacti is built from the U-rooted factorial connection. The cactus in the left contains a stem and four buds. Both the cactus in the middle and the cactus in the right are just stems.

Definition 3 (stem)**.** *A* stem *is an elementary path originating from an input vertex. The initial (or terminal) vertex of a stem is called the* root *(or* top*) of the stem.*

Definition 4 (bud)**.** *A* bud *is an elementary cycle C with an additional edge e that ends, but not begins, in a vertex of the cycle. This additional edge e is called the distinguished edge of the bud, e.g. $(u_1 \to x_3)$ or $(x_2 \to x_6)$ in Fig. S2d.*

Definition 5 (U-rooted factorial connection)**.** *A set of vertex-disjoint stems and elementary cycles such that the union of all the stems and all the cycles spans $G(\mathbf{A}, \mathbf{B})$ is called a* U-rooted factorial

connection *in* $G(\mathbf{A}, \mathbf{B})$.

Remark It is easy to show that if and only if there are no dilations in $G(\mathbf{A}, \mathbf{B})$ then there is a U-rooted factorial connection in $G(\mathbf{A}, \mathbf{B})$. See Fig. S2c.

Definition 6 (cactus)**.** *A* cactus *is a subgraph defined recursively as follows. A stem is a cactus. Given a stem* S_0 *and buds* $B_1, B_2, \cdots, B_l$*, then* $S_0 \cup B_1 \cup B_2 \cup \cdots \cup B_l$ *is a cactus if for every* i $(1 \leq i \leq l)$ *the initial vertex of the distinguished edge of* B_i *is not the top of* S_0 *and is the only vertex belonging at the same time to* B_i *and* $S_0 \cup B_1 \cup B_2 \cup \cdots \cup B_{i-1}$*. A set of vertex-disjoint cacti is called a* cacti*.*

Remark Cactus (or cacti) is the *minimal structure* which contains neither inaccessible nodes nor dilations. That is, for a given cactus, the removal of an arbitrary edge will cause either inaccessibility or dilation.

Theorem 1 (Lin's Structural Controllability Theorem)**.**
The following three statements are equivalent:

1. *A linear control system* $(\mathbf{A}, \mathbf{B})$ *is structurally controllable.*
2. *i) The digraph* $G(\mathbf{A}, \mathbf{B})$ *contains no inaccessible nodes.*
   ii) The digraph $G(\mathbf{A}, \mathbf{B})$ *contains no dilation.*
3. $G(\mathbf{A}, \mathbf{B})$ *is spanned by cacti.*

Remarks

1. Lin's theorem has pure algebraic meanings [18]. The presence of inaccessible state vertices means that the structured matrix $[\mathbf{A}; \mathbf{B}]$ is *reducible*, i.e. there exists a permutation matrix $\mathbf{P}$ such that

$$\mathbf{PAP}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}; \quad \mathbf{PB} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B}_2 \end{bmatrix} \tag{S3}$$

   where $\mathbf{A}_{11} \in \mathbb{R}^{K \times K}$, $\mathbf{A}_{21} \in \mathbb{R}^{(N-K) \times K}$, $\mathbf{A}_{22} \in \mathbb{R}^{(N-K) \times (N-K)}$, and $\mathbf{B}_2 \in \mathbb{R}^{(N-K) \times M}$ with $1 \leq K \leq N$. The presence of a dilation is equivalent to the statement that the structured matrix $[\mathbf{A}; \mathbf{B}]$ has *generic rank* less than N:

$$\mathrm{rank}_{\mathrm{g}} [\mathbf{A}; \mathbf{B}] < N. \tag{S4}$$

Here, the generic rank of a structured matrix $\mathbf{M}$ is defined to be the maximum rank that $\mathbf{M}$ can attain as a function of all the free parameters in $\mathbf{M}$. Then Lin's theorem can be written as: A linear control system $(\mathbf{A}, \mathbf{B})$ is structurally controllable iff the structured matrix $[\mathbf{A}; \mathbf{B}]$ is irreducible and has generic rank N.

2. Lin's theorem also has intuitive explanations: (1) A system is uncontrollable if there are inaccessible nodes, which cannot be accessed or "influenced" by the external inputs. A naive example would be isolated nodes. Another example is the following: By controlling a directed path with a driver node which is not the starting node of the path, we cannot fully control the path, simply because we lose control of all the nodes upstream from the driver node, which are inaccessible from the driver node. (2) A system is also uncontrollable if there are dilations. Roughly speaking, dilations are subgraphs in which there are more nodes pointed or "ruled" by less other nodes. (In other words, there are more "subordinates" than "superiors".) For example, in a directed star with many leaf-nodes pointed by one central hub, any two leaf-nodes together with the central hub form a dilation. If we control the central hub only, the system is uncontrollable because we cannot independently control the difference of any two leaf-nodes state variables. In other words, we cannot independently control two subordinates if they share one superior.

   To fully control a network, one has to remove all possible dilations and make sure every node is accessible from the external inputs. Colloquially, this means: each node must have its own "superior". If a node has no "superiors", i.e. no nodes pointing to it, then obviously it is inaccessible and we lose control of it. If two or more nodes share one "superior", then dilation occurs, we cannot fully control the system either.

D. Minimum Inputs Theorem

In this section, we prove that in order to fully control a network $G(\mathbf{A})$ the minimum number of input vertices (or equivalently the minimum number of driver nodes) we need is related to the size of *maximum matching* in the corresponding digraph $G(\mathbf{A})$, which is one of our key results. To achieve this, we first generalize the concept of matching in undirected graph to digraph.

Definition 7. *For an undirected graph, a* matching *M is an independent edge set, i.e. a set of edges without common vertices. A vertex is* matched *if it is incident to an edge in the matching. Otherwise the vertex is* unmatched.

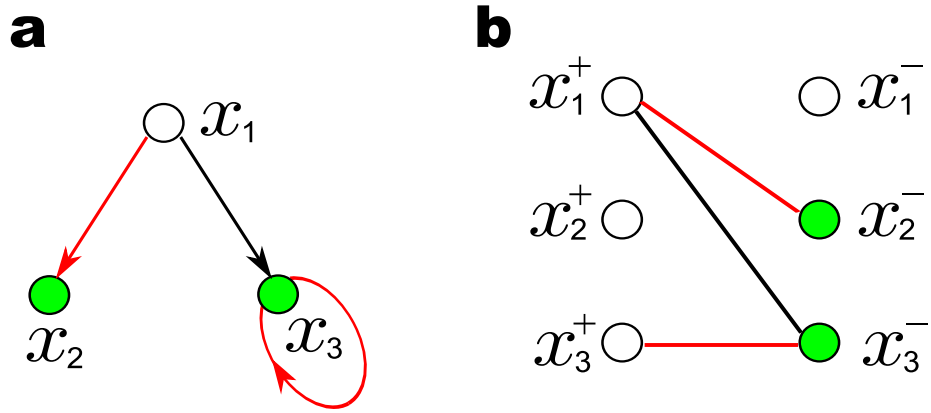

Figure S3: **Matching in digraph and its bipartite representation.** **a**, A simple digraph. **b**, The bipartite representation of the digraph shown in **a**. It has a unique maximum matching, which is shown in red. Matched (or unmatched) nodes are shown in green (or white), respectively.

Definition 8. *For a digraph, an edge subset M is a* matching *if no two edges in M share a common starting vertex or a common ending vertex. A vertex is* matched *if it is an ending vertex of an edge in the matching. Otherwise, it is* unmatched.

In both cases, a matching of maximum cardinality/size is called a *maximum matching*. (Note that in general there could be many different maximum matchings for a given graph or digraph.) A maximum matching is called *perfect* if all vertices are matched. For example, in a directed elementary cycle, all vertices are matched. Note that a maximum matching of a digraph $G(\mathbf{A})$ can be easily found in its bipartite representation, denoted as $H(\mathbf{A})$. See Fig. S3. The bipartite graph is defined in the following way. $H(\mathbf{A}) = (V_A^+ \cup V_A^-, \Gamma)$. Here, $V_A^+ = \{x_1^+, \cdots, x_N^+\}$ and $V_A^- = \{x_1^-, \cdots, x_N^-\}$ are the set of vertices corresponding to the N columns and rows of the state matrix $\mathbf{A}$, respectively. Edge set $\Gamma = \{(x_j^+, x_i^-) \mid a_{ij} \neq 0\}$. For a general bipartite graph, its maximum matching can be found efficiently using the well-known Hopcroft-Karp algorithm, which runs in $O(\sqrt{V}E)$ time [32].

Denote $|M^*|$ as the size of the maximum matching in the directed network $G(\mathbf{A})$. We have

Theorem 2 (Minimum Input Theorem)**.**
The minimum number of inputs (N_I) or equivalently the minimum number of driver nodes (N_D) need to fully control a network $G(\mathbf{A})$ is one if there is a perfect matching in $G(\mathbf{A})$. (In this case, any single node can be chosen as the driver node.) Otherwise, it equals the number of unmatched nodes with respective to any maximum matchings. (In this case, the driver nodes are just the

unmatched nodes.)

$$N_{\mathrm{I}} = N_{\mathrm{D}} = \max\{N - |M^*|, 1\}. \tag{S5}$$

Proof. Case 1) If $|M^*| < N$, the maximum matching in $G(\mathbf{A})$ is not perfect and there are $N - |M^*|$ unmatched vertices. The matching edges form elementary paths and circles. We call them matching paths and matching circles hereafter (see Fig. S2c). We can connect an input vertex to each unmatched vertex, and form $N - |M^*|$ stems. All other state vertices are spanned by matching cycles. For a given matching cycle C, if there is an edge e in $G(\mathbf{A})$ such that the starting vertex of e belongs to one of the stems and the ending vertex of e belongs to C, then $e \cup C$ forms a bud. For a matching cycle which can not form a bud in such a way, we can connect one of the input vertices to it and form a bud. In both cases, matching cycles do not require extra inputs to form buds. Thus, we have a disjoint set of cacti with $N - |M^*|$ inputs (origins). According to Lin's theorem, the system is structurally controllable. Since $|M^*|$ is the size of the maximum matching in $G(\mathbf{A})$, $N_{\mathrm{D}} = N - |M^*|$ is therefore the the minimum number of inputs we need.
Case 2) If $|M^*| = N$, then all vertices are spanned by one or more matching cycles. The simplest cactus can be constructed by introducing one input vertex (origin) and connecting it to all cycles to form buds. Modify an arbitrary bud to a stem, we obtain a cactus with one input $N_{\mathrm{D}} = 1$. □

Remarks

1. The above theorem has an intutive explanation. To make sure that each node has its own "superior" so that we can fully control the network, we must inject enough inputs to the system. The minimum inputs we need is determined by the maximum matching of the network. Roughly speaking, a matched node has already been controlled by its "superior", i.e. the node pointing to it. But unmatched node has to be controlled directly by an external "superior" or input. Thus, they are the driver nodes for the whole system. Once we have injected inputs into all the driver nodes, then each node has its own "superior" and the system is fully controllable.

2. Note that when a directed network is strongly connected (i.e. between any pair of distinct nodes i and j there exists a directed path from i to j), and if $N_{\mathrm{D}} = 1$, then there is a directed spanning tree, i.e. a directed tree with at least one root r such that any other nodes can be reached by r along edges in the tree [33]. This is a simple application of Lin's Structural

Controllability Theorem. However, if a strongly connected network has a directed spanning tree, then one cannot claim that $N_{\rm D} = 1$. The reason is that the network can have more than one dilations, which require more than one driver nodes. The existence of dilations has nothing to do with the existence of directed spanning tree(s).

3. According to Lin's theorem, adding more links will never weaken a system's structural controllability (which is not generally true for the usual concept of controllability). This makes the above minimum input theorem meaningful in dealing with missing links, a common occurrence in biological network data. In other words, the minimum input theorem yields an upper bound on the minimum inputs for the network with all missing links added.

IV. MAXIMUM MATCHING

For a random digraph ensemble with a given degree distribution $P(k_{\rm in}, k_{\rm out})$, we derived the average size and number of the maximum matching, using the cavity method developed in statistical physics [34]. (Note that similar results for undirected graph have been obtained using the same method [35].) Then the average number or density of unmatched nodes, i.e. the driver nodes, can be easily obtained.

A. Statistical physics description

Describe a matching M in a digraph $G = \{V(G), E(G)\}$ by the binary variables $s_a = s_{(i\to j)} \in \{0, 1\}$ assigned to each directed edge or arc $a = (i \to j) \in E(G)$ with $s_a = 1$ if $a \in M$ and $s_a = 0$ otherwise. According to the definition of matching in a digraph, one has two constraints on each vertex $i \in V(G)$,

$$\sum_{j\in\partial^+ i} s_{(i\to j)} \ \le \ 1 \tag{S6}$$

$$\sum_{k\in\partial^- i} s_{(k\to i)} \ \le \ 1 \tag{S7}$$

with $\partial^- i$ and $\partial^+ i$ indicating the sets of nodes which point to i or are pointed by i, respectively.

If we define

$$E_i(\{s\}) = 1 - \sum_{k\in\partial^- i} s_{(k\to i)}, \tag{S8}$$

then $E_i(\{s\}) \geq 0$ and it just means whether vertex i is matched ($E_i(\{s\}) = 0$) or unmatched ($E_i(\{s\}) = 1$). Define the cost (or energy) function which gives, for each matching $M = \{s\}$, twice the number of unmatched vertices:

$$E_G(\{s\}) = 2 \sum_{i \in V(G)} E_i(\{s\}) = 2(N - |M|). \tag{S9}$$

We define the Boltzmann probability in the space of matchings as

$$P_G(\{s\}) = \frac{e^{-\beta E_G(\{s\})}}{Z_G(\beta)}, \tag{S10}$$

where β is the inverse temperature and $Z_G(\beta)$ is the partition function:

$$Z_G(\beta) = \sum_{\{s\}} e^{-\beta E_G(\{s\})}. \tag{S11}$$

For $\beta \to \infty$ (zero temperature limit), the internal energy $E_G(\beta)$ gives the ground state property, i.e. the property of the maximum matchings.

B. Internal energy

For each arc $a = (i \to j)$, define two cavity fields: $h_{i\to j}$ and $\widehat{h}_{j\to i}$. They satisfy the recursion relation

$$h_{i\to j} = -\frac{1}{\beta} \log \left(e^{-\beta} + \sum_{k \in \partial^+ i \backslash j} e^{\beta \widehat{h}_{k\to i}} \right) \tag{S12}$$

$$\widehat{h}_{j\to i} = -\frac{1}{\beta} \log \left(e^{-\beta} + \sum_{k \in \partial^- j \backslash i} e^{\beta h_{k\to j}} \right). \tag{S13}$$

The free entropy is given by

$$\begin{aligned} \log Z_G(\beta) &= -\beta F_G(\beta) \\ &= \sum_i \log \left(e^{-\beta} + \sum_{k \in \partial^+ i} e^{\beta \widehat{h}_{k\to i}} \right) + \sum_j \log \left(e^{-\beta} + \sum_{k \in \partial^- j} e^{\beta h_{k\to j}} \right) \\ &\quad - \sum_{(i\to j)} \log \left(1 + e^{\beta(h_{i\to j} + \widehat{h}_{j\to i})} \right). \end{aligned} \tag{S14}$$

Note that this form of free entropy is variational, i.e. the derivatives $\frac{\partial(\beta F(\beta))}{\partial h_{i\to j}}$ and $\frac{\partial(\beta F(\beta))}{\partial \widehat{h}_{j\to i}}$ vanish if and only if the fields $h_{i\to j}$ and $\widehat{h}_{i\to j}$ satisfy the recursion relations Eq. S12 and S13.

