# Heterogeneity of central nodes explains the benefits of time-varying control scheduling in complex dynamical networks

Erfan Nozari[†]

*Department of Mechanical and Aerospace Engineering, University of California, San Diego, USA*
[†]*Corresponding author. Email: enozari@ucsd.edu*

Fabio Pasqualetti

*Department of Mechanical Engineering, University of California, Riverside, USA*

and

Jorge Cortés

*Department of Mechanical and Aerospace Engineering, University of California, San Diego, USA*

Edited by: Adilson Motter

Despite extensive research and remarkable advancements in the control of complex dynamical networks, most studies and practical control methods limit their focus to time-invariant control schedules (TICS). This is both due to their simplicity and the fact that the benefits of time-varying control schedules (TVCS) have remained largely uncharacterized. In this article, we study networks with linear and discrete-time dynamics and analyse the role of network structure in TVCS. First, we show that TVCS can significantly enhance network controllability over TICS both in small and large networks. Through the analysis of a scale-dependent notion of nodal centrality, we then show that optimal TVCS involves the actuation of the most central nodes at appropriate spatial scales at all times. Consequently, it is the scale-heterogeneity of the central nodes in a network that determine whether, and to what extent, TVCS outperforms conventional policies based on TICS. Here, scale-heterogeneity of a network refers to how diverse the central nodes of the network are at different spatial (local vs. global) scales. Several analytical results and case studies support and illustrate this relationship.

*Keywords*: complex dynamical networks; control scheduling; time-varying actuation; scale heterogeneity.

## 1. Introduction

Many natural and man-made systems, ranging from the nervous system to power and transportation grids to societies, exhibit dynamic behaviours that evolve over a sparse and complex network. The ability to control such network dynamics is not only a theoretically challenging problem but also a barrier to fundamental breakthroughs across science and engineering. While multiple studies have addressed various aspects of this problem, several fundamental questions remain unanswered, including to what extent the capability of controlling a different set of nodes over time can improve the controllability of large-scale, complex networked systems.

Controllability of a dynamical network (i.e., a network that supports the temporal evolution of a well-defined set of nodal *states*) is classically defined as the possibility of steering its state arbitrarily around the state space through the application of external inputs to (i.e., *actuation* of) certain *control nodes* [1]. This raises a fundamental question: how does the choice of control nodes affect network controllability? Hereafter, we refer to this as the *control scheduling problem* [2–4]. Notice that in this classical setting, attention is only paid to the *possibility* of arbitrarily steering the network state, but not to the *difficulty and energy cost* of doing so. This has motivated the introduction of several controllability metrics to quantify the required effort in the control scheduling problem [5–9]. While a comprehensive solution has remained elusive, these works have collectively revealed the role of several factors in the control scheduling problem such as the network size and structure [6], nodal dynamics [3] and centralities [2, 7], the number of control nodes [6] and the choice of controllability metric [8]. This problem has also close connections with the optimal sensor scheduling problem, see e.g., [10–13] and the references therein.

The majority of the above literature, however, implicitly relies on the assumption of time-invariant control schedules (TICS), namely, that the control node(s) is fixed over time. Depending on the specific network structure, this assumption may come at the expense of a significant limitation on its controllability, especially for large-scale systems where distant nodes inevitably exist relative to any control node. Intuitively, the possibility of time-varying control schedules (TVCS), namely, the ability to control different nodes at different times, allows for targeted interventions at different network locations and can ultimately decrease the control effort to accomplish a desired task. On the other hand, from a practical standpoint, the implementation of TVCS requires the ability to geographically relocate actuators or the presence of actuation mechanisms at different, ideally all, network nodes, and more sophisticated control policies. This leads to a critical trade-off between the benefits of TVCS and its implementation costs which has not received enough, if any, attention in the literature.

The significant potential of time-varying schedules for control (and also sensing, which has a dual interpretation to control) has led to the design of (sub)optimal sensor [14, 15] and control [16, 17] scheduling algorithms in recent years. While constituting a notable leap forward and the benchmark for the methods developed in this article, these works are oblivious to the fundamental question of whether, and to what extent, TVCS provides an improvement in network controllability compared with TICS. Our previous work [18] has studied the former question (i.e., whether TVCS provides *any* improvement over TICS) in the case of undirected networks, but did not consider directed networks or, more importantly, addressed the latter question of how large the relative improvement in network controllability is. Given the trade-off between benefits and costs of TVCS, a clear answer to this question is vital for the practical application of TVCS in real-world complex networks.

In this article, we address these two questions in the context of discrete-time linear dynamics evolving over directed networks. Since the implementation costs of TVCS are greatly domain-specific and do not follow any common pattern of dependence on the control schedule, we here provide an in-depth analysis of the benefits of TVCS. This provides the necessary information for comparison with the costs of implementing TVCS in any specific application in order to decide between TICS and TVCS.

To this end, we show that $2k$-*communicability*, a new notion of nodal centrality that we define here, plays a fundamental role in TVCS. This notion measures the centrality of each node in the network at different spatial scales. Throughout this work, the *spatial scale* (or simply *scale*) of any notion of centrality is defined as the maximum topological distance between pairs of nodes that allows them to affect the centrality of each other, where topological distance between a pair of nodes refers to the minimum number of edges in the graph of the network that should be traversed to go from one to the other. In particular, the spatial scale of degree centrality is 1, while the spatial scale of eigenvector centrality is $\infty$. Based on the distinction between local and global nodal centralities (i.e., centralities with small and large spatial scales,

respectively), we show that the optimal control node at every time instance is the node with the largest centrality at the appropriate scale (i.e., the node with the largest $2k$-communicability at an appropriate $k$). Accordingly, our main conclusion is that the benefit of TVCS is directly related to the *scale-heterogeneity of central nodes* in the network: the most benefit is gained in networks where the highest centrality is attained by various nodes at different spatial scales, while this benefit starts to decay as fewer nodes dominate the network at all scales (i.e., scale-homogeneity).