The internal energy is given by

$$\begin{aligned} E_G(\beta) &= -\frac{\partial \log Z_G(\beta)}{\partial \beta} \\ &= \sum_i \frac{e^{-\beta} - \sum_{k\in\partial^+ i} \widehat{h}_{k\to i} e^{\beta \widehat{h}_{k\to i}}}{e^{-\beta} + \sum_{k\in\partial^+ i} e^{\beta \widehat{h}_{k\to i}}} + \sum_j \frac{e^{-\beta} - \sum_{k\in\partial^- j} h_{k\to j} e^{\beta h_{k\to j}}}{e^{-\beta} + \sum_{k\in\partial^- j} e^{\beta h_{k\to j}}} \\ &\quad + \sum_{(i\to j)} \frac{(h_{i\to j} + \widehat{h}_{j\to i}) e^{\beta(h_{i\to j} + \widehat{h}_{j\to i})}}{1 + e^{\beta(h_{i\to j} + \widehat{h}_{j\to i})}}. \end{aligned} \tag{S15}$$

Equations (S14, S15) hold on a single digraph G. Now we average over the random digraph ensemble with a prescribed degree distribution $P(k_{\rm in}, k_{\rm out})$. Call $\mathcal{P}_\beta(h)$, $\widehat{\mathcal{P}}_\beta(\widehat{h})$ the distributions of cavity fields along and against all the arcs of a large typical digraph from the digraph ensemble. They satisfy the following self-consistent equations:

$$\begin{aligned} \mathcal{P}_\beta(h) &= \sum_{k_{\rm in}=0}^{\infty} \sum_{k_{\rm out}=1}^{\infty} \frac{k_{\rm out}}{\langle k_{\rm out}\rangle} P(k_{\rm in}, k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}-1} [\mathrm{d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \delta\left[h + \frac{1}{\beta}\log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm out}-1} e^{\beta\widehat{h}_i}\right)\right] \\ &= \sum_{k_{\rm out}=1}^{\infty} Q(k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}-1} [\mathrm{d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \delta\left[h + \frac{1}{\beta}\log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm out}-1} e^{\beta\widehat{h}_i}\right)\right] \end{aligned} \tag{S16}$$

$$\begin{aligned} \widehat{\mathcal{P}}_\beta(\widehat{h}) &= \sum_{k_{\rm in}=1}^{\infty} \sum_{k_{\rm out}=0}^{\infty} \frac{k_{\rm in}}{\langle k_{\rm in}\rangle} P(k_{\rm in}, k_{\rm out}) \int \prod_{i=1}^{k_{\rm in}-1} [\mathrm{d}h_i \mathcal{P}_\beta(h_i)] \delta\left[\widehat{h} + \frac{1}{\beta}\log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm in}-1} e^{\beta h_i}\right)\right] \\ &= \sum_{k_{\rm in}=1}^{\infty} \widehat{Q}(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}-1} [\mathrm{d}h_i \mathcal{P}_\beta(h_i)] \delta\left[\widehat{h} + \frac{1}{\beta}\log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm in}-1} e^{\beta h_i}\right)\right]. \end{aligned} \tag{S17}$$

Here, the term

$$Q(k_{\rm out}) \equiv \frac{k_{\rm out} \sum_{k_{\rm in}=0}^{\infty} P(k_{\rm in}, k_{\rm out})}{\langle k_{\rm out}\rangle} = \frac{k_{\rm out} P(k_{\rm out})}{\langle k_{\rm out}\rangle} \tag{S18}$$

$$\widehat{Q}(k_{\rm in}) \equiv \frac{k_{\rm in} \sum_{k_{\rm out}=0}^{\infty} P(k_{\rm in}, k_{\rm out})}{\langle k_{\rm in}\rangle} = \frac{k_{\rm in} \widehat{P}(k_{\rm in})}{\langle k_{\rm in}\rangle} \tag{S19}$$

are the out- and in- degree distributions of the node i when one picks up uniformly at random an arc $(i \to j)$ from the digraph. This equations for distributions $\mathcal{P}_\beta(h)$ and $\widehat{\mathcal{P}}_\beta(\widehat{h})$ can be solved numerically using the technique of *population dynamics* [34].

The average of the free energy density is given by

$$\begin{aligned} f(\beta) &= \frac{\overline{F_G(\beta)}}{N} \\ &= -\frac{1}{\beta}\sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}} [\mathrm{d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm out}} e^{\beta \widehat{h}_i}\right) \\ &\quad -\frac{1}{\beta}\sum_{k_{\rm in}=0}^{\infty} \widehat{P}(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [\mathrm{d}h_i \mathcal{P}_\beta(h_i)] \log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm in}} e^{\beta h_i}\right) \\ &\quad +\frac{z}{2\beta}\int [\mathrm{d}h\mathrm{d}\widehat{h}\mathcal{P}_\beta(h)\widehat{\mathcal{P}}_\beta(\widehat{h})] \log\left(1+e^{\beta(h+\widehat{h})}\right), \end{aligned} \tag{S20}$$

where $z = \langle k \rangle$ is the mean degree.

Similarly, the average energy density is given by

$$\begin{aligned} \epsilon(\beta) &= \frac{\overline{E_G(\beta)}}{N} \\ &= \sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}} [\mathrm{d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \frac{e^{-\beta} - \sum_{i=1}^{k_{\rm out}} \widehat{h}_i e^{\beta \widehat{h}_i}}{e^{-\beta} + \sum_{i=1}^{k_{\rm out}} e^{\beta \widehat{h}_i}} \\ &\quad + \sum_{k_{\rm in}=0}^{\infty} \widehat{P}(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [\mathrm{d}h_i \mathcal{P}_\beta(h_i)] \frac{e^{-\beta} - \sum_{i=1}^{k_{\rm in}} h_i e^{\beta h_i}}{e^{-\beta} + \sum_{i=1}^{k_{\rm in}} e^{\beta h_i}} \\ &\quad + \frac{z}{2}\int [\mathrm{d}h\mathrm{d}\widehat{h}\mathcal{P}_\beta(h)\widehat{\mathcal{P}}_\beta(\widehat{h})] \frac{(h+\widehat{h})e^{\beta(h+\widehat{h})}}{1+e^{\beta(h+\widehat{h})}}. \end{aligned} \tag{S21}$$

The zero temperature limit ($\beta \to \infty$) corresponds to the ground state (maximum matching) of our system. Numerical studies of Eq. S16, S17 show that for large β the cavity field distribution $\mathcal{P}_\beta(h)$ peaks around three different values $h \in \{-1, 0, 1\}$:

$$\mathcal{P}(h) = w_1\delta(h-1) + w_2\delta(h+1) + w_3\delta(h), \tag{S22}$$

where w_1,w_2, and w_3 are the weights (probabilities) of $h = 1, -1$ and 0. Similarly,

$$\widehat{\mathcal{P}}(\widehat{h}) = \widehat{w}_1\delta(\widehat{h}-1) + \widehat{w}_2\delta(\widehat{h}+1) + \widehat{w}_3\delta(\widehat{h}). \tag{S23}$$

This is called the *energetic zero-temperature limit*. After some algebra, we have

$$\begin{aligned} \mathcal{P}_\beta(h) &= \sum_{k_{\rm out}=1}^{\infty} Q(k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}-1} [\mathrm{d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \delta\left[h + \frac{1}{\beta}\log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm out}-1} e^{\beta \widehat{h}_i}\right)\right] \\ &\overset{\beta\to\infty}{=} \sum_{k_{\rm out}=1}^{\infty} Q(k_{\rm out}) \Big[(1-(1-\widehat{w}_1)^{k_{\rm out}-1})\delta(h+1) + \widehat{w}_2^{k_{\rm out}-1}\delta(h-1) \\ &\quad + ((1-\widehat{w}_1)^{k_{\rm out}-1} - \widehat{w}_2^{k_{\rm out}-1})\delta(h)\Big] \end{aligned} \tag{S24}$$

$$\begin{aligned}\widehat{\mathcal{P}}_\beta(\widehat{h}) &= \sum_{k_{\rm in}=1}^{\infty}\widehat{Q}(k_{\rm in})\int\prod_{i=1}^{k_{\rm in}-1}[{\rm d}h_i\mathcal{P}_\beta(h_i)]\delta\left[\widehat{h}+\frac{1}{\beta}\log\left(e^{-\beta}+\sum_{i=1}^{k_{\rm in}-1}e^{\beta h_i}\right)\right]\\ &\overset{\beta\to\infty}{=} \sum_{k_{\rm in}=1}^{\infty}\widehat{Q}(k_{\rm in})\Big[(1-(1-w_1)^{k_{\rm in}-1})\delta(\widehat{h}+1)+w_2^{k_{\rm in}-1}\delta(\widehat{h}-1)\\ &\quad+((1-w_1)^{k_{\rm in}-1}-w_2^{k_{\rm in}-1})\delta(\widehat{h})\Big]. \qquad \text{(S25)}\end{aligned}$$

Compare Eq. S22 with S24, one has

$$w_1 = \sum_{k_{\rm out}=1}^{\infty}Q(k_{\rm out})\widehat{w}_2^{k_{\rm out}-1}=\sum_{k_{\rm out}=0}^{\infty}Q(k_{\rm out}+1)\widehat{w}_2^{k_{\rm out}}=H(\widehat{w}_2) \qquad \text{(S26)}$$

$$w_2 = \sum_{k_{\rm out}=1}^{\infty}Q(k_{\rm out})[1-(1-\widehat{w}_1)^{k_{\rm out}-1}]=1-H(1-\widehat{w}_1) \qquad \text{(S27)}$$

$$w_3 = \sum_{k_{\rm out}=1}^{\infty}Q(k_{\rm out})[(1-\widehat{w}_1)^{k_{\rm out}-1}-\widehat{w}_2^{k_{\rm out}-1}]=1-w_2-w_1. \qquad \text{(S28)}$$

Compare Eq. S23 with S25, one has

$$\widehat{w}_1 = \sum_{k_{\rm in}=1}^{\infty}\widehat{Q}(k_{\rm in})w_2^{k_{\rm in}-1}=\sum_{k_{\rm in}=0}^{\infty}\widehat{Q}(k_{\rm in}+1)w_2^{k_{\rm in}}=\widehat{H}(w_2) \qquad \text{(S29)}$$

$$\widehat{w}_2 = \sum_{k_{\rm in}=1}^{\infty}\widehat{Q}(k_{\rm in})[1-(1-w_1)^{k_{\rm in}-1}]=1-\widehat{H}(1-w_1) \qquad \text{(S30)}$$

$$\widehat{w}_3 = \sum_{k_{\rm in}=1}^{\infty}\widehat{Q}(k_{\rm in})[(1-w_1)^{k_{\rm in}-1}-w_2^{k_{\rm in}-1}]=1-\widehat{w}_2-\widehat{w}_1. \qquad \text{(S31)}$$

Here we defined the generating functions

$$G(x) = \sum_{k_{\rm out}=0}^{\infty}P(k_{\rm out})x^{k_{\rm out}} \qquad \text{(S32)}$$

$$\widehat{G}(x) = \sum_{k_{\rm in}=0}^{\infty}\widehat{P}(k_{\rm in})x^{k_{\rm in}} \qquad \text{(S33)}$$

$$H(x) = \sum_{k_{\rm out}=0}^{\infty}Q(k_{\rm out}+1)x^{k_{\rm out}} \qquad \text{(S34)}$$

$$\widehat{H}(x) = \sum_{k_{\rm in}=0}^{\infty}\widehat{Q}(k_{\rm in}+1)x^{k_{\rm in}}. \qquad \text{(S35)}$$

In the zero temperature limit, the average energy density can be calculated as follows

$$\begin{aligned}
\epsilon(\beta) &= \sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}} [{\rm d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \frac{e^{-\beta} - \sum_{i=1}^{k_{\rm out}} \widehat{h}_i e^{\beta \widehat{h}_i}}{e^{-\beta} + \sum_{i=1}^{k_{\rm out}} e^{\beta \widehat{h}_i}} \\
&\quad + \sum_{k_{\rm in}=0}^{\infty} \widehat{P}(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [{\rm d}h_i \mathcal{P}_\beta(h_i)] \frac{e^{-\beta} - \sum_{i=1}^{k_{\rm in}} h_i e^{\beta h_i}}{e^{-\beta} + \sum_{i=1}^{k_{\rm in}} e^{\beta h_i}} \\
&\quad + \frac{z}{2} \int [{\rm d}h {\rm d}\widehat{h} \mathcal{P}_\beta(h) \widehat{\mathcal{P}}_\beta(\widehat{h})] \frac{(h+\widehat{h}) e^{\beta(h+\widehat{h})}}{1+e^{\beta(h+\widehat{h})}} \\
&\overset{\beta\to\infty}{=} \sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out}) [\widehat{w}_2^{k_{\rm out}} + ((1-\widehat{w}_1)^{k_{\rm out}} - 1)] + \sum_{k_{\rm in}=0}^{\infty} \widehat{P}(k_{\rm in}) [w_2^{k_{\rm in}} + ((1-w_1)^{k_{\rm in}} - 1)] \\
&\quad + \frac{z}{2} [\widehat{w}_1(1-w_2) + w_1(1-\widehat{w}_2)] \\
&= \big[G(\widehat{w}_2) + G(1-\widehat{w}_1) - 1\big] + \big[\widehat{G}(w_2) + \widehat{G}(1-w_1) - 1\big] \\
&\quad + \frac{z}{2} \big[\widehat{w}_1(1-w_2) + w_1(1-\widehat{w}_2)\big]. \qquad \text{(S36)}
\end{aligned}$$

The minimum density of unmatched nodes or equivalently the minimum density of driver nodes is then given by

$$n_{\rm D} = \frac{1}{2} \left\{ \big[G(\widehat{w}_2) + G(1-\widehat{w}_1) - 1\big] + \big[\widehat{G}(w_2) + \widehat{G}(1-w_1) - 1\big] + \frac{z}{2} \big[\widehat{w}_1(1-w_2) + w_1(1-\widehat{w}_2)\big] \right\}. \qquad \text{(S37)}$$

In the following, we apply the above calculation onto some examples of specific digraphs. Though the calculation can be performed on any given $P(k_{\rm in}, k_{\rm out})$, to simplify the discussion, we present the results obtained from digraphs in which in-degrees and out-degrees share the same distribution, which is of course not always true for real-world networks. We will briefly discuss the results obtained from digraphs where in-degree and out-degree distributions are different.