Moreover, we provide an extensive discussion of how the dynamical adjacency matrix of a network can (and should) be extracted from its static connectivity, a vital step that is often ignored in the literature. Indeed, our simulation results show that this step has a significant effect on the benefit of TVCS, with *transmission* networks (networks with states that represent physical quantities transmitted over the network) benefiting significantly more than *induction* networks (those with non-physical states that induce state dynamics over the network) from TVCS.

## 2. Notation and preliminaries

In this section, we introduce our notation and briefly review some preliminary concepts that will be used throughout the work. We use $\mathbb{R}$ and $\mathbb{N}$ to denote the set of reals and positive integers, respectively. Given $x \in \mathbb{R}^n$, $x_i$ and $(x)_i$ refer to its $i$'th component. Similarly, $a_{ij}$ and $(A)_{ij}$ refer to the $(i,j)$th entry of $A$, and $a_i$ refers to its $i$'th column. Given a matrix $M \in \mathbb{R}^{n \times n}$, its trace, determinant, and eigenvalue with smallest magnitude are denoted by $\text{tr}(M)$, $\det(M)$ and $\lambda_{\min}(M)$, respectively.

### 2.1 Graph theory

A weighted undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E}, A)$ consists of a vertex set $\mathcal{N} = \{1, \ldots, n\}$, an edge set

$$\mathcal{E} = \{\{i,j\} \mid i \text{ is connected to } j\},$$

and an adjacency matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$ where, for any $i, j \in \mathcal{N}$, $a_{ij} \geq 0$ is the weight of the edge from node $j$ to node $i$. A path in $\mathcal{G}$ from node $i$ to $j$ is a finite sequence $\ell_0, \ell_1, \ldots, \ell_p$ of nodes where $\ell_0 = i$, $\ell_p = j$, and $\{\ell_{m-1}, \ell_m\} \in \mathcal{E}$ for $\ell \in \{1, \ldots, p\}$. A cycle is a path with $\ell_0 = \ell_p$. For $k \geq 1$, $(A^k)_{ij}$ gives the (weighted) number of paths of length $k$ between nodes $i$ and $j$. A regular graph of degree $k$ is a graph where all the vertices have $k$ neighbours. A strongly regular graph with parameters $(n, k, \lambda, \mu)$ is a regular graph of $n$ nodes with degree $k$ where any two adjacent vertices have $\lambda$ common neighbours and any pair of non-adjacent vertices have $\mu$ neighbours in common. Given a network $\mathcal{G}$ with $n$ nodes, a cone on $\mathcal{G}$ is a network with $n + 1$ nodes where the last one is connected to all others.

### 2.2 Network centrality

We briefly review here three centrality measures with spectral characterizations. Consider a network of size $n$ represented by the adjacency matrix $A$.

### Eigenvector centrality [19, 20]:

Let $v_i \in \mathbb{R}_{\geq 0}$ denote the centrality value of node $i \in \mathcal{N}$. Eigenvector centrality is based on the idea that the influential nodes are the ones that are connected to other influential nodes. In other words, $v_i \propto \sum_{j=0}^{n} a_{ij} v_j$ for all $i$. This requires the existence of a constant $\lambda > 0$ such that $\lambda v_i = \sum_{j=0}^{n} a_{ij} v_j$ for all $i$. In matrix

notation, $v = [v_1 \cdots v_n]^T$, this becomes $Av = \lambda v$, which is an eigenvalue problem. Since $A$ is non-negative, by the Perron-Frobenius Theorem [21, Fact 4.11.4], there always exists a pair $(\lambda, v) \in \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0}^n$ such that $Av = \lambda v$. This vector $v$ is thus defined as the vector of (right) eigenvector centralities. The same argument can be repeated by reversing the direction of influence flow in the network, leading to the vector of left eigenvector centralities (i.e., a positive vector $u$ such that $u^T A = \lambda u^T$).

**Exponential and resolvent communicability [22, 23]:**

The communicability of a node measures its ability to communicate with the rest of the network. Different notions of communicability have been proposed for complex networks. For a given node $i$, these include exponential communicability $(e^{\beta A})_{ii}$ and the resolvent communicability $((I - \beta A)^{-1})_{ii}$, respectively, where $\beta > 0$. From the power series expansion of $e^{\beta A}$ and $(I - \beta A)^{-1}$, it follows that the exponential and resolvent communicabilities count the total number of cycles that pass through node $i$, weighting the "importance" of cycles of length $k$ by $\beta^k/k!$ and $\beta^k$, respectively. Thus, the role of $\beta$ is to determine how local/global these measures are: increasing $\beta$ increases the weights of longer cycles. One can show [23] that in the extreme cases of $\beta \to \infty$ in the exponential case and $\beta \to \frac{1}{\lambda_{\max}(A)}$ in the resolvent case, both notions result in the same rankings of nodes as eigenvector centrality.

**Degree centrality:**

The degree centrality of node $i$ is the sum of the $i$'th row (or column) of $A$ and provides a measure of the immediate influence of node $i$ on its neighbours.

## 3. Problem statement: comparison of time-varying and time-invariant control scheduling

We consider a network of $n$ nodes that communicate over a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E}, A)$ that is in general weighted and directed (see Appendix A for methods of obtaining $A$ from network connectivity structure). Each node $i$ has a *state* value $x_i \in \mathbb{R}$ that evolves over time through the interaction of node $i$ with its neighbours in $\mathcal{G}$ and an external control $u$. Assuming that these interactions are linear and time-invariant, we have

$$x(k + 1) = Ax(k) + b(k)u(k), \quad k \in \{0, \ldots, K - 1\}, \tag{1}$$

where $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is the network state, $u(k) \in \mathbb{R}$ is the control input, $b(k) \in \mathbb{R}^n$ is the *time-varying* input vector and $K$ is the time horizon. For simplicity of exposition, we consider only one control input at a time, but the discussion is generalizable to multi-input networks (cf. Appendix E). Define

$$\iota_k \in \mathcal{N}, \tag{2}$$

to be the index of the node to which the control signal $u(k)$ is applied at time $k$. Then, $b(k)$ is equal to the $\iota_k$'th column of the identity matrix. For the sake of simplicity, we here assume that all the network nodes are actuatable, so $\iota_k \in \mathcal{N}$. If a subset of nodes are *latent*, (i.e., not actuatable), further challenges arise and thus we postpone the analysis of this case to Section 4.4.

The dynamical network (1) is *controllable* if its state can be steered from arbitrary $x(0) = x_0$ to arbitrary $x(K) = x_f$ using the control input $\{u(k)\}_{k=0}^{K-1}$ or, equivalently, if the *controllability Gramian*

$$\mathcal{W}_K = \sum_{k=0}^{K-1} A^k b(K-1-k) b(K-1-k)^T (A^T)^k, \tag{3}$$

is non-singular [24]. In general, the eigenvalues of $\mathcal{W}_K$ determine how large the *unit-energy reachability set* (the set of states $x_f$ that can be reached from the origin $x_0 = 0$ using controls with unit energy) is (cf. Appendix B for derivation). Therefore, various measures of controllability based on the eigenvalues of $\mathcal{W}_K$ have been proposed, most notably $\mathrm{tr}(\mathcal{W}_K)$, $\mathrm{tr}(\mathcal{W}_K^{-1})^{-1}$, $\det(\mathcal{W}_K)$, $\lambda_{\min}(\mathcal{W}_K)$. Each metric has its own benefits and limitations, on which we elaborate more in the following.

Assume, for now, that $f(\mathcal{W}_K) \geq 0$ is any of the aforementioned controllability measures. In *optimal control scheduling*, we seek to choose the control nodes $\{\iota_k\}_{k=0}^{K-1}$ (or, equivalently, $\{b(k)\}_{k=0}^{K-1}$) optimally. The conventional approach in the literature [2–9] is to assume a constant control node, thus called the *time-invariant control scheduling (TICS) problem*:

$$\text{TICS:} \qquad \max_{\iota_0,\dots,\iota_{K-1} \in \mathcal{N}} \quad f(\mathcal{W}_K) \tag{4a}$$

$$\text{s.t.} \qquad \iota_0 = \cdots = \iota_{K-1} \tag{4b}$$

The main advantage of TICS is its simplicity, from theoretical, computational and implementation perspectives. However, this simplicity comes at a possibly significant cost in terms of network controllability, compared with the case where the control nodes $\{\iota_k\}_{k=0}^{K-1}$ are independently chosen, namely,

$$\text{TVCS:} \qquad \max_{\iota_0,\dots,\iota_{K-1} \in \mathcal{N}} \quad f(\mathcal{W}_K). \tag{5}$$

This approach, namely, time-varying control scheduling (TVCS), is at least as good as TICS, but has the potential to improve network controllability significantly. Figure 1(a and b) illustrates a small network of $n = 5$ nodes together with the optimal values of equations (4) and (5) and the relative advantage of TVCS over TICS, defined as

$$\chi = \frac{f_{\max}^{\text{TV}} - f_{\max}^{\text{TI}}}{f_{\max}^{\text{TI}}}. \tag{6}$$

Three observations are worth highlighting. First, the value of $\chi$ is extremely dependent on the choice of controllability measure $f$, and different choices lead to orders of magnitude change in $\chi$. Second, the relative advantage of TVCS over TICS is significant for all choices of the controllability measure, with the minimum improvement of $\chi = 35\%$ for the choice of $f(\cdot) = \mathrm{tr}(\cdot)$. The fact that $f(\cdot) = \mathrm{tr}(\cdot)$ results in the smallest value of $\chi$ relative to other measures is consistently observed in synthetic and real-world networks, and stems from the fact that $\mathrm{tr}(\mathcal{W}_K)$ has the smallest sensitivity (greatest robustness) to the choice of control schedule. Finally, even with optimal TVCS, $\lambda_{\min}(\mathcal{W}_K)$ is orders of magnitude less than 1, indicating the inevitable existence of very hard-to-reach directions in the state space. This shows that efficient controllability cannot be maintained in all directions in the state space even using TVCS and even in very small networks with control over $1/5 = 20\%$ of the nodes. Except for $\mathrm{tr}(\mathcal{W}_K)$, all the measures rely heavily on this least-controllable direction, while $\mathrm{tr}(\mathcal{W}_K)$ trades this off for improved controllability in the most efficient directions in the state space. See Appendix B for further discussion of this trade-off.
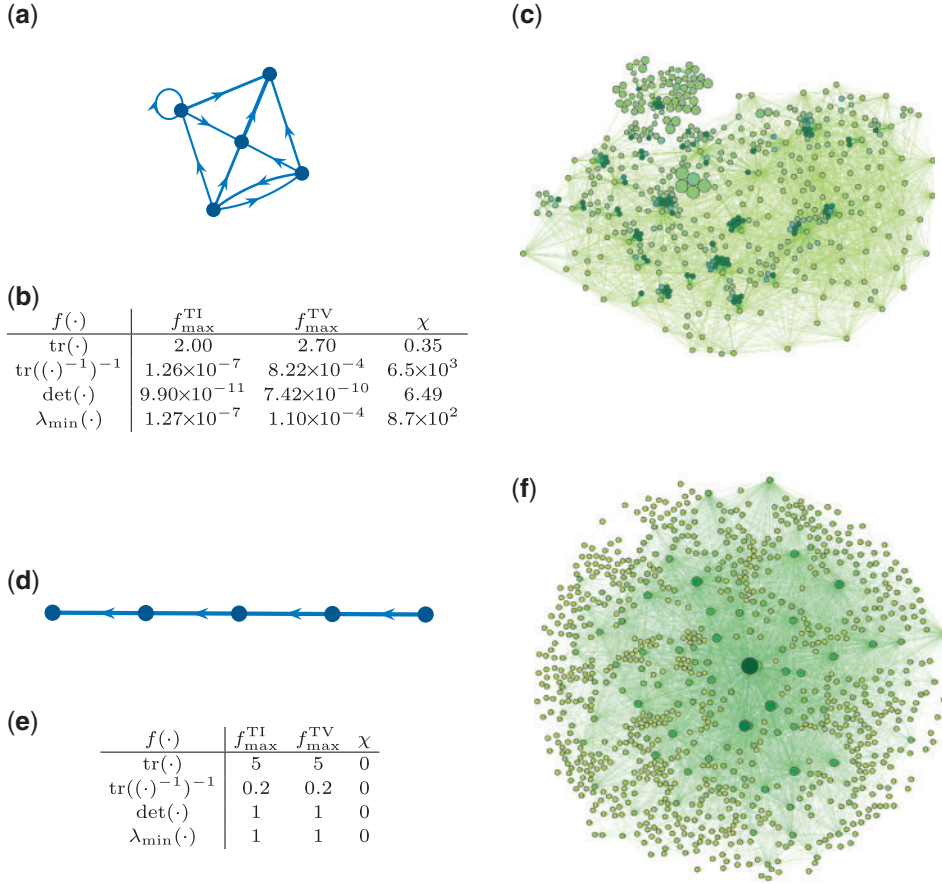
**(a)**