1. r−regular random digraph

A r-regular graph is a graph without self-edges and multiple edges where each vertex has the same degree r. A r-regular digraph (with r = even) must also satisfy the stronger condition that the in-degree and out-degree of each vertex are equal to each other, i.e.

$$P(k_{\rm out}) = \delta(k_{\rm out} - \frac{r}{2}) = Q(k_{\rm out}) \qquad \text{(S38)}$$

$$\widehat{P}(k_{\rm in}) = \delta(k_{\rm in} - \frac{r}{2}) = \widehat{Q}(k_{\rm in}). \qquad \text{(S39)}$$

In this case, the self-consistent equations of $\mathcal{P}_\beta(h)$, $\widehat{\mathcal{P}}_\beta(\widehat{h})$ can be simplified to

$$\mathcal{P}_\beta(h) = \int \prod_{i=1}^{r/2-1} [\mathrm{d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \delta \left[h + \frac{1}{\beta} \log \left(e^{-\beta} + e^{\beta \widehat{h}_1} + \cdots + e^{\beta \widehat{h}_{r/2-1}} \right) \right] \tag{S40}$$

$$\widehat{\mathcal{P}}_\beta(\widehat{h}) = \int \prod_{i=1}^{r/2-1} [\mathrm{d}h_i \mathcal{P}_\beta(h_i)] \delta \left[\widehat{h} + \frac{1}{\beta} \log \left(e^{-\beta} + e^{\beta h_1} + \cdots + e^{\beta h_{r/2-1}} \right) \right] \tag{S41}$$

with solutions given by

$$\mathcal{P}_\beta(h) = \delta(h - h_r) \tag{S42}$$

$$\widehat{\mathcal{P}}_\beta(\widehat{h}) = \delta(\widehat{h} - \widehat{h}_r). \tag{S43}$$

For the recursion equations of h_r and $\widehat{h}_r$,

$$h_r = -\frac{1}{\beta} \log \left[e^{-\beta} + (r/2-1) e^{\beta \widehat{h}_r} \right] \tag{S44}$$

$$\widehat{h}_r = -\frac{1}{\beta} \log \left[e^{-\beta} + (r/2-1) e^{\beta h_r} \right] \tag{S45}$$

which have obvious solutions $h_r = \widehat{h}_r$. By letting $x = e^{\beta h_r}$ and solving $x^{-1} = e^{-\beta} + (r/2-1)x$, we have

$$h_r = \widehat{h}_r = \frac{1}{\beta} \log \left[\frac{\sqrt{4(r/2-1) + e^{-2\beta}} - e^{-\beta}}{2(r/2-1)} \right]. \tag{S46}$$

The average of the free energy density can be simplified

$$\begin{aligned} f(\beta) &= -\frac{1}{\beta} \log \left(e^{-\beta} + \frac{r}{2} e^{\beta h_r} \right) - \frac{1}{\beta} \log \left(e^{-\beta} + \frac{r}{2} e^{\beta h_r} \right) + \frac{r}{2\beta} \log \left(1 + e^{2\beta h_r} \right) \\ &= -\frac{2}{\beta} \log \left(e^{-\beta} + \frac{r}{2} e^{\beta h_r} \right) + \frac{r}{2\beta} \log \left(1 + e^{2\beta h_r} \right). \end{aligned} \tag{S47}$$

And the average energy density is simplified to

$$\begin{aligned} \epsilon(\beta) &= \frac{e^{-\beta} - \frac{r}{2} h_r e^{\beta h_r}}{e^{-\beta} + \frac{r}{2} e^{\beta h_r}} + \frac{e^{-\beta} - \frac{r}{2} h_r e^{\beta h_r}}{e^{-\beta} + \frac{r}{2} e^{\beta h_r}} + \frac{r h_r e^{2\beta h_r}}{1 + e^{2\beta h_r}} \\ &= 2 \frac{e^{-\beta} - \frac{r}{2} h_r e^{\beta h_r}}{e^{-\beta} + \frac{r}{2} e^{\beta h_r}} + \frac{r h_r e^{2\beta h_r}}{1 + e^{2\beta h_r}}. \end{aligned} \tag{S48}$$

For $r/2 \geq 2$ in the zero temperature limit $\beta \to \infty$ we have $h_r = 0$. The ground state energy density is then $\epsilon = 0$, which means that asymptotically almost all the vertices may be matched for almost every r-regular digraph ($r \geq 4$). For $r/2 = 1$, we know that 2-regular digraph consists of disconnected cycles, which indicates the maximum matching is perfect.

Notice that the result obtained from cavity method is consistent with the graph theoretical result that every regular bipartite graph has a perfect matching [36].

2. *Poisson-distributed digraph*

For directed Erdős-Rényi random networks, in the thermodynamic limit, both $P(k_{\rm in})$ and $P(k_{\rm out})$ follow a Poisson distribution, i.e. $e^{-z_0}z_0^k/k!$ with $z_0 = z/2 = \langle k_{\rm in}\rangle = \langle k_{\rm out}\rangle$ the half mean degree, so we have

$$G(x) = H(x) = \widehat{G}(x) = \widehat{H}(x) = e^{-z_0(1-x)}. \tag{S49}$$

It is easy to check that

$$n_{\rm D} = \frac{1}{2}\epsilon_0 = w_1 - w_2 + z_0 w_1(1-w_2) \tag{S50}$$

and w_1, w_2 satisfy the following equations $w_1 = e^{-z_0(1-w_2)}$, $w_2 = 1 - e^{-z_0 w_1}$. One solves the self-consistent equation

$$w_1 = \exp\left(-z_0 e^{-z_0 w_1}\right) \tag{S51}$$

for w_1 and then $n_{\rm D}$ is obtained (see Fig. S4). Note that $n_{\rm D}$ decreases exponentially as $z_0 = z/2$ increases. The asymptotic behavior of $n_{\rm D}$ in the large z limit can be estimated [37]: As $z \gg 1$, one has $w_1 \sim e^{-z_0}$, $w_2 = 1 - e^{-z_0 w_1} \sim z_0 w_1$, and thus $n_{\rm D} \sim w_1 - z_0 w_1 w_2 \sim e^{-z_0} - z_0^2 e^{-2z_0}$. Therefore, in the large $z = \langle k\rangle$ limit we have

$$n_{\rm D} \sim e^{-z_0} = e^{-\langle k\rangle/2}. \tag{S52}$$

To study the effect of degree heterogeneity on $n_{\rm D}$ for an arbitrary network, we defined the degree heterogeneity as the relative mean difference of its degree distribution,

$$H = \frac{\Delta}{\langle k\rangle} = \frac{\sum_i \sum_j |i-j| P(i)P(j)}{\langle k\rangle} = 2\frac{\sum_i \sum_{j<i} (i-j)P(i)P(j)}{\langle k\rangle}, \tag{S53}$$

where Δ is the average absolute degree difference of two degrees (i and j) drawn from the degree distribution. As a measure of statistical dispersion, H has a well-defined range $[0, 2]$ and is twice of the Gini coefficient of the given distribution. (Note that an often used measure of degree heterogeneity, $\langle k^2\rangle/\langle k\rangle^2$, diverges for scale-free networks with $\gamma \leq 3$ in the $N \to \infty$ limit, thus it is not appropriate to explore the impact of heterogeneity on controllability for scale-free networks.)

For Poisson distribution: $P(k) = \frac{e^{-z}z^k}{k!}$, $k = 0, 1, 2, \cdots$, one can show that

$$H = 2e^{-2z}\left[I_0(2z) + I_1(2z)\right], \tag{S54}$$

where $I_n(x) = \sum_{i=0}^{\infty} \frac{1}{i!(n+i)!}\left(\frac{x}{2}\right)^{2i+n}$ is the modified Bessel function of the first kind. For a general distribution $P(k)$, it is usually hard to get an analytical formula for H. Nevertheless, it can be numerically calculated by using its definition.

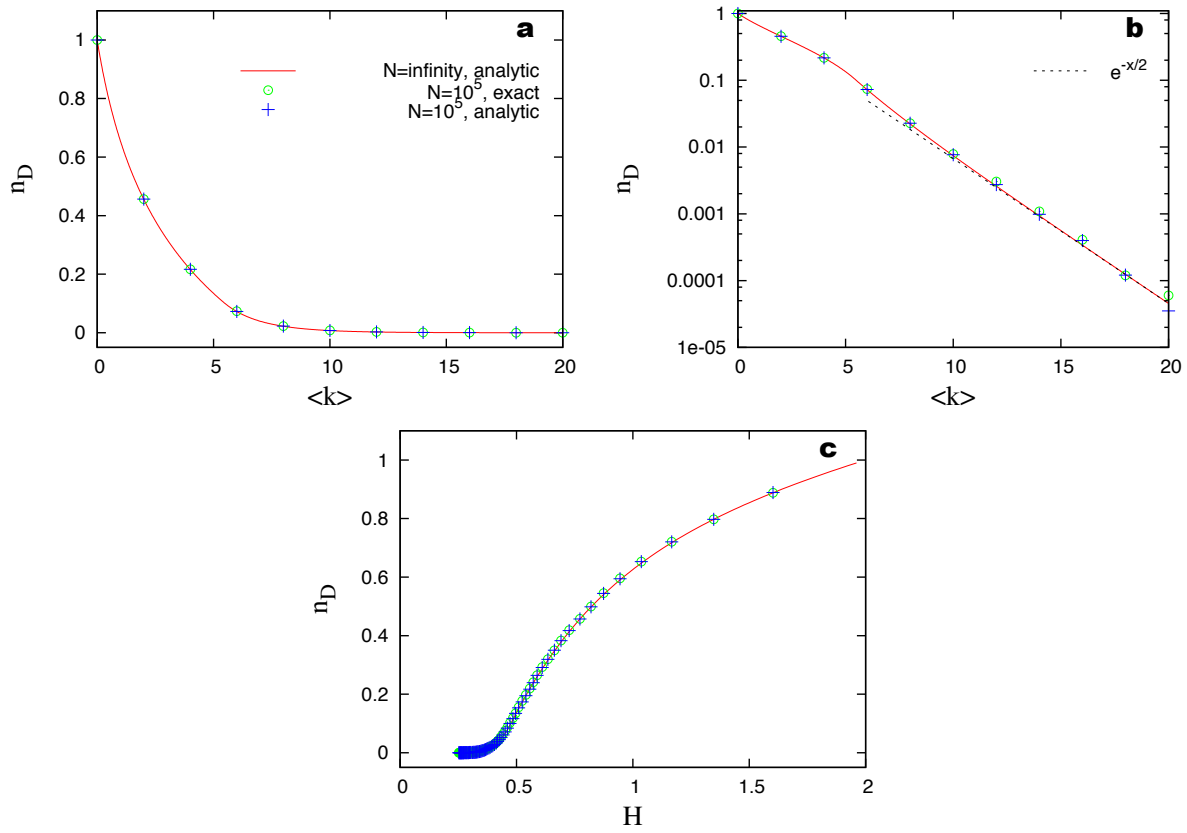

Figure S4: n_{D} **for Poisson-distributed digraphs. a**, n_{D} decreases as $\langle k\rangle$ increases. Lines are analytical results calculated with the cavity method by solving Eq. S51. Symbols are results calculated from finite networks with size $N = 10^5$, constructed by the static model [38]: 'o' is the exact result calculated from the maximum matching algorithm and '+' is the analytical result of the cavity method using the exact degree sequence of the constructed network. **b**, For large $\langle k\rangle$, n_{D} decays exponentially as $n_{\mathrm{D}} \sim e^{-\langle k\rangle/2}$. **c**, Effect of degree heterogeneity H on n_{D}.

3. *Exponentially distributed digraph*

For exponentially distributed digraphs, we assume both $P(k_{\mathrm{in}})$ and $P(k_{\mathrm{out}})$ follow the same exponential distribution, i.e. $(1-e^{-1/\kappa})e^{-k/\kappa}$. Then the half mean degree is given by $z_0 = z/2 = \frac{e^{-1/\kappa}}{1-e^{-1/\kappa}} = \langle k_{\mathrm{in}}\rangle = \langle k_{\mathrm{out}}\rangle$ and the generating functions are

$$G(x) = \widehat{G}(x) = \frac{1-e^{-1/\kappa_0}}{1-xe^{-1/\kappa_0}} \tag{S55}$$

and

$$H(x) = \widehat{H}(x) = \left(\frac{1-e^{-1/\kappa_0}}{1-xe^{-1/\kappa_0}}\right)^2. \tag{S56}$$

The calculated n_{D} is shown in Fig. S5.

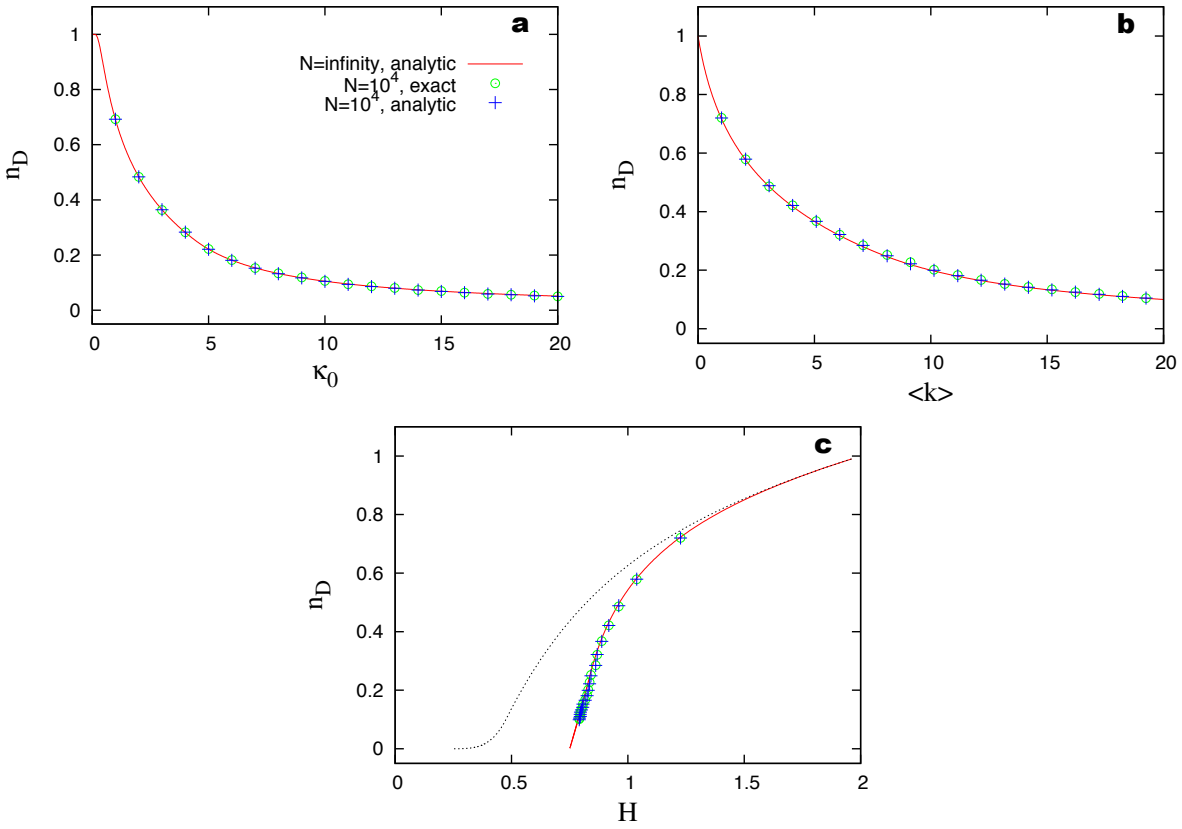

Figure S5: $n_{\rm D}$ **for exponentially distributed digraphs.** **a**, $n_{\rm D}$ decreases as κ_0 increases. Lines are analytical results calculated by the cavity method using Eq. S37. Symbols are results calculated from finite networks with size $N = 10^4$, constructed by the configuration model [39]: 'o' is the exact result calculated from the maximum matching algorithm and '+' is the analytical result of the cavity method using the exact degree sequence of the constructed network. **b**, As $\langle k \rangle$ increases, $n_{\rm D}$ also increases. **c**, Effect of degree heterogeneity H on $n_{\rm D}$. For exponential distribution, we do not have an analytical form for H as a function of κ_0. Instead, it is numerically calculated from the degree distribution. The $n_{\rm D}(H)$ curve for Poisson distributed digraphs is shown (dotted line) just for comparison.

4. Power-law distributed digraphs

Assuming $P(k_{\rm in})$ and $P(k_{\rm out})$ can be described by the same degree distribution with power-law exponent γ and exponential cutoff

$$P(k_{\rm in}) = C\, k_{\rm in}^{-\gamma}\, e^{-k/\kappa} \tag{S57}$$

$$P(k_{\rm out}) = C\, k_{\rm out}^{-\gamma}\, e^{-k/\kappa}. \tag{S58}$$

The normalization constant is given by $C = \left[\mathrm{Li}_\gamma(e^{-k/\kappa})\right]^{-1}$ with $\mathrm{Li}_n(x)$ is the nth polylogarithm of x. Note that with the exponential cutoff $e^{-k/\kappa}$, the distribution is normalizable for any γ.

It is easy to check that

$$G(x) = \widehat{G}(x) = \frac{\mathrm{Li}_\gamma(xe^{-k/\kappa})}{\mathrm{Li}_\gamma(e^{-k/\kappa})} \tag{S59}$$

$$H(x) = \widehat{H}(x) = \frac{\mathrm{Li}_{\gamma-1}(xe^{-k/\kappa})}{x\mathrm{Li}_{\gamma-1}(e^{-k/\kappa})}. \tag{S60}$$

In the $\kappa \to \infty$ limit, one has

$$G(x) = \widetilde{G}(x) = \frac{\mathrm{Li}_\gamma(x)}{\zeta(\gamma)} \tag{S61}$$

$$H(x) = \widetilde{H}(x) = \frac{\mathrm{Li}_{\gamma-1}(x)}{x\zeta(\gamma-1)} \tag{S62}$$

with $\zeta(\gamma) = \sum_{k=1}^{\infty} k^{-\gamma}$ the Riemann Zeta function. Then one can show that as $\gamma \to 2$, $n_\mathrm{D} \to 1$, which means one has to control almost all the nodes to achieve full control over the network. The derivation follows. As $\gamma \to 2$, $H(x) = \widetilde{H}(x) \to 0$ for $0 \le x < 1$. Thus, $w_1 = H(\widehat{w}_2) \to 0$, $\widehat{w}_2 = 1 - \widetilde{H}(1-w_1) \to 1$. Similarly $\widehat{w}_1 \to 0$ and $w_2 \to 1$. So we have

$$\begin{aligned} n_\mathrm{D} &= \frac{1}{2}\left\{\left[G(\widehat{w}_2) + G(1-\widehat{w}_1) - 1\right] + \left[\widetilde{G}(w_2) + \widetilde{G}(1-w_1) - 1\right] + \frac{z}{2}\left[\widehat{w}_1(1-w_2) + w_1(1-\widehat{w}_2)\right]\right\} \\ &\to \left[G(1) + G(1) - 1\right] + \frac{\zeta(\gamma-1)}{\zeta(\gamma)} \cdot \frac{\mathrm{Li}_{\gamma-1}(w_2)}{w_2\zeta(\gamma-1)} \cdot (1-w_2) \\ &\to 1 - \frac{1}{\zeta(2)} \cdot \left[(1-w_2)\ln(1-w_2)\right] \\ &\to 1 \end{aligned} \tag{S63}$$

where we have used the fact that $\mathrm{Li}_1(x) = -\ln(1-x)$ and $\lim_{x\to 0} x\ln x = 0$.

This analytical prediction is confirmed by numerical calculation. One solves the self-consistent equation

$$w_1 = \frac{\mathrm{Li}_{\gamma-1}\left(1 - \frac{\mathrm{Li}_{\gamma-1}(1-w_1)}{(1-w_1)\zeta(\gamma-1)}\right)}{\left[1 - \frac{\mathrm{Li}_{\gamma-1}(1-w_1)}{(1-w_1)\zeta(\gamma-1)}\right]\zeta(\gamma-1)} \tag{S64}$$

for w_1, and similar equation for $\widehat{w}_1$, then n_D is can be calculated from Eq. S37. The result is shown in Fig. S6a. For comparisons, the results for power-law digraphs with finite κ are also shown. It is clearly seen that as $\kappa \to \infty$ and $\gamma \to 2$, $n_\mathrm{D} \to 1$. This suggests that $\gamma = 2$ is a critical value for the controllability of infinite pure scale-free networks. Only for $\gamma > 2$ can we have full controllability by controlling only a subset of nodes.

We remark that, for $\gamma \to 2$, super-hubs will emerge and connect to almost all nodes in the network[40]. We know that for a star-like digraph with one central hub and $N-1$ leaves, one has to control $N_\mathrm{D} = N-1$ nodes (the central hub and any $N-2$ leaves). In the large N limit, $N-1 \approx N$, which explains intuitively why we have to control almost all nodes when $\gamma \to 2$.

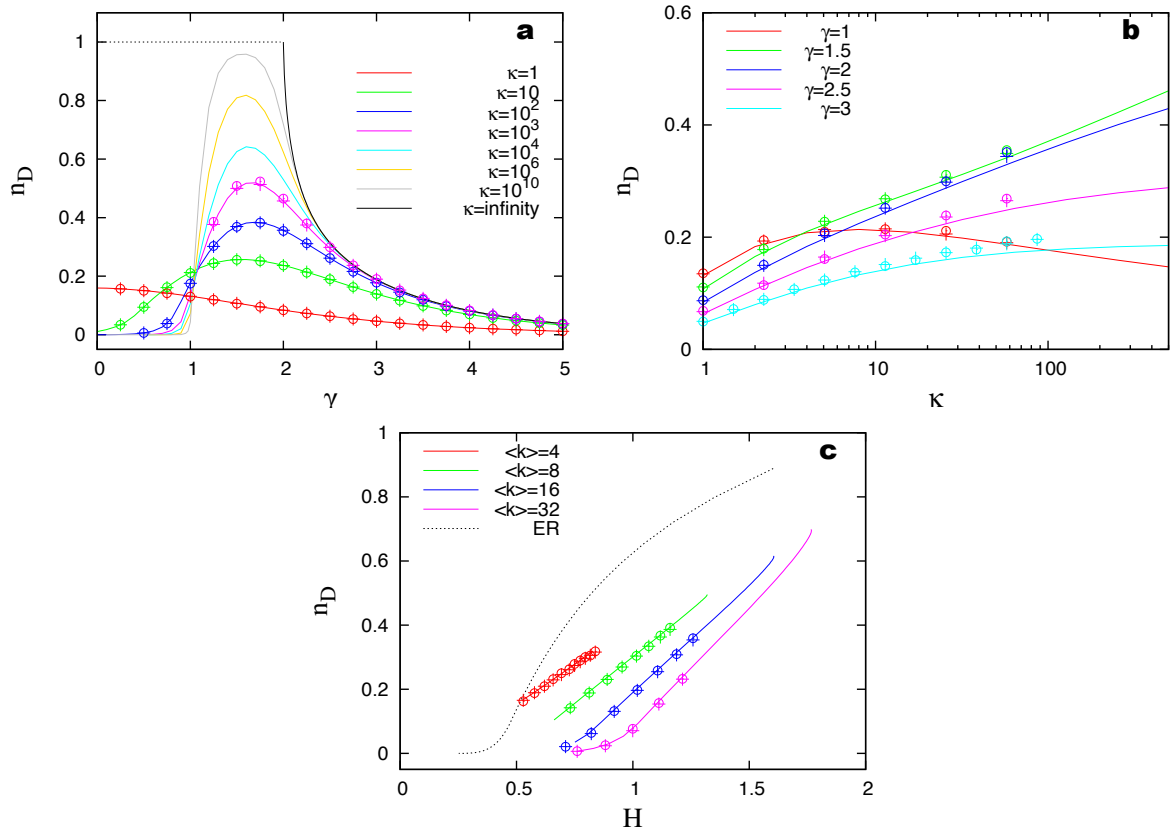

Figure S6: n_D **for power-law distributed digraphs with exponential degree cutoff.** **a**, n_D shows non-monotonic behavior as γ increases at fixed finite κ. At $\gamma = 2$, $n_D \to 1$ as $\kappa \to \infty$. Lines are analytical results calculated with the cavity method using Eq. S57 and S58. Symbols are results calculated from finite networks with size $N = 10^5$ and $\kappa \leq 10^3$, constructed by the configuration model [39]: 'o' is the exact result calculated from the maximum matching algorithm and '+' is the analytical result of the cavity method using the exact degree sequence of the constructed network. **b**, n_D as a function of κ at fixed γ. **c**, n_D as a function of degree heterogeneity H at fixed $\langle k \rangle$. Since $\langle k \rangle = 2\frac{\mathrm{Li}_{\gamma-1}(e^{-1/\kappa})}{\mathrm{Li}_{\gamma}(e^{-1/\kappa})}$, to keep $\langle k \rangle$ fixed, one has to vary κ and γ simultaneously.

5. *Static model*

An often used model to generate static (i.e. not growing) scale-free undirected networks with $\gamma > 2$ is the so-called static model [38]. We start from N disconnected nodes indexed by integer number i ($i = 1, \ldots N$). We assign a weight or expected degree $w_i = c\, i^{-\alpha}$ to each node, with α a real number in the range $[0, 1)$ and c is a constant such that $\sum_{i=1}^{N} w_i = 2E = N\langle k \rangle$. Two different nodes i and j are randomly selected from the set of N vertices, with probability proportional to w_i and w_j, respectively. If they have not been connected, then connect them. Otherwise randomly choose another pair. This process is repeated until $E = mN$ links are created, resulting

in $\langle k\rangle = 2m$. Note that in case $\alpha = 0$, this model is equivalent to the classical Erdős-Rényi random graph model.

Many properties of the static model have been analytically derived [41, 42]. For example, in the thermodynamic limit the degree distribution is given by $P(k) = \frac{[m(1-\alpha)]^{1/\alpha}}{\alpha}\frac{\Gamma(k-1/\alpha, m[1-\alpha])}{\Gamma(k+1)}$ with $\Gamma(s)$ the gamma function and $\Gamma(s,x)$ the upper incomplete gamma function. In the large k limit, $P(k) \sim k^{-(1+\frac{1}{\alpha})} = k^{-\gamma}$ where $\gamma = 1 + \frac{1}{\alpha}$.

Note that this model can be easily generalized to construct directed networks with different $\gamma_{\rm in}$ and $\gamma_{\rm out}$. Here, we assume $\gamma_{\rm in} = \gamma_{\rm out} = \gamma = 1 + \frac{1}{\alpha}$. The generating functions are then given by

$$G(x) = \widetilde{G}(x) = \frac{1}{\alpha}\mathrm{E}_{1+\frac{1}{\alpha}}[(1-x)m(1-\alpha)] \tag{S65}$$

$$H(x) = \widetilde{H}(x) = \frac{1-\alpha}{\alpha}\mathrm{E}_{\frac{1}{\alpha}}[(1-x)m(1-\alpha)]. \tag{S66}$$

Here $\mathrm{E}_n(x) := \int_1^\infty \mathrm{d}y\, e^{-xy}y^{-n}$ is the exponential integral. The asymptotic behavior of $n_{\rm D}(\langle k\rangle, \gamma)$ for large $\langle k\rangle$ can be derived as follows. Consider the self-consistent equations of w_1 and $\widehat{w}_2$:

$$\begin{aligned} w_1 &= H(\widehat{w}_2) = \frac{1-\alpha}{\alpha}\,\mathrm{E}_{\frac{1}{\alpha}}[(1-\widehat{w}_2)m(1-\alpha)] \\ \widehat{w}_2 &= 1 - \widetilde{H}(1-w_1) = 1 - \frac{1-\alpha}{\alpha}\,\mathrm{E}_{\frac{1}{\alpha}}[w_1 m(1-\alpha)]. \end{aligned} \tag{S67}$$

We know that as $m \to \infty$, $w_1 \to 0$, $\widehat{w}_2 \to 0$ and $[(1-\widehat{w}_2)m(1-\alpha)] \to \infty$. Using the asymptotic expansion of $\mathrm{E}_n(x)$ with $x = (1-\widehat{w}_2)m(1-\alpha)$, one can show that $w_1 \approx \frac{e^{-m(1-\alpha)}}{m\alpha} \to 0$ and similarly $\widetilde{G}(w_2) \approx \frac{1}{\alpha}\frac{e^{-m(1-\alpha)}}{m(1-\alpha)} \to 0$. Notice that $mw_1(1-\alpha) \approx \frac{1-\alpha}{\alpha}e^{-m(1-\alpha)} \to 0$, so by using the series expansion of $\mathrm{E}_n(x)$ with $x = mw_1(1-\alpha)$, we have $\widetilde{G}(1-w_1) \approx 1$.

Finally,

$$n_{\rm D} = G(\widehat{w}_2) + \widetilde{G}(1-w_1) - 1 + \frac{z}{2}\,w_1(1-\widehat{w}_2) \approx \frac{e^{-m(1-\alpha)}}{\alpha} \sim e^{-\frac{1}{2}(1-\frac{1}{\gamma-1})\langle k\rangle}, \tag{S68}$$

where we have used the fact that $P(k_{\rm in})$ and $P(k_{\rm out})$ share the same functional form. Notice that in case $\gamma_{\rm in} \neq \gamma_{\rm out}$, one can show that it is the minimum value $\min[\gamma_{\rm in}, \gamma_{\rm out}]$ that determines the asymptotic behavior of $n_{\rm D}$.