**(b)**

| $f(\cdot)$ | $f_{\max}^{\mathrm{TI}}$ | $f_{\max}^{\mathrm{TV}}$ | $\chi$ |
|---|---|---|---|
| $\mathrm{tr}(\cdot)$ | 2.00 | 2.70 | 0.35 |
| $\mathrm{tr}((\cdot)^{-1})^{-1}$ | $1.26{\times}10^{-7}$ | $8.22{\times}10^{-4}$ | $6.5{\times}10^{3}$ |
| $\det(\cdot)$ | $9.90{\times}10^{-11}$ | $7.42{\times}10^{-10}$ | 6.49 |
| $\lambda_{\min}(\cdot)$ | $1.27{\times}10^{-7}$ | $1.10{\times}10^{-4}$ | $8.7{\times}10^{2}$ |

**(d)**



**(e)**

| $f(\cdot)$ | $f_{\max}^{\mathrm{TI}}$ | $f_{\max}^{\mathrm{TV}}$ | $\chi$ |
|---|---|---|---|
| $\mathrm{tr}(\cdot)$ | 5 | 5 | 0 |
| $\mathrm{tr}((\cdot)^{-1})^{-1}$ | 0.2 | 0.2 | 0 |
| $\det(\cdot)$ | 1 | 1 | 0 |
| $\lambda_{\min}(\cdot)$ | 1 | 1 | 0 |

**(c)**



**(f)**



FIG. 1. Advantage of TVCS in dynamic networks. (a) A small example network of $n = 5$ nodes. The thickness of each edge $(i,j)$ illustrates its weight $a_{ij}$. (b) The optimal values of TICS and TVCS (equations (4) and (5), respectively) and the relative TVCS advantage (equation (6)) for the network in (a). (c) An air transportation network among the busiest airports in the United States (see 'air500' in Table 1 for details). The network is undirected, and the dynamical adjacency matrix $A$ is computed from static connectivity using the transmission method (cf. Appendix A). This is an example of a network that significantly benefits from TVCS with $\chi \simeq 20\%$. (d) A small example network of the same size as (a) but with no benefit from TVCS. (e) The optimal values of TICS and TVCS (equations (4) and (5), respectively) and the relative TVCS advantage (equation (6)) for the network in (d). We see that the network does not benefit from TVCS independently of the choice of controllability metric. (f) A social network of students at the University of California, Irvine (see 'UCI Forum' in Table 1 for details). Similar to (c), the network is undirected and the adjacency matrix is computed using the transmission method. This network, however, does not benefit from TVCS ($\chi = 0$). In (c) and (f), the controllability measure of equation (7) is used due to the large size of the network. In both cases, the colour intensity and size of nodes represent their values of $R_i(1)$ and $R_i(K-1)$, respectively ($K = 10$). While there is a close correlation between nodal size and colour intensity in (f) (i.e., the darkest nodes are also the largest), this is not the case in (c), which is the root cause for the difference in their $\chi$-values. The interested reader can find comprehensive discussions of the network control problem for air transportation in [25–28], social opinion in [29–33] and social epidemic dynamics in [34–39] and references therein.

Despite the significant increase in size and complexity, the same core principles outlined above apply to controllability of real-world networks. The large size of these networks, however, imposes new constraints on the choice of the controllability measure $f$ that make the use of $f(\cdot) = \lambda_{\min}(\cdot)$, $\mathrm{tr}((\cdot)^{-1})^{-1}$ and $\det(\cdot)$ numerically infeasible and theoretically over-conservative, as discussed in detail in Appendix B. As a

result, we resort to the particular choice of controllability measure

$$f(\mathcal{W}_K) = \text{tr}(\mathcal{W}_K), \tag{7}$$

for networks beyond $n \simeq 15$. Since this measure has the smallest sensitivity to the choice of $\{\iota_k\}_{k=0}^{K-1}$ (Fig. 1(b)), we expect any network that benefits from TVCS using the choice of equation (7) to also benefit from it using other Gramian-based measures (while the converse is not necessarily true, i.e., there are networks that significantly benefit from TVCS using other measures but show no benefit in terms of $\text{tr}(\mathcal{W}_K)$). Figure 1(c) illustrates an air transportation network among the busiest airports in the United States, comprising of $n = 500$ nodes. Using (7), we see $\chi \simeq 20\%$ improvement in controllability, verifying our expectation about the benefits of TVCS.