Equation (S68) indicates that as $\gamma \to 2$, $\alpha \to 1$, $n_{\rm D} \to 1$, which is consistent with the result that $\gamma_{\rm c} = 2$ obtained from the pure SF network. The asymptotic behavior of $n_{\rm D}$ at large $\langle k\rangle$ is shown in Fig. S7a, b, by numerically solving the self-consistent equations.

We also compared the results calculated from finite networks constructed by the static model with the analytical results obtained from infinite systems. We find that for $\gamma > 3$, they agree well with each other. See Fig. S7c. However, for $\gamma < 3$, especially when $\gamma \to 2$, significant

finite size effect is observed. Moreover, even for the constructed finite network itself, the analytical result calculated from the degree sequence only deviated from the exact result calculated from the maximum matching. These deviations are rooted in the intrinsic degree correlations in the static model when $\gamma < 3$ [41, 42]. This raises a natural question: How will the degree correlation affect n_{D}, i.e. the controllability of a network? This will be systematically explored in a future work.

6. *Chung-Lu model*

The built-in degree correlations in the static model can be eliminated by introducing structural degree cutoff, i.e. $k_{\max} < (N\langle k\rangle)^{\frac{1}{2}}$ [43]. In fact, this is a standard procedure to generate uncorrelated *simple graphs*, i.e. graphs without multiple-edges and self-edges [44]. Here, we discuss the results from the Chung-Lu model, which implicitly considers the degree cutoff and therefore has no degree correlations [45]. Consider random graphs with a given *expected* degree sequence $\mathbf{w} = (w_1, \cdots, w_N)$. The node i is assigned a weight or expected degree w_i. The probability p_{ij} that there is an edge between i and j is proportional to $w_i w_j$, i.e. $p_{ij} = \frac{w_i w_j}{\sum_{k=1}^{N} w_k}$. We assume $\max_i w_i^2 < \sum_k w_k$ to ensure $p_{ij} \leq 1$ for all i and j. Note that this condition is equivalent to set $k_{\max} < (N\langle k\rangle)^{\frac{1}{2}}$, i.e. a structural cutoff is implicitly introduced. A simple way to generate uncorrelated scale-free networks with degree exponent $\gamma = 1 + \frac{1}{\alpha}$ using Chung-Lu model is to assign weight $w_i = c\,(i + i_0 - 1)^{-\alpha}$ to node i with $i = 1, \cdots, N$ and c is a constant such that $\sum_{i=1}^{N} w_i = 2E = N\langle k\rangle$. The constant i_0 is chosen to ensure that $k_{\max} \sim c i_0^{-\alpha} < (N\langle k\rangle)^{\frac{1}{2}}$ such that the degree correlation is eliminated [46]. The construction of SF networks using the Chung-Lu model is then quite similar to the static model, except the introduced constant i_0. Note that for $\gamma > 3$, one simplify chooses $i_0 = 1$, then the two models are exactly the same.

The results of n_{D} for ER and SF networks constructed by the Chung-Lu model are shown in Fig. S8. We find excellent agreement between the exact n_{D} calculated from the maximum matching algorithm and the analytical n_{D} calculated from the cavity method using the exact degree sequence of the constructed network. Both agree with the analytical result calculated from the cavity method using the expected degree sequence only without constructing the network.

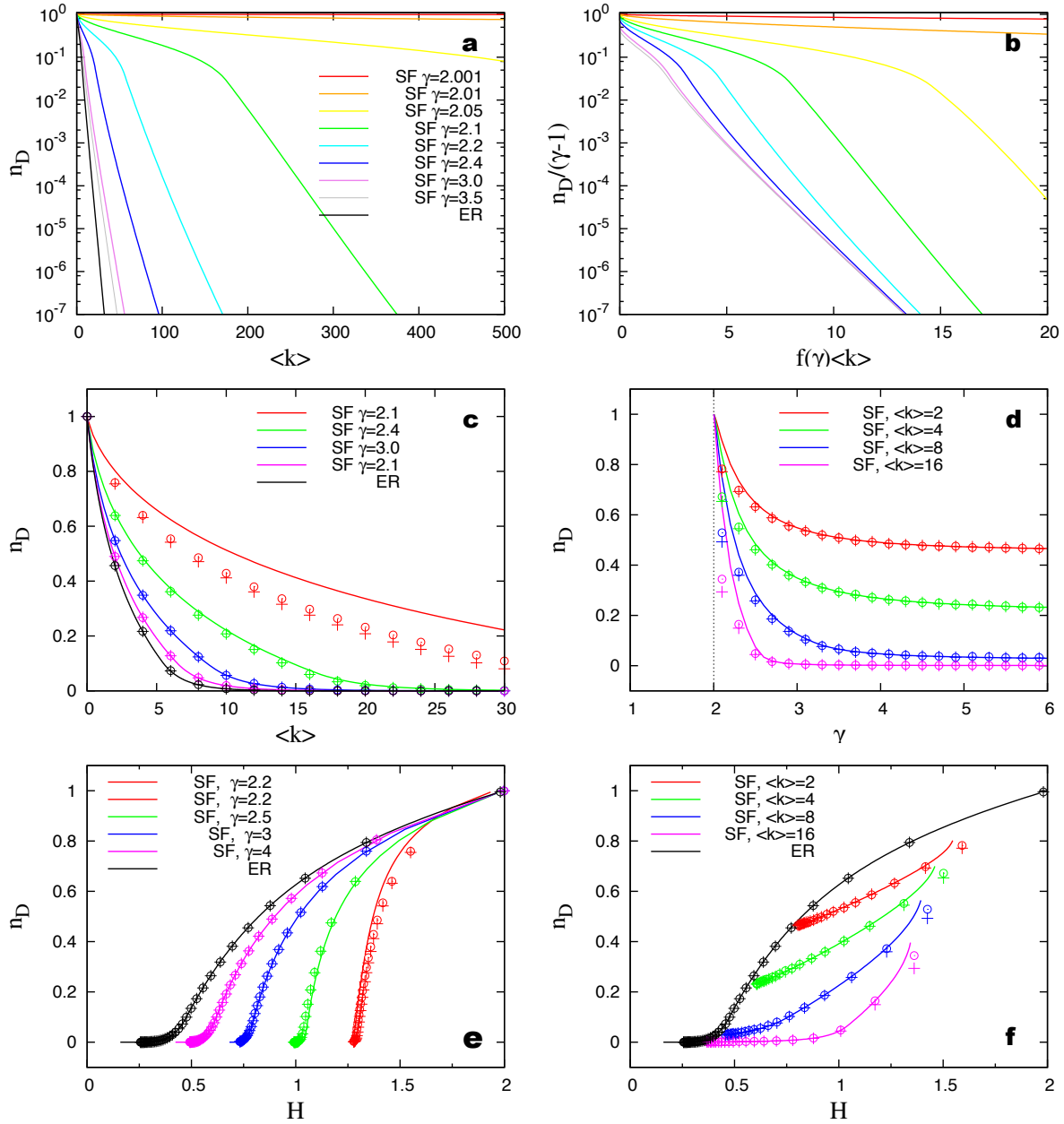

Figure S7: n_D **of the static model.** **a**, n_D in the thermodynamic limit. Those lines are analytical results calculated by the cavity method using Eq. S37. Result for ER network is shown for comparison. **b**, In the thermodynamic limit, scaling behavior of n_D in the large $\langle k\rangle$ limit can be shown by plotting $n_D/(\gamma-1)$ vs. $f(\gamma)\langle k\rangle$ with $f(\gamma) = -\frac{1}{2}(1-\frac{1}{1-\gamma})$. For $\gamma \to 2$, one needs large $\langle k\rangle$ to get the asymptotic behavior as predicted. Only for large γ, the asymptotic behavior can be easily observed. **c**, n_D decreases as $\langle k\rangle$ increases. Lines are analytical results for the infinite system. Symbols are results calculated from finite networks with size $N = 10^5$, constructed by the static model : 'o' is the exact result calculated from the maximum matching algorithm and '+' is the analytical result of the cavity method. Strong finite size effects are seen as $\gamma \to 2$. **d**, n_D increases as γ decreases at fixed $\langle k\rangle$. Again, strong finite size effects are seen as $\gamma \to 2$. **e**, n_D as a function of degree heterogeneity H for ER and SF networks with fixed γ values and variable $\langle k\rangle$. **f**, n_D as a function of degree heterogeneity H for ER and SF networks at fixed $\langle k\rangle$ values and variable γ. As γ increases the curves converge to the ER result (black) at the corresponding $\langle k\rangle$ value.

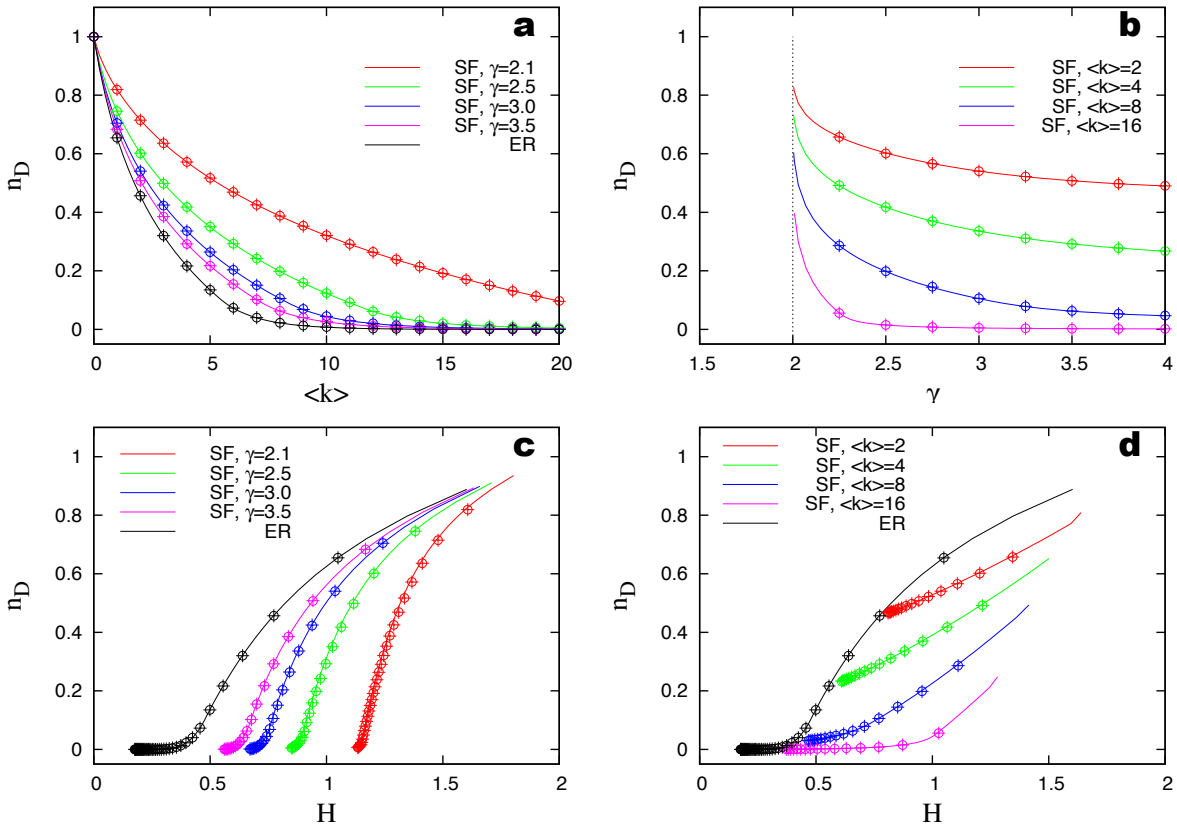

Figure S8: n_D **of the Chung-Lu model. a**, n_D as a function of $\langle k \rangle$ for ER and SF networks with different γ values. Both ER and SF networks are generated from the Chung-Lu model with $N = 10^5$. Lines are analytical results calculated by the cavity method using the expected degree sequence of the Chung-Lu model at the given system size without constructing the network; symbols are results calculated from the constructed network : 'o' is the exact result calculated from the maximum matching algorithm and '+' is the analytical result of the cavity method using the exact degree sequence of the constructed network. **b**, n_D as a function of γ for SF networks at given $\langle k \rangle$. For infinite SF networks, as $\gamma \to \gamma_c = 2$, $n_D \to 1$, i.e. one has to control almost all nodes to fully control the network. For finite SF networks, n_D reaches its maximum as γ approaches 2 at given $\langle k \rangle$. **c**, n_D as a function of degree heterogeneity H for ER and SF networks with fixed γ values and variable $\langle k \rangle$. **d**, n_D as a function of degree heterogeneity H for ER and SF networks at fixed $\langle k \rangle$ values and variable γ. As γ increases the curves converge to the ER result (black) at the corresponding $\langle k \rangle$ value.

C. Entropy

To count the ground states (maximum matchings), we calculate the ground state entropy S_0, which gives the leading exponential behavior of the number of maximum matchings. Here, we again follow the methodology developed in the study of maximum matchings in undirected graph [35]. It is a simple exercise to extend this methodology to the digraph case, as shown below.

The ground state entropy density $s_0 = S_0/N$ can be computed by expanding the free energy density at low temperatures $f(\beta \to \infty) = e_0 - s_0/\beta + \mathcal{O}(1/\beta^2)$ and studying the 'evanescent' parts of the cavity fields [35], i.e. the leading corrections to their value at $\beta = \infty$. Numerically, it has been observed that at $\beta \gg 1$ the three delta peaks in $\mathcal{P}(h)$ keep their weights (w_1, w_2, w_3) and spread as

$$\begin{aligned} h &= 1 + \frac{\log \nu}{\beta} \qquad \text{for the peak around} \quad h = 1 \\ h &= -1 + \frac{\log \mu}{\beta} \qquad \text{for the peak around} \quad h = -1 \\ h &= \frac{\log \gamma}{\beta} \qquad \text{for the peak around} \quad h = 0. \end{aligned} \tag{S69}$$

Similarly, the three delta peaks in $\widehat{\mathcal{P}}(\widehat{h})$ spread as

$$\begin{aligned} \widehat{h} &= 1 + \frac{\log \widehat{\nu}}{\beta} \qquad \text{for the peak around} \quad \widehat{h} = 1 \\ \widehat{h} &= -1 + \frac{\log \widehat{\mu}}{\beta} \qquad \text{for the peak around} \quad \widehat{h} = -1 \\ \widehat{h} &= \frac{\log \widehat{\gamma}}{\beta} \qquad \text{for the peak around} \quad \widehat{h} = 0. \end{aligned} \tag{S70}$$

In the $\beta \to \infty$ limit,

$$\begin{aligned} \mathcal{P}_\beta(h) &= \sum_{k_{\text{out}}=1}^{\infty} Q(k_{\text{out}}) \int \prod_{i=1}^{k_{\text{out}}-1} \left[\mathrm{d}\widehat{h}_i \overline{\widehat{\mathcal{P}}_\beta(\widehat{h}_i; \widehat{\nu}_i, \widehat{\mu}_i, \widehat{\gamma}_i)}^{\widehat{\nu},\widehat{\mu},\widehat{\gamma}} \right] \delta \left[h + \frac{1}{\beta} \log \left(e^{-\beta} + \sum_{i=1}^{k_{\text{out}}-1} e^{\beta \widehat{h}_i} \right) \right] \\ &= \sum_{k_{\text{out}}=1}^{\infty} Q(k_{\text{out}}) \cdot \\ &\quad \int \prod_{i=1}^{k_{\text{out}}-1} \left[\mathrm{d}\widehat{h}_i \left(\widehat{w}_1 \overline{\delta(\widehat{h}_i - 1 - \frac{\log \widehat{\nu}_i}{\beta})}^{\widehat{\nu}} + \widehat{w}_2 \overline{\delta(\widehat{h}_i + 1 - \frac{\log \widehat{\mu}_i}{\beta})}^{\widehat{\mu}} + \widehat{w}_3 \overline{\delta(\widehat{h}_i - \frac{\log \widehat{\gamma}_i}{\beta})}^{\widehat{\gamma}} \right) \right] \\ &\quad \delta \left[h + \frac{1}{\beta} \log \left(e^{-\beta} + \sum_{i=1}^{k_{\text{out}}-1} e^{\beta \widehat{h}_i} \right) \right] \end{aligned} \tag{S71}$$

and

$$\begin{aligned}\widehat{\mathcal{P}}_\beta(\widehat{h}) &= \sum_{k_{\rm in}=1}^{\infty}\widetilde{Q}(k_{\rm in})\int\prod_{i=1}^{k_{\rm in}-1}\left[{\rm d}h_i\overline{\mathcal{P}_\beta(h_i;\nu_i,\mu_i,\gamma_i)}^{\nu,\mu,\gamma}\right]\delta\left[\widehat{h}+\frac{1}{\beta}\log\left(e^{-\beta}+\sum_{i=1}^{k_{\rm in}-1}e^{\beta h_i}\right)\right]\\ &= \sum_{k_{\rm in}=1}^{\infty}\widetilde{Q}(k_{\rm in})\cdot\\ &\int\prod_{i=1}^{k_{\rm in}-1}\left[{\rm d}h_i\left(w_1\overline{\delta(h_i-1-\frac{\log\nu_i}{\beta})}^{\nu}+w_2\overline{\delta(h_i+1-\frac{\log\mu_i}{\beta})}^{\mu}+w_3\overline{\delta(h_i-\frac{\log\gamma_i}{\beta})}^{\gamma}\right)\right]\\ &\delta\left[\widehat{h}+\frac{1}{\beta}\log\left(e^{-\beta}+\sum_{i=1}^{k_{\rm in}-1}e^{\beta h_i}\right)\right].\end{aligned}\tag{S72}$$

This is called the *entropic zero-temperature limit*. Note that hereafter, the overline and supscripts denote expectations over independent random variables with distribution $\mathcal{A}_1$ (for ν-variables), $\mathcal{A}_2$ (for μ-variables), $\mathcal{A}_3$ (for γ-variables); and $\widehat{\mathcal{A}}_1$ (for $\widehat{\nu}$-variables), $\widehat{\mathcal{A}}_2$ (for $\widehat{\mu}$-variables), $\widehat{\mathcal{A}}_3$ (for $\widehat{\gamma}$-variables). If confusions cannot occur, subscripts will be neglected.

From the above self-consistent equations of $\mathcal{P}(h)$ and $\widehat{\mathcal{P}}(\widehat{h})$, we get the self-consistent equations satisfied by $\mathcal{A}_1(\nu)$, $\mathcal{A}_2(\mu)$, $\mathcal{A}_3(\gamma)$ and $\widehat{\mathcal{A}}_1(\widehat{\nu})$, $\widehat{\mathcal{A}}_2(\widehat{\mu})$, $\widehat{\mathcal{A}}_3(\widehat{\gamma})$.

$$\mathcal{A}_1(\nu) = \sum_{k_{\rm out}=0}^{\infty}\mathcal{C}_1(k_{\rm out})\int\prod_{i=1}^{k_{\rm out}}[{\rm d}\widehat{\mu}_i\widehat{\mathcal{A}}_2(\widehat{\mu}_i)]\,\delta\left(\nu-\frac{1}{1+\sum_{i=1}^{k_{\rm out}}\widehat{\mu}_i}\right)\tag{S73}$$

$$\mathcal{A}_2(\mu) = \sum_{k_{\rm out}=1}^{\infty}\mathcal{C}_2(k_{\rm out})\int\prod_{i=1}^{k_{\rm out}}[{\rm d}\widehat{\nu}_i\widehat{\mathcal{A}}_1(\widehat{\nu}_i)]\,\delta\left(\mu-\frac{1}{\sum_{i=1}^{k_{\rm out}}\widehat{\nu}_i}\right)\tag{S74}$$

$$\mathcal{A}_3(\gamma) = \sum_{k_{\rm out}=1}^{\infty}\mathcal{C}_3(k_{\rm out})\int\prod_{i=1}^{k_{\rm out}}[{\rm d}\widehat{\gamma}_i\widehat{\mathcal{A}}_3(\widehat{\gamma}_i)]\,\delta\left(\gamma-\frac{1}{\sum_{i=1}^{k_{\rm out}}\widehat{\gamma}_i}\right)\tag{S75}$$

with

$$\mathcal{C}_1(k_{\rm out}) = \frac{1}{w_1}\cdot\widehat{w}_2^{k_{\rm out}}Q(k_{\rm out}+1)\tag{S76}$$

$$\mathcal{C}_2(k_{\rm out}) = \frac{1}{w_2}\cdot\sum_{m=k_{\rm out}}^{\infty}\binom{m}{k_{\rm out}}\widehat{w}_1^{k_{\rm out}}\,(\widehat{w}_2+\widehat{w}_3)^{m-k_{\rm out}}Q(m+1)\tag{S77}$$

$$\mathcal{C}_3(k_{\rm out}) = \frac{1}{w_3}\cdot\sum_{m=k_{\rm out}}^{\infty}\binom{m}{k_{\rm out}}\widehat{w}_3^{k_{\rm out}}\,\widehat{w}_2^{m-k_{\rm out}}Q(m+1)\tag{S78}$$

and

$$\widehat{\mathcal{A}}_1(\widehat{\nu}) = \sum_{k_{\rm in}=0}^{\infty} \widehat{\mathcal{C}}_1(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [{\rm d}\mu_i \mathcal{A}_2(\mu_i)]\, \delta\left(\widehat{\nu} - \frac{1}{1+\sum_{i=1}^{k_{\rm in}} \mu_i}\right) \tag{S79}$$

$$\widehat{\mathcal{A}}_2(\widehat{\mu}) = \sum_{k_{\rm in}=1}^{\infty} \widehat{\mathcal{C}}_2(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [{\rm d}\nu_i \mathcal{A}_1(\nu_i)]\, \delta\left(\widehat{\mu} - \frac{1}{\sum_{i=1}^{k_{\rm in}} \nu_i}\right) \tag{S80}$$

$$\widehat{\mathcal{A}}_3(\widehat{\gamma}) = \sum_{k_{\rm in}=1}^{\infty} \widehat{\mathcal{C}}_3(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [{\rm d}\gamma_i \mathcal{A}_3(\gamma_i)]\, \delta\left(\widehat{\gamma} - \frac{1}{\sum_{i=1}^{k_{\rm in}} \gamma_i}\right) \tag{S81}$$

with

$$\widehat{\mathcal{C}}_1(k_{\rm in}) = \frac{1}{\widehat{w}_1} \cdot w_2^{k_{\rm in}} \widetilde{Q}(k_{\rm in}+1) \tag{S82}$$

$$\widehat{\mathcal{C}}_2(k_{\rm in}) = \frac{1}{\widehat{w}_2} \cdot \sum_{m=k_{\rm in}}^{\infty} \binom{m}{k_{\rm in}} w_1^{k_{\rm in}} (w_2+w_3)^{m-k_{\rm in}} \widetilde{Q}(m+1) \tag{S83}$$

$$\widehat{\mathcal{C}}_3(k_{\rm in}) = \frac{1}{\widehat{w}_3} \cdot \sum_{m=k_{\rm in}}^{\infty} \binom{m}{k_{\rm in}} w_3^{k_{\rm in}} w_2^{m-k_{\rm in}} \widetilde{Q}(m+1). \tag{S84}$$

Using the above equations we can expand the free energy to order $1/\beta$ and get the ground state entropy density of maximum matchings. (As a byproduct, the ground state energy is also obtained, which should be equal to the result obtained from the calculation in the energetic zero-temperature limit.)

$$\begin{aligned} f(\beta) &= \frac{\overline{F_G(\beta)}}{N} \\ &= -\frac{1}{\beta} \sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}} [{\rm d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm out}} e^{\beta \widehat{h}_i}\right) \\ &\quad -\frac{1}{\beta} \sum_{k_{\rm in}=0}^{\infty} \widetilde{P}(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [{\rm d}h_i \mathcal{P}_\beta(h_i)] \log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm in}} e^{\beta h_i}\right) \\ &\quad +\frac{z}{2\beta} \int [{\rm d}h {\rm d}\widehat{h} \mathcal{P}_\beta(h) \widehat{\mathcal{P}}_\beta(\widehat{h})] \log\left(1+e^{\beta(h+\widehat{h})}\right) \\ &\equiv [1]+[2]+[3]. \end{aligned} \tag{S85}$$

In the $\beta \to \infty$ limit, one has

$$\mathcal{P}_\beta(h) = w_1 \overline{\delta(h-1-\frac{\log \nu}{\beta})}^{\nu} + w_2 \overline{\delta(h+1-\frac{\log \mu}{\beta})}^{\mu} + w_3 \overline{\delta(h-\frac{\log \gamma}{\beta})}^{\gamma} \tag{S86}$$

$$\widehat{\mathcal{P}}_\beta(\widehat{h}) = \widehat{w}_1 \overline{\delta(\widehat{h}-1-\frac{\log \widehat{\nu}}{\beta})}^{\widehat{\nu}} + \widehat{w}_2 \overline{\delta(\widehat{h}+1-\frac{\log \widehat{\mu}}{\beta})}^{\widehat{\mu}} + \widehat{w}_3 \overline{\delta(\widehat{h}-\frac{\log \widehat{\gamma}}{\beta})}^{\widehat{\gamma}}. \tag{S87}$$

Consider the first term in the free energy density, one has

$$\begin{aligned}
[1] &= -\frac{1}{\beta}\sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out}) \int \prod_{i=1}^{k_{\rm out}} [\mathrm{d}\widehat{h}_i \widehat{\mathcal{P}}_\beta(\widehat{h}_i)] \log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm out}} e^{\beta \widehat{h}_i}\right) \\
&= -1 + G(\widehat{w}_2 + \widehat{w}_3) + G(\widehat{w}_2) \\
&\quad -\frac{1}{\beta}\Bigg[\sum_{k_{\rm out}=1}^{\infty}\sum_{m=k_{\rm out}}^{\infty} \binom{m}{k_{\rm out}} \widehat{w}_1^{k_{\rm out}} (1-\widehat{w}_1)^{m-k_{\rm out}} P(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm out}} \widehat{\nu}_i} \\
&\quad + \sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out})\, \widehat{w}_2^{k_{\rm out}} \cdot \overline{\log(1+\sum_{i=1}^{k_{\rm out}} \widehat{\mu}_i)} \\
&\quad + \sum_{k_{\rm out}=1}^{\infty}\sum_{m=k_{\rm out}}^{\infty} \binom{m}{k_{\rm out}} \widehat{w}_3^{k_{\rm out}} \widehat{w}_2^{m-k_{\rm out}} P(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm out}} \widehat{\gamma}_i} \Bigg].
\end{aligned} \tag{S88}$$

Similarly, for the second term, one has

$$\begin{aligned}
[2] &= -\frac{1}{\beta}\sum_{k_{\rm in}=0}^{\infty} \widetilde{P}(k_{\rm in}) \int \prod_{i=1}^{k_{\rm in}} [\mathrm{d}h_i \mathcal{P}_\beta(h_i)] \log\left(e^{-\beta} + \sum_{i=1}^{k_{\rm in}} e^{\beta h_i}\right) \\
&= -1 + \widetilde{G}(w_2 + w_3) + \widetilde{G}(w_2) \\
&\quad -\frac{1}{\beta}\Bigg[\sum_{k_{\rm in}=1}^{\infty}\sum_{m=k_{\rm in}}^{\infty} \binom{m}{k_{\rm in}} w_1^{k_{\rm in}} (1-w_1)^{m-k_{\rm in}} \widetilde{P}(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm in}} \nu_i} \\
&\quad + \sum_{k_{\rm in}=0}^{\infty} \widetilde{P}(k_{\rm in})\, w_2^{k_{\rm in}} \cdot \overline{\log(1+\sum_{i=1}^{k_{\rm in}} \mu_i)} \\
&\quad + \sum_{k_{\rm in}=1}^{\infty}\sum_{m=k_{\rm in}}^{\infty} \binom{m}{k_{\rm in}} w_3^{k_{\rm in}} w_2^{m-k_{\rm in}} \widetilde{P}(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm in}} \gamma_i} \Bigg].
\end{aligned} \tag{S89}$$

For the third term, one has

$$\begin{aligned}
[3] &= \frac{z}{2\beta} \int [\mathrm{d}h \mathrm{d}\widehat{h} \mathcal{P}_\beta(h) \widehat{\mathcal{P}}_\beta(\widehat{h})] \log\left(1 + e^{\beta(h+\widehat{h})}\right) \\
&= z\widehat{w}_1 w_1 + \frac{z}{2}(\widehat{w}_1 w_3 + w_1 \widehat{w}_3) \\
&\quad -\frac{1}{\beta}\Big\{ -\frac{z}{2}\widehat{w}_1(w_1+w_3)\,\overline{\log \widehat{\nu}} - \frac{z}{2}(\widehat{w}_1+\widehat{w}_3) w_1\, \overline{\log \nu} \\
&\quad -\frac{z}{2}\widehat{w}_1 w_2\, \overline{\log(1+\hat{\nu}\mu)} - \frac{z}{2} w_1 \widehat{w}_2\, \overline{\log(1+\hat{\mu}\nu)} \\
&\quad -\frac{z}{2}\widehat{w}_1 w_3\, \overline{\log \gamma} - \frac{z}{2} w_1 \widehat{w}_3\, \overline{\log \widehat{\gamma}} - \frac{z}{2} w_3 \widehat{w}_3\, \overline{\log(1+\widehat{\gamma}\gamma)} \Big\}.
\end{aligned} \tag{S90}$$