In spite of this potential benefit, TVCS has usually higher computational and implementation costs. These include the higher computational cost of computing the optimal TVCS, and that of installing an actuator at several (ideally all) nodes of the network. Further, not all networks benefit from TVCS alike. A simple directed chain network with the same size as that of Fig. 1(a) gains absolutely no benefit from TVCS, independently of the choice of $f$ (Fig. 1(d and e)). Similarly, $\chi = 0$ is also observed in larger, complex networks, indicating that the optimal TVCS and the optimal TICS are the same (Fig. 1(f)).

These observations collectively raise a fundamental question that constitutes the main problem studied in this article. Before formally stating the problem, we need a definition for ease of reference.

DEFINITION 3.1 *(Class $\mathcal{V}$ and $\mathcal{I}$ networks).* Consider a dynamical network described by (1) and the measure $\chi$ introduced in (6). We say that the network belongs to class $\mathcal{V}$ if it has $\chi > 0$ and we say it belongs to class $\mathcal{I}$ otherwise ($\chi = 0$).

In words, class $\mathcal{V}$ networks are those that benefit from TVCS and class $\mathcal{I}$ networks are those that do not. Our main problem of interest is then as follows.

PROBLEM 1 Given the set of all dynamical networks described by dynamics of the form (1), characterize the sets $\mathcal{V}$ and $\mathcal{I}$ in terms of the network structure $A$ and develop efficient and interpretable methods for distinguishing between them.

In the following, we restrict our attention to the choice of controllability measure in equation (7) due to its applicability to all network sizes and carry a thorough analysis of its properties in order to address Problem 1.

## 4. Main results

In this section, we present our main results regarding Problem 1. First, we introduce a new notion of communicability that is pivotal to the solution of Problem 1. Then, we present our results regarding the characterization of class $\mathcal{V}$ and $\mathcal{I}$ networks and, finally, study the case of networks with latent nodes declared earlier.

### 4.1 2k-*Communicability and scale-heterogeneity*

Consider the TVCS problem in equation (5) with $f(\cdot) = \text{tr}(\cdot)$. Using the definition of the controllability Gramian in (3) and the invariance property of trace under cyclic permutations, we can write

$$\text{tr}(\mathcal{W}_K) = \sum_{k=0}^{K-1} b(K-1-k)^T (A^k)^T A^k b(K-1-k).$$

Therefore,

$$\max_{\iota_0,\dots,\iota_{K-1}} \text{tr}(\mathcal{W}_K) = \sum_{k=0}^{K-1} \max_{\iota_{K-1-k}} b(K-1-k)^T (A^k)^T A^k b(K-1-k),$$

where each term $b(K-1-k)^T (A^k)^T A^k b(K-1-k)$ is the $\iota_{K-1-k}$'th diagonal entry of $(A^k)^T A^k$ (cf. equation (2)). Therefore, the optimization in (5) boils down to finding the largest diagonal element of $(A^k)^T A^k$ and applying $u(K-1-k)$ to this node. On the other hand, for the TICS problem in (4) we have

$$\text{tr}(\mathcal{W}_K) = b^T \left( \sum_{k=0}^{K-1} (A^k)^T A^k \right) b,$$

so one has to instead find the largest diagonal entry of $\sum_{k=0}^{K-1} (A^k)^T A^k$ and apply all the control inputs $u(0), \dots, u(K-1)$ to this same node, which is clearly sub-optimal with respect to TVCS. This discussion motivates the following definition.

DEFINITION 4.1 *(2k-**communicability**).* Given the network dynamics (1), the 2k-communicability of a node $i \in \mathcal{N}$ is defined as

$$R_i(k) = ((A^k)^T A^k)_{ii}, \qquad i \in \mathcal{N}, \quad k \geq 0. \tag{8}$$

Figure 2(a and b) illustrates the evolution of $R_i(k)$ as a function of $k$ for all $i \in \mathcal{N}$ for a sample network of $n = 20$ nodes.

Perhaps the most salient property of 2k-communicability is the extent to which it relies on the local interactions among the nodes. Recall, cf. [40], that for any $k$, the $(i, j)$ entry of $A^k$ equals the total number of paths of length $k$ from node $i$ to $j$ (if the graph is weighted, each path counts as its weight, equal to the product of the weights of its edges). From equation (8), we see that $R_i(k)$ equals the sum of the squares of the total (weighted) number of paths of length $k$ ending in node $i$. In other words, $R_i(k)$ only depends on connections of node $i$ with its $k$-hop out-neighbours and is independent of the rest of the network. Therefore, $R_i(k)$ is a local notion of centrality for small $k$ and it incorporates more global information as $k$ grows. In particular, as shown in Appendix C, $R_i(k)$ is closely related to

- the out-degree centrality of node $i$ for $k = 1$;
- the left eigenvector centrality of node $i$ for $k \to \infty$.