Since

$$f(\beta \to \infty) = e_0 - \frac{s_0}{\beta} + \mathcal{O}(1/\beta^2), \tag{S91}$$

from Eq. S88, S89, and S90, one recognizes that the ground state energy density

$$\begin{aligned} e_0 &= [-1 + G(\widehat{w}_2 + \widehat{w}_3)] + G(\widehat{w}_2)] + [-1 + \widetilde{G}(w_2 + w_3) + \widetilde{G}(w_2)] + \frac{z}{2}[\widehat{w}_1(1 - w_2) + w_1(1 - \widehat{w}_2)] \\ &= [G(\widehat{w}_2) + G(1 - \widehat{w}_1) - 1] + [\widetilde{G}(w_2) + \widetilde{G}(1 - w_1) - 1] + \frac{z}{2}[\widehat{w}_1(1 - w_2) + w_1(1 - \widehat{w}_2)] \end{aligned} \quad \text{(S92)}$$

and the ground state entropy density

$$\begin{aligned} s_0 = & \left[\sum_{k_{\rm out}=1}^{\infty} \sum_{m=k_{\rm out}}^{\infty} \binom{m}{k_{\rm out}} \widehat{w}_1^{k_{\rm out}} (1 - \widehat{w}_1)^{m-k_{\rm out}} P(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm out}} \widehat{\nu}_i} \right. \\ & + \sum_{k_{\rm out}=0}^{\infty} P(k_{\rm out})\, \widehat{w}_2^{k_{\rm out}} \cdot \overline{\log(1 + \sum_{i=1}^{k_{\rm out}} \widehat{\mu}_i)} \\ & \left. + \sum_{k_{\rm out}=1}^{\infty} \sum_{m=k_{\rm out}}^{\infty} \binom{m}{k_{\rm out}} \widehat{w}_3^{k_{\rm out}} \widehat{w}_2^{m-k_{\rm out}} P(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm out}} \widehat{\gamma}_i} \right] \\ & + \left[\sum_{k_{\rm in}=1}^{\infty} \sum_{m=k_{\rm in}}^{\infty} \binom{m}{k_{\rm in}} w_1^{k_{\rm in}} (1 - w_1)^{m-k_{\rm in}} \widetilde{P}(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm in}} \nu_i} \right. \\ & + \sum_{k_{\rm in}=0}^{\infty} \widetilde{P}(k_{\rm in})\, w_2^{k_{\rm in}} \cdot \overline{\log(1 + \sum_{i=1}^{k_{\rm in}} \mu_i)} \\ & \left. + \sum_{k_{\rm in}=1}^{\infty} \sum_{m=k_{\rm in}}^{\infty} \binom{m}{k_{\rm in}} w_3^{k_{\rm in}} w_2^{m-k_{\rm in}} \widetilde{P}(m) \cdot \overline{\log \sum_{i=1}^{k_{\rm in}} \gamma_i} \right] \\ & + \left[-\frac{z}{2} \widehat{w}_1 (w_1 + w_3)\, \overline{\log \widehat{\nu}} - \frac{z}{2} (\widehat{w}_1 + \widehat{w}_3) w_1\, \overline{\log \nu} \right. \\ & -\frac{z}{2} \widehat{w}_1 w_2\, \overline{\log(1 + \hat{\nu}\mu)} - \frac{z}{2} w_1 \widehat{w}_2\, \overline{\log(1 + \hat{\mu}\nu)} \\ & \left. -\frac{z}{2} \widehat{w}_1 w_3\, \overline{\log \gamma} - \frac{z}{2} w_1 \widehat{w}_3\, \overline{\log \widehat{\gamma}} - \frac{z}{2} w_3 \widehat{w}_3\, \overline{\log(1 + \widehat{\gamma}\gamma)} \right]. \end{aligned} \quad \text{(S93)}$$

Note that e_0 is exact what we obtained from the direct calculation of e_0 in the energetic zero-temperature limit, i.e. Eq. S36.

Fig. S9 shows the ground state entropy density s_0 for various model networks. Similar to the undirected case [35], we find that s_0 is not monotonic with $\langle k \rangle$ (Fig. S9a, c): it develops a nontrivial local minimum around $\langle k \rangle \sim \langle k \rangle_{\rm c}$ for both ER and SF networks. This complex behavior results from the competition between entropy contributions of two topologically distinct regions of a network, the *core* and *leaves*[35]. The core represents a compact cluster of nodes left in the network after applying a *greedy leaf removal* procedure[47]. While leaves are nodes with $k_{\rm in} = 1$ or $k_{\rm out} = 1$ before or during the greedy leaf removal. The core emerges through a percolation process (Fig. S9b, d): for $k < \langle k \rangle_{\rm c}$, $n_{\rm core} = N_{\rm core}/N = 0$, so only the leaves contribute to s_0, their

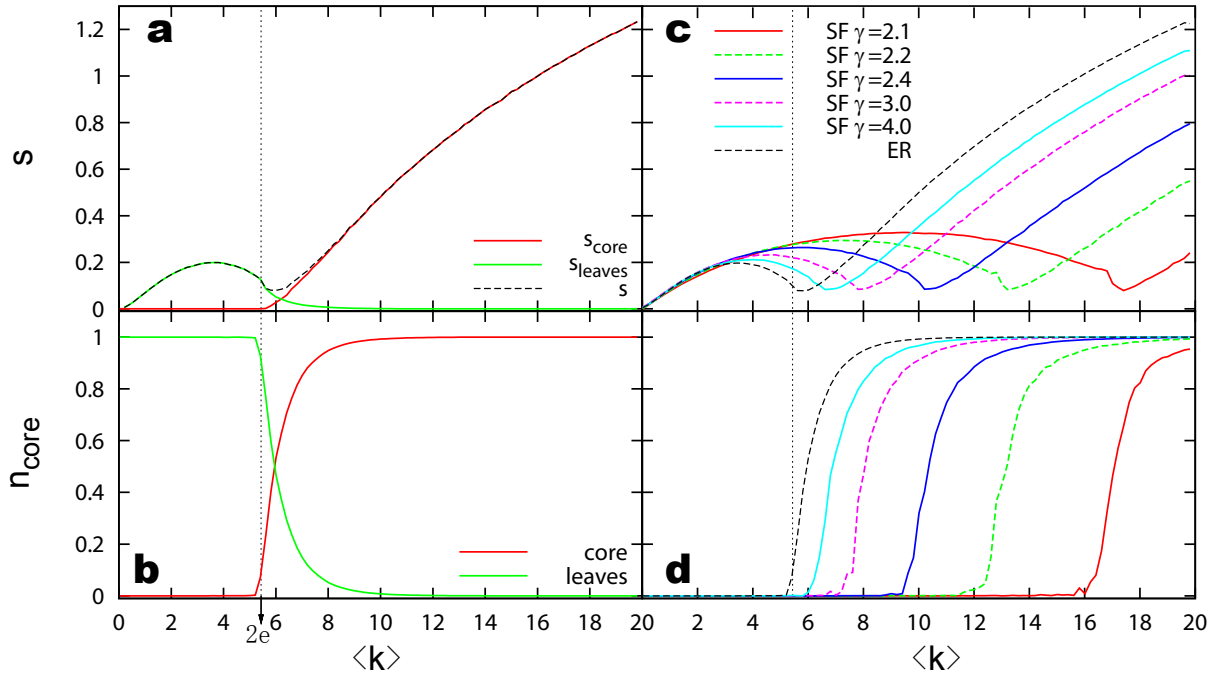

Figure S9: **Entropy density of maximum matchings and the core percolation of random networks.** (a,b) Erdős-Rényi (ER). (c,d) ER and scale-free (SF). (a) The entropy density of maximum matchings (dashed line) has two contributions due to leaves (green line) and the core (red line): $s = s_{\rm leaves} + s_{\rm core}$. (b) For ER, the core percolation occurs at $k = \langle k\rangle_{\rm c} = 2e$, which explains the entropy valley developed around $\langle k\rangle_{\rm c}$. (c,d) For SF random networks with smaller γ, the core percolation is shifted to higher $\langle k\rangle_{\rm c}$ value. Both the ER and SF networks are generated from the Chung-Lu model [45] with $N = 10^4$.

contribution reaching a maximum value at some $\langle k\rangle < \langle k\rangle_{\rm c}$. As $\langle k\rangle$ increases beyond $\langle k\rangle_{\rm c}$, the core emerges and the number of leaves decreases. For ER network, $\langle k\rangle_{\rm c} = 2\,e \approx 5.436564$ (see Fig. S9b). For SF networks the overall behavior remains unchanged, s_0 also develops a significant valley but $\langle k\rangle_{\rm c}$ decreases with γ (see Fig. S9d).

V. CONTROL ROBUSTNESS

In the main text, we classify each link into three different types according to its role in maintaining controllability: (1) *critical* if in its absence we need to increase the number of driver nodes to maintain full control of the system; (2) *redundant* if it can be removed without affecting the current set of driver nodes; (3) *ordinary* if it is neither critical nor redundant, i.e., its removal can eliminate some control configurations, but the network can still be controllable in its absence with

the same number of driver nodes N_D. Note that, in engineered systems, this classification implies that in order to make controllability more robust, one can simply double (or triple) each critical link, formally making each of these links redundant and therefore ensuring there is no critical link in the controlled system.

Now, we show that nodes can be classified similarly. Given a network, denote the minimum number of driver nodes as N_D. After a node is removed, denote the minimum number of driver nodes as N'_D. Then we classify each node into one of the following three categories: (1) A node is *critical* if in its absence we have to control more driver nodes, i.e. $N'_D > N_D$. For example, remove one node in the middle of a directed path will cause N_D increase. (2) A node is *redundant* if in its absence we have $N'_D < N_D$. For example, remove one leaf node in a star will decrease N_D by 1. (3) A node is *ordinary* if in its absence we have $N'_D = N_D$. For example, remove the central hub in a star will not change N_D at all.

Denote the density of critical, redundant, and ordinary nodes as $n_c = N_c/N$, $n_r = N_r/N$, and $n_o = N_o/N$, respectively. In Fig. S10 we show n_c, n_r, and n_o for all the studied real networks. For comparison, the result on link-category is also shown. In general, most networks have few or no critical nodes and considerable number of redundant nodes, suggesting that networks are relatively stable against node failure.

The $\langle k \rangle$ dependent n_c, n_r and n_o for ER and SF networks are shown in Fig. S11b, e, respectively. The overall trend looks similar to that of l_c, l_r and l_o (Fig. S11a, d). And it is also related to the core percolation (Fig. S11c, f). Surprisingly, for ER networks we find that $l_r \approx n_c$ and $l_c \approx n_r$ for any $\langle k \rangle$. For SF networks, this does not hold.

VI. NETWORK DATASETS

All the real-world networks analyzed in the paper are listed and briefly described in Table S1.

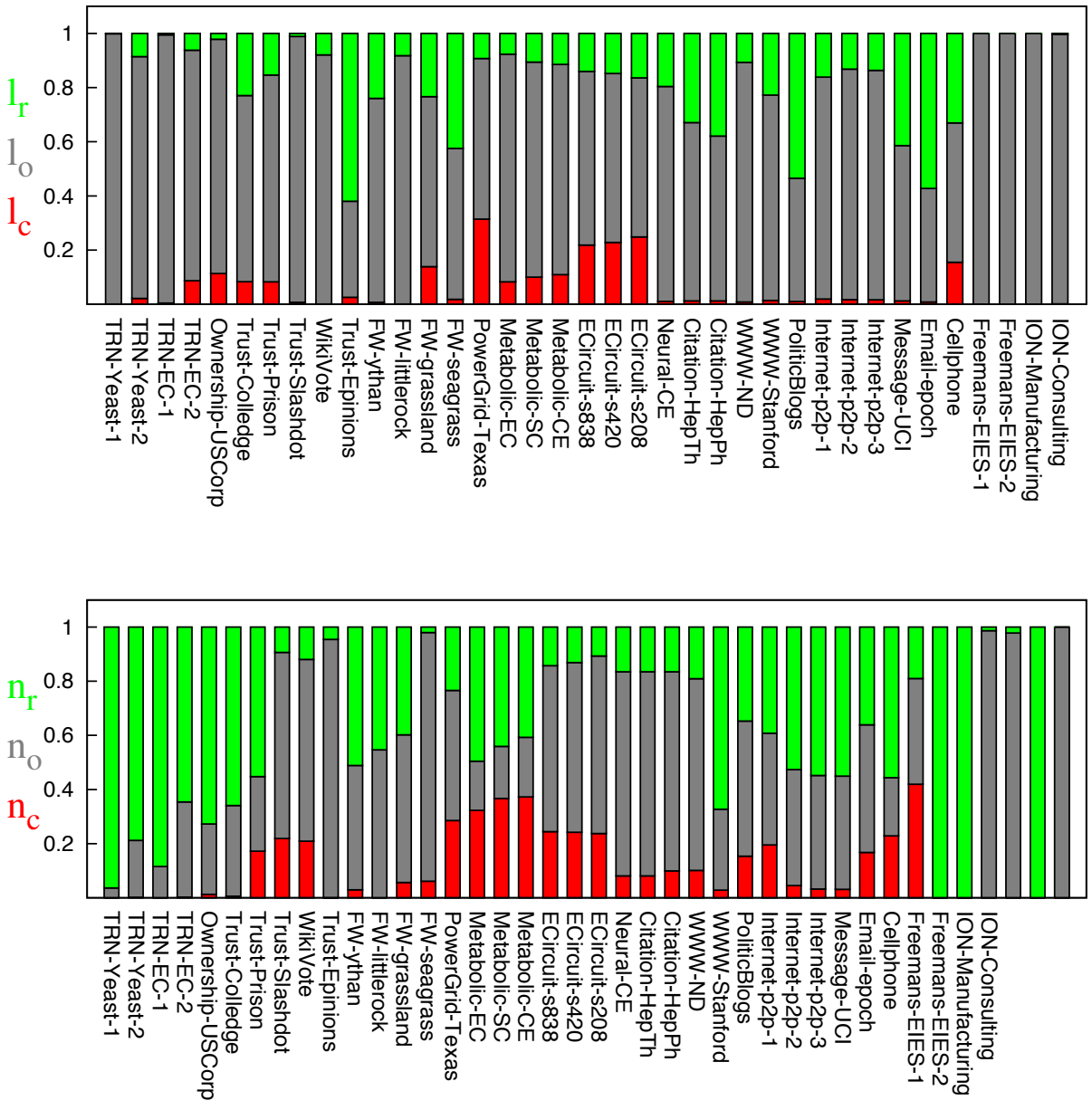

Figure S10: **Link- and node-categories for real networks listed in Table 1 of the main text.**

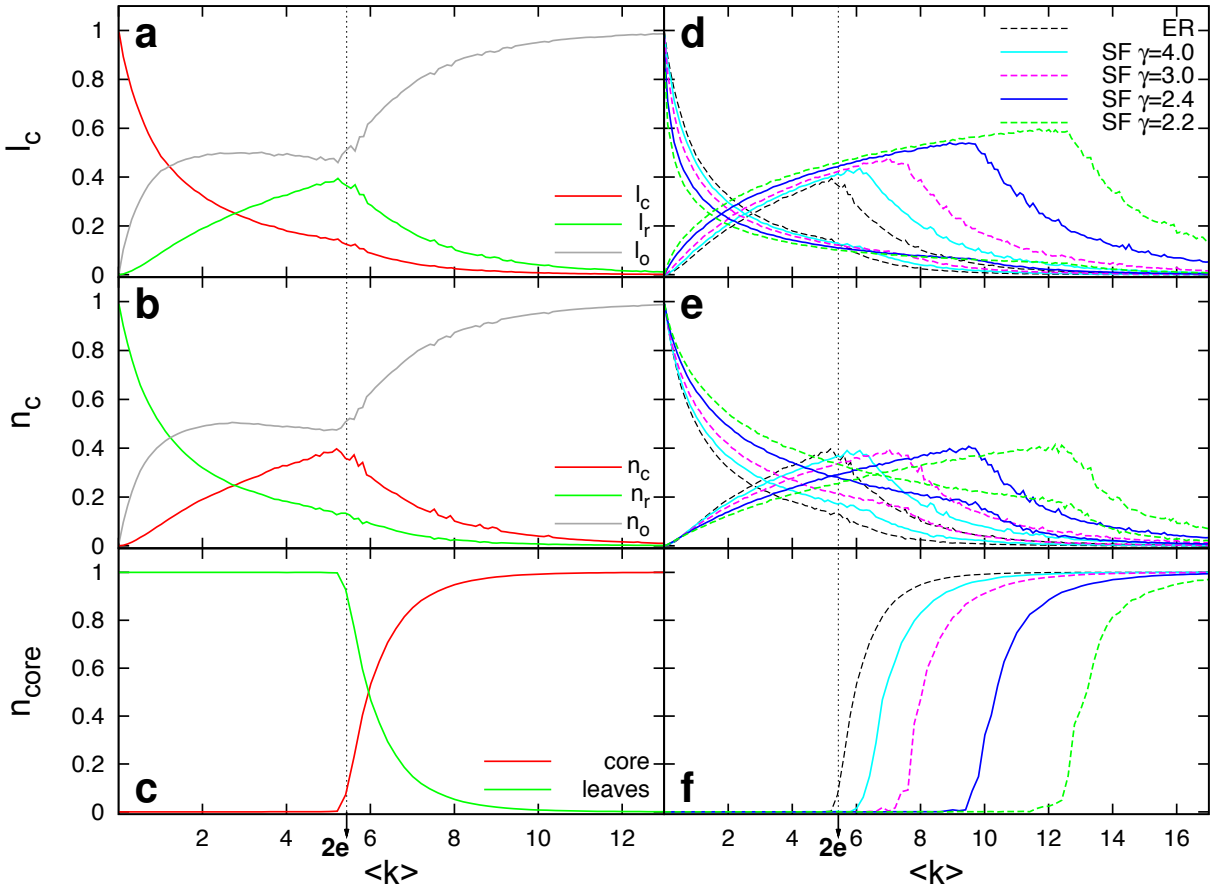

Figure S11: $\langle k \rangle$ **dependent link- and node-categories and core percolation of random networks.** (a,b,c) Erdős-Rényi (ER). (d,e,f) ER and scale-free (SF). (**a**) The fractions of critical, redundant and ordinary links (l_c, l_r and l_o). (**b**) The fractions of critical, redundant and ordinary nodes (n_c, n_r and n_o). (**c**) For ER, the core percolation occurs at $k = \langle k \rangle_c = 2e$. (**d, e, f**) For SF random networks with smaller γ, the core percolation is shifted to higher $\langle k \rangle_c$ value. Both the ER and SF networks are generated from the Chung-Lu model [45] with $N = 10^4$.

TABLE SI: **Real networks analyzed in the paper.** For each network, we show its type, name and reference; number of nodes (N) and edges (L); and brief description.

	name	N	L	description
Regulatory	TRN-Yeast-1 [48]	4,441	12,873	Transcriptional regulatory network of *S. cerevisiae*
	TRN-Yeast-2 [49]	688	1,079	Same as above (compiled by different group).
	TRN-EC-1 [50]	1,550	3,340	Transcriptional regulatory network of *E. coli*
	TRN-EC-2 [49]	418	519	Same as above (compiled by different group).
	Ownership-USCorp [51]	7,253	6,726	Ownership network of US corporations.
Trust	college student [52, 53]	32	96	Social networks of positive sentiment (college students).
	prison inmate [52, 53]	67	182	Same as above (prison inmates).
	Slashdot [54]	82,168	948,464	Social network (friend/foe) of Slashdot users.
	WikiVote [54]	7,115	103,689	Who-vote-whom network of Wikipedia users.
	Epinions [55]	75,888	508,837	Who-trust-whom network of Epinions.com users.
Food Web	Ythan [56]	135	601	Food Web in Ythan Estuary.
	Little Rock [57]	183	2,494	Food Web in Little Rock lake.
	Grassland [56]	88	137	Food Web in Grassland.
	Seagrass [58]	49	226	Food Web in St. Marks Seagrass.
Power Grid	TexasPowerGrid [59]	4,889	5,855	Power grid in Texas.
Metabolic	*E. coli* [60]	2,275	5,763	Metabolic network of *E. coli*.
	S. cerevisiae [60]	1,511	3,833	Metabolic network of *S. cerevisiae*.
	C. elegans [60]	1,173	2,864	Metabolic network of *C. elegans*.
Electronic	s838 [49]	512	819	Electronic sequential logic circuit.
Circuits	s420 [49]	252	399	Same as above.
	s208 [49]	122	189	Same as above.
Neuronal	*C. elegans* [61]	297	2,345	Neural network of *C. elegans*.
Citation	ArXiv-HepTh [62]	27,770	352,807	Citation networks in HEP-TH category of Arxiv.
	ArXiv-HepPh [62]	34,546	421,578	Citation networks in HEP-PH category of Arxiv.
WWW	nd.edu [63]	325,729	1,497,134	WWW from nd.edu domain.
	stanford.edu [54]	281,903	2,312,497	WWW from stanford.edu domain.
	Political blogs [64]	1,224	19,025	Hyperlinks between weblogs on US politics.
Internet	p2p-1 [65]	10,876	39,994	Gnutella peer-to-peer file sharing network.
	p2p-2 [65]	8,846	31,839	Same as above (at different time).
	p2p-3 [65]	8,717	31,525	Same as above (at different time).
Social	UCIonline [66]	1,899	20,296	Online message network of students at UC, Irvine.
Communication	Email-epoch [67]	3,188	39,256	Email network in a university.
	Cellphone [68]	36,595	91,826	Call network of cell phone users.
Intra-	Freemans-2 [69]	34	830	Social network of network researchers.
organizational	Freemans-1 [69]	34	695	Same as above (at different time).
	Manufacturing [70]	77	2,228	Social network from a manufacturing company.
	Consulting [70]	46	879	Social network from a consulting company.

[S1] H. G. Tanner, Decision and Control, 2004. CDC. 43rd IEEE Conference on **3**, 2467 (2004).

[S2] A. Lombardi and M. Hörnquist, Phys. Rev. E **75**, 056110 (2007).

[S3] B. Liu, T. Chu, L. Wang, and G. Xie, IEEE Trans. Auto. Contr. **53**, 1009 (2008).

[S4] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, SIAM J. Control Optim. **48**, 162 (2009).

[S5] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks* (Princeton University Press, Princeton, 2010).

[S6] C.-T. Lin, IEEE Trans. Auto. Contr. **19**, 201 (1974).

[S7] X. F. Wang and G. Chen, Physica A **310**, 521 (2002).

[S8] F. Sorrentino, M. di Bernardo, F. Garofalo, and G. Chen, Phys. Rev. E **75**, 046103 (2007).

[S9] Y. Zou and G. Chen, Europhys. Lett. **84**, 58005 (2008).

[S10] W. Yu, G. Chen, and J. Lü, Automatica **45**, 429 (2009).

[S11] W. Wang and J.-J. Slotine, Biol. Cybern. **92** (2005).

[S12] Q.-C. Pham and J.-J. Slotine, Neural Networks **20** (2007).

[S13] M. Golubitsky and I. Stewart, *The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space* (Birkhäuser, Boston, 2000).

[S14] W. Ren, R. Beard, and E. Atkins, in *American Control Conference, 2005. Proceedings of the 2005* (2005), pp. 1859 – 1864 vol. 3, ISSN 0743-1619.

[S15] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, J. Oper. Res. Soc. **49**, 237 (1998).

[S16] R. Srikant, *The mathematics of Internet congestion control* (Birkhäuser, Boston, 2004).

[S17] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, in *Proceedings of the IEEE* (2007).

[S18] R. W. Shields and J. B. Pearson, IEEE Trans. Auto. Contr. **21**, 203 (1976).

[S19] H. Mayeda and T. Yamada, SIAM J. Control Optim. **17**, 123 (1979).

[S20] D. Ben-Zvi, B.-Z. Shilo, A. Fainsod, and N. Barkai, Nature **453**, 1205 (2008).

[S21] J.-J. Slotine and W. Li, *Applied Nonlinear Control* (Prentice-Hall, 1991).

[S22] W. Lohmiller and J.-J. E. Slotine, Automatica **34**, 6 (1998).

[S23] A. Isidori, *Nonlinear Control Systems* (Springer-Verlag, 1995).

[S24] S. Poljak, IEEE Trans. Auto. Contr. **37**, 1961 (1992).

[S25] D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models, & Applications* (John Wiley & Sons, New York, 1979).

[S26] T. Kailath, *Linear Systems* (Prentice-Hall, Inc., 1980).

[S27] K. Glover and L. M. Silverman, IEEE Trans. Auto. Contr. **21**, 534 (1976).

[S28] E. J. Davison, Automatica **13**, 109 (1977).

[S29] S. Hosoe and K. Matsumoto, IEEE Trans. Auto. Contr. **24**, 963 (1979).

[S30] H. Mayeda, IEEE Trans. Auto. Contr. **26**, 795 (1981).

[S31] A. Linnemann, IEEE Trans. Auto. Contr. **31**, 638 (1986).

[S32] J. E. Hopcroft and R. M. Karp, SIAM J. Comput. **2**, 225 (1973).

[S33] W. Yu, G. Chen, M. Cao, and J. Kurths, IEEE Trans Syst Man Cybern B Cybern **40**, 881 (2010).

[S34] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).

[S35] L. Zdeborová and M. Mézard, J. Stat. Mech. **05**, P05003 (2006).

[S36] L. Lovász and M. D. Plummer, *Matching Theory* (American Mathematical Society, Rhode Island, 2009).

[S37] H. Zhou and Z. Ou-Yang, *Maximum matching on random graphs*, *arXiv:cond-mat/0309348v1* (2003).

[S38] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).

[S39] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).

[S40] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[S41] M. Catanzaro and R. Pastor-Satorras, Eur. Phys. J. B **44**, 241 (2005).

[S42] J.-S. Lee, K.-I. Goh, B. Kahng, and D. Kim, Eur. Phys. J. B **49**, 231 (2006).

[S43] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **38**, 205 (2004).

[S44] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, Phys. Rev. E **71**, 027103 (2005).

[S45] F. Chung and L. Lu, Annals of Combinatorics **6**, 125 (2002).

[S46] Y. S. Cho, J. S. Kim, J. Park, B. Kahng, and D. Kim, Phys. Rev. Lett. **103**, 135702 (2009).

[S47] M. Bauer and O. Golinelli, Eur. Phys. J. B **24**, 339 (2001).

[S48] S. Balaji, M. Madan Babu, L. Iyer, N. Luscombe, and L. Aravind, J. Mol. Biol. **360**, 213 (2006).

[S49] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).

[S50] S. Gama-Castro, V. J. Jacinto, M. Peralta-Gil, A. Santos-Zavaleta, M. I. Pealoza-Spindola, B. Contreras-Moreira, J. Segura-Salazar, L. Muiz-Rascado, I. Martinez-Flores, H. Salgado, et al., Nucleic Acids Research **36**, Database issue:D120 (2008).

[S51] K. Norlen, G. Lucas, M. Gebbie, and J. Chuang, *Eva: Extraction, visualization and analysis of the telecommunications and media ownership network*, *Proceedings of International Telecommunications Society 14th Biennial Conference,Seoul Korea* (2002).

[S52] M. A. J. Van Duijn, M. Huisman, F. N. Stokman, F. W. Wasseur, and E. P. H. Zeggelink, J. Math. Sociol. **27**, 153 (2003).

[S53] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, Science **303**, 1538 (2004).

[S54] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney, *Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters*, *arXiv.org:0810.1355* (2008).

[S55] M. Richardson, R. Agrawal, and P. Domingos, in *Proceedings of the second international semantic web conference* (2003), pp. 351–368.

[S56] J. A. Dunne, R. J. Williams, and N. D. Martinez, Proc. Natl. Acad. Sci. USA **99**, 12917 (2002).

[S57] N. Martinez, Econological Monographs **61**, 367 (1991).

[S58] R. R. Christian and J. J. Luczkovich, Ecological Modelling **117**, 99 (1999).

[S59] G. Bianconi, N. Gulbahce, and A. E. Motter, Phys. Rev. Lett. **100**, 118701 (2008).

[S60] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature **407**, 651 (2000).

[S61] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).

[S62] J. K. J. Leskovec and C. Faloutsos, *Graphs over time: Densification laws, shrinking diameters and possible explanations*, *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* (2005).

[S63] R. Albert, H. Jeong, and A.-L. Barabási, Nature **401**, 130 (1999).

[S64] L. A. Adamic and N. Glance, *The political blogosphere and the 2004 us election*, *Proceedings of the WWW-2005 Workshop on the Weblogging Ecosystem* (2005).

[S65] J. K. J. Leskovec and C. Faloutsos, *Graph evolution: Densification and shrinking diameters*, *ACM Transactions on Knowledge Discovery from Data (ACM TKDD)*, 1(1), 2007.

[S66] T. Opsahl and P. Panzarasa, Social Networks **31**, 155 (2009).

[S67] J.-P. Eckmann, E. Moses, and D. Sergi, Proc. Natl. Acad. Sci. USA **101**, 14333 (2004).

[S68] C. Song, Z. Qu, N. Blumm, and A.-L. Barabsi, Science **327**, 1018 (2010).

[S69] S. Freeman and L. Freeman, Social Science Research Reports 46. University of California, Irvine, CA. (1979).

[S70] R. Cross and A. Parker, *The Hidden Power of Social Networks* (Harvard Business School Press, Boston, MA, 2004).