This scaling property of 2k-communicability is illustrated in Fig. 2(a–d) for an example network of $n = 100$ nodes. Accordingly, we take the left eigenvector centrality squared as the definition of $R_i(\infty)$ in the sequel.
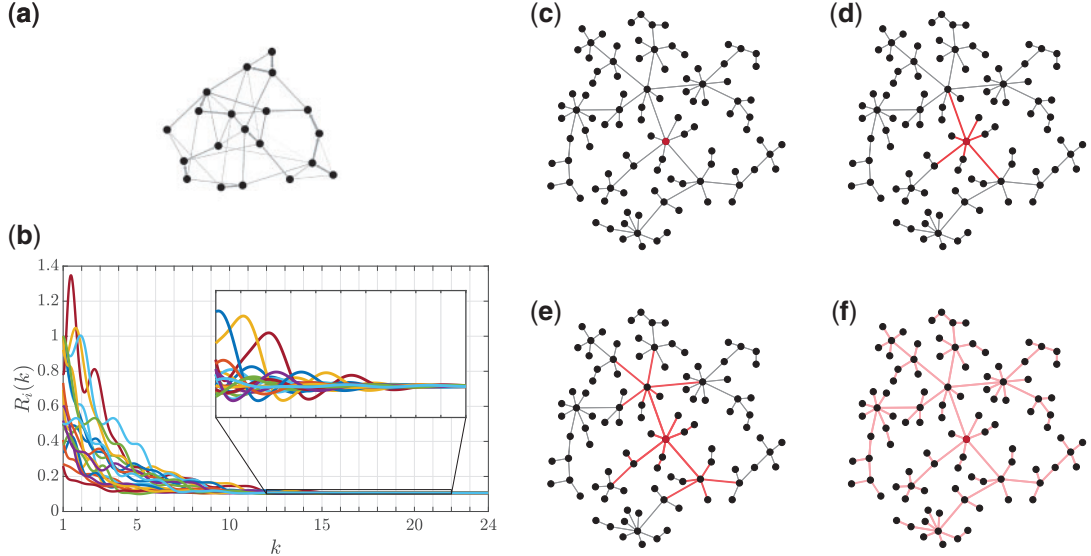
FIG. 2. $2k$-communicability of dynamical networks. (a) An example network of $n = 20$ nodes for illustration of the dependence on $k$ of nodal $2k$-communicabilities. The thickness of the edges is proportional to their weights. (b) The evolution of the functions $\{R_i(k)\}_{i=1}^n$. Although these functions are originally only defined over integer values of $k$, we have extended their domain to real numbers for better illustration of their crossings and oscillatory behaviour. Oscillatory behaviour only arises when $A$ has complex-valued eigenvalues (otherwise, $R_i(k)$ is strictly convex). (c) An example network of $n = 100$ nodes for illustration of the scaling property of $2k$-communicability. The node whose $2k$-communicabilities are to be computed (i.e., "node $i$") is depicted in red. (d–f) The 2-, 4- and 14-communicability of the node depicted in red, as determined by its 1-, 2- and 7-hop incoming paths. We see that $R_i(1)$ only depends on the immediate (out-)neighbours of $i$, but as $k$ grows, $R_i(k)$ encodes more global information.

The scaling property of $2k$-communicability also plays an important role in Problem 1. For ease of reference, let

$$r(k) \in \mathcal{N}$$

denote the index of the node that has the largest $R_i(k)$. Then, according to the discussion above,

$$\iota_k^* = r(K - 1 - k), \tag{9}$$

which forms the core connection between $2k$-communicability and TVCS. From this, we see that the optimal TVCS involves the application of $u(0)$ to the node $r(K - 1)$ with the highest global centrality and gradually moving the control node until we apply $u(K - 2)$ to the node $r(1)$ with the highest local centrality (the control node at time $K - 1$ is arbitrary as $R_i(0) = 1$ for all $i$). The intuition behind this procedure is simple. At $k = 0$, the control input has enough time to propagate through the network, which is why the highest globally-central node should be controlled. As we reach the control horizon $K$, the control input has only a few time steps to disseminate through the network, hence the optimality of locally-central nodes. This further motivates our definition of *scale-heterogeneity*, as follows.

DEFINITION 4.2 *(Scale-heterogeneity of dynamical networks).* Consider the network dynamics (1) subject to the TVCS problem (5) with $2k$-communicability as defined in Definition 4.1. The network is called

*scale-homogeneous* if $r(1) = r(2) = \cdots = r(\infty)$ and scale-heterogeneous otherwise. Accordingly, the more varied $\{r(k)\}_{k=1}^{\infty}$ and $\{R_{r(k)}(k)\}_{k=1}^{\infty}$ are, the more scale-heterogeneous the network is.

Based on this definition, we see that the scale-heterogeneity is the main factor in the benefit of TVCS over TICS. In fact, scale-homogeneous and scale-heterogenous networks are the same as class $\mathcal{I}$ and $\mathcal{V}$ networks, respectively, due to (9). Further, note that the degree of scale-heterogeneity provides a *geometric and qualitative* characterization of the amount of benefit TVCS has over TICS and distinguishes between networks in $\mathcal{V}$ that only marginally benefit from TVCS and those the benefit significantly (while $2k$-communicability is a more *quantitative* notion used for computational assignment of networks to class $\mathcal{V}$ or $\mathcal{I}$).

It follows immediately from Definition 4.2 that determining the scale-heterogeneity of a network requires computation of all $\{r(k)\}_{k=1}^{\infty}$ which is expensive for large networks. Next, we seek simple and computationally efficient conditions to be used as a proxy for scale-heterogeneity.

### 4.2 *Identifying class $\mathcal{V}$ networks*

In this section, we discuss a sufficient condition for scale-heterogeneity that, when satisfied, ensures that a network belongs to class $\mathcal{V}$. This condition, given next, relies on the fact that $r(1)$ and $r(\infty)$ are particularly important elements of $\{r(k)\}_{k=1}^{\infty}$ in determining scale-heterogeneity. The proof of this theorem is given in Appendix G.

THEOREM 4.3  *(Class $\mathcal{V}$ networks).* Consider the TVCS problem (5) for the network dynamics (1). Assume that the adjacency matrix $A$ is irreducible, aperiodic and diagonalizable. If

$$\arg\max_{i \in \mathcal{N}} R_i(1) \cap \arg\max_{i \in \mathcal{N}} R_i(\infty) = \emptyset,$$

then the network belongs to class $\mathcal{V}$ for sufficiently large $K$.                                                           □

The condition of $A$ being irreducible is equivalent to the network being strongly connected, and thus not restrictive. Likewise, $A$ being aperiodic is not restrictive as it requires that there exists no integer number greater than 1 that divides the length of every cycle in the network (satisfied, in particular, if any self-loops exist). Finally, $A$ is almost always diagonalizable in the Lebesgue sense, i.e., the set of non-diagonalizable $A$ has Lebesgue measure zero.

Consider again the networks of Fig. 1(c and f). Here, the colour intensity of each node indicates its value $R_i(1)$, while its size corresponds to its value $R_i(K-1)$. Clearly, the first few largest and darkest nodes are distinct in Fig. 1(c), while there is a close correlation between nodal size and darkness in Fig. 1(f), illustrating the root cause of their difference in benefiting from TVCS.

If a network has $r(0) = r(K-1)$, it is still possible that the network belongs to class $\mathcal{V}$. In fact, about half of the networks with $r(0) = r(K-1)$ still belong to $\mathcal{V}$ (Fig. 3(a)). However, these networks have a value of $\chi$ of no more than 3% on average, and in turn this value quickly decreases with the dominance of the node $r(0)$ over the rest of the network nodes (Fig. 3(b)). This is a strong indication that, for most practical purposes, the test based on $2k$-communicability is a valid indicator of whether a network benefits from TVCS. Furthermore, in the case of undirected networks, it is possible to analytically prove that a network belongs to class $\mathcal{I}$ ($\chi = 0$) if certain conditions based on the eigen-decomposition of the adjacency matrix $A$ are satisfied, as shown next.
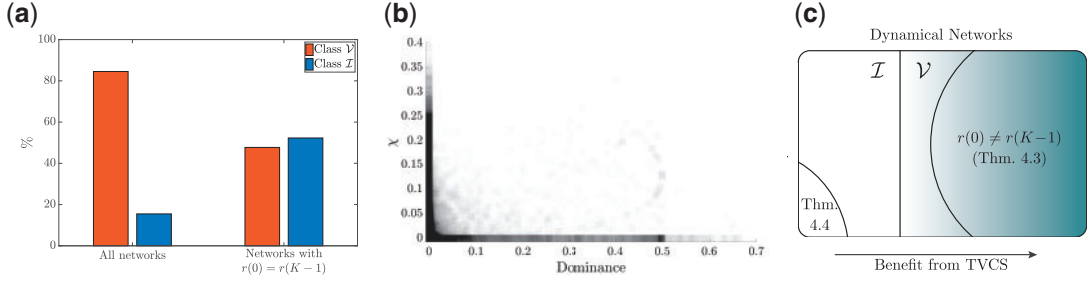
FIG. 3. The role of $2k$-communicability in distinguishing between networks of class $\mathcal{V}$ ($\chi > 0$) and $\mathcal{I}$ ($\chi = 0$). (a) The proportion of random networks in $\mathcal{V}$ and $\mathcal{I}$. A total of $10^5$ random connectivity matrices were generated with logarithmically-uniform $n$ in $[10^1, 10^3]$, uniform sparsity $p$ in $[0, 1]$, and uniform pairwise connectivity weight in $[0, 1]$, and then transformed to adjacency matrices $A$ using the transmission method (cf. Appendix A). A time-horizon of $K = 10$ is used for all networks. While more than 80% of all networks belong to class $\mathcal{V}$, this number drops to less than 50% among networks with $r(1) = r(K - 1)$ (i.e., networks where the same node has the greatest local and global centralities). (b) The $\chi$-value of the same networks as in (a) that have $r(1) = r(K - 1)$ as a function of the dominance of the node $r(0)$. For the node $r(0)$, its *dominance* (over the rest of the network) is a measure of how distinctly $R_{r(0)}(1)$ and $R_{r(0)}(K - 1)$ are larger than $R_i(1)$ and $R_i(K - 1)$, respectively, for $i \neq r(0)$ (cf. Appendix D). Each grey square represents one randomly generated network, so the darkness of each area represents the probability of observing random networks with that value of (dominance, $\chi$). A rapid decay of $\chi$ with dominance is clear, such that networks with positive dominance have very low probability of having $\chi > 0$. (c) A Venn diagram illustrating the decomposition of dynamical networks based on the extent to which they benefit from TVCS. The colour gradient is a depiction of this extent, as measured by $\chi$ (equation (6)), where darker areas correspond to higher $\chi$. As shown in (a) and (b), the class of networks for which $r(0) \neq r(K - 1)$ is only a subset of $\mathcal{V}$ but provides a good approximation for it.

### 4.3 Identifying class $\mathcal{I}$ networks

Complementary to Section 4.2, here we discuss some necessary conditions for scale-heterogeneity based on the eigen-structure of the network that characterize subsets of $\mathcal{I}$. Let $A = V\Lambda V^T$ be the eigen-decomposition of $A$, where $V = [v_{ij}]_{n \times n}$ and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues with $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. Further, let $W = [w_{ij}]_{n \times n}$ be the doubly stochastic matrix such that $w_{ij} = v_{ij}^2$ for all $i, j \in \{1, \ldots, n\}$. The next result, proven in Appendix G, characterizes three undirected sub-classes of $\mathcal{I}$.

THEOREM 4.4 *(Class $\mathcal{I}$ networks).* Consider the TVCS problem (5) for the network dynamics (1). Assume that the network is undirected (i.e., $A = A^T$) and that, without loss of generality, the node with the largest eigenvector centrality is labelled as node 1. If any of the following conditions holds:

(i) $\frac{1 - w_{11}}{w_{11}} \leq \frac{|\lambda_1| - |\lambda_2|}{|\lambda_1| - |\lambda_n|}$,

(ii) $w_{11} + w_{12} = 1$,

(iii) the network has three or fewer nonzero eigenvalues with different absolute values and $1 \in \arg\max_i R_i(1)$,

then,

$$1 \in \arg\max_{1 \leq i \leq n} R_i(k), \quad \forall k \in \{0, \ldots, K - 1\}, \tag{10}$$

i.e., selecting the node with the largest eigenvector centrality at every time step is the solution to (5). □

The conditions in Theorem 4.4 are based on the eigen-decomposition of the network adjacency matrix $A$ and thus abstract. However, these conditions can be interpreted as follows:

(i) Condition *(i)* holds for networks where there is a sufficiently distinct central node, in the sense of eigenvector centrality, and the network dynamics is dominated by the largest eigenvalue. An extreme case of such networks is a totally disconnected network where $W = I$ and the highest authority is the node with the largest self-loop.

(ii) Condition *(ii)* holds for networks where the eigenvector centrality of all nodes is determined by the weight of the link to the most eigenvector-central node. To see this, note that we have $w_{1j} = 0$ for $j \geq 3$, implying $v_{1j} = 0, j \geq 3$. Since the rows of $V$ are orthogonal, we deduce $v_{i2} = \alpha v_{i1}$ for all $i \geq 2$, where $\alpha = -v_{11}/v_{12}$ is constant. Using $A = V\Lambda V^T$, we have

$$a_{1i} = \lambda_1 v_{11} v_{i1} + \lambda_2 v_{12} v_{i2} = (v_{11}\lambda_1 + \alpha v_{12}\lambda_2)v_{i1},$$

so $v_{i1} \propto a_{1i}$ for all $i \geq 2$. Examples of such networks are star networks with no (or small-weight) self-loops (cf. Proposition F.3).

(iii) Regarding condition *(iii)*, the most well-known families of networks with three distinct eigenvalues are the complete bipartite networks and connected strongly regular networks. Moreover, cones on $(n, k, \lambda, \mu)$-strongly regular graphs satisfying $\lambda_{\min}(A)(\lambda_{\min}(A) - k) = n$ are also known to have three distinct eigenvalues [41]. The other condition $1 \in \arg\max_i R_i(1)$ holds when the node with the largest eigenvector centrality (i.e., $r(\infty)$) has also the largest 2-communicability. The simplest example of a network with these properties is the star network (with no or equal self-loops).

The general abstraction from these cases is that a network belongs to class $\mathcal{I}$ if it contains a sufficiently distinct central node, which reinforces our main conclusion that $\mathcal{V}$ is the class of networks with multiple scale-heterogeneous central nodes. The inclusion relationships between the various classes of networks introduced in this section are summarized in Fig. 3(c).

While Theorem 4.4 is only applicable to undirected networks, it has a straightforward extension to normal networks (i.e., directed networks with normal $A$). Using the same proof technique as in Theorem 4.4, it can be shown that the exact same results hold if one replaces the eigenvalues and eigenvectors with singular values and singular vectors of $A$. Interestingly in this case, $R_i(\infty)$ coincides with HITS hub/authority centrality of node $i$ squared [42].

### 4.4 *Networks with latent nodes*

As mentioned in Section 3, in many real-world applications of TVCS not all the nodes are available/accessible for control. In this case, we call a node *manifest* if it can be actuated and *latent* if it cannot. The natural solution would then be to choose the control nodes optimally among the manifest nodes. If the adjacency matrix $A$ of the network is fixed and given, this is the best solution. However, there are cases where $A$ itself can be changed, at least among the manifest nodes. We call such a change of structure an *(edge) manipulation*. Edge manipulations are primarily possible in man-made (power, transportation, etc.) networks, since the edges are originally engineered, but are also becoming increasingly feasible in biological networks due to advances in bioengineering, see e.g., [43, 44] for brain and [45, 46] for gene networks. When manipulation is possible in a network with latent nodes, another solution to TVCS is to manipulate the network among the manifest nodes such that the optimal control nodes (when

computed *without* any restrictions on control scheduling) lie among the manifest nodes for all time. The following result provides a guarantee that this is always possible, provided that the manipulation is sufficiently strong and not acyclic.

THEOREM 4.5 *(Network manipulation and TVCS in networks with latent nodes).* Consider the optimal node selection problem (5) over a time horizon $K$. Given a network of $n$ nodes with adjacency matrix $A_0 \in \mathbb{R}^{n \times n}$, let $E \in \mathbb{R}^{n \times n}$ be a non-negative matrix of the form

$$
E = \begin{bmatrix} \overbrace{\star}^{n_1} & \overbrace{0}^{n-n_1} \\ 0 & 0 \end{bmatrix} \begin{matrix} \}n_1 \\ \}n-n_1 \end{matrix},
$$

corresponding to the manifest subnetwork involving the first $n_1 < n$ nodes (this is without loss of generality, since nodes can be renumbered) and consider the dynamic network described by (1) with adjacency matrix $A = A_0 + \alpha E$, where $\alpha > 0$. Then, if $E$ is not acyclic, there exists $\bar{\alpha} > 0$ such that for $\alpha > \bar{\alpha}$,

$$
r(k) \in \{1, \ldots, n_1\}, \tag{11}
$$

for all $k \in \{0, \ldots, K-1\}$. Furthermore, if $A_0$ and $E$ are symmetric (the corresponding networks are undirected), $\bar{\alpha}$ can be found in closed form and (11) holds for all $k \geq 1$. $\square$

Both requirements of Theorem 4.5 (that $\alpha E$ is sufficiently strong and acyclic) have clear interpretations. First, depending on how large the size of the manifest subnetwork is and how central its nodes already are (pre-manipulation), larger manipulation may be necessary to turn them into central nodes at various scales (i.e., $r(k)$ for $k = \{0, \ldots, K-1\}$). Second, for the manifest nodes to become central at arbitrarily global scales (i.e., $r(k)$ for $k \sim K \to \infty$), the manipulation must contain paths of arbitrarily long lengths, which are absent in acyclic networks.

According to Theorem 4.5, manipulation of the manifest subnetwork is effective even when the manifest nodes are among the least central nodes of the network (before the manipulation). In this case, as we increase $\alpha$ from 0, the manifest nodes usually first turn into the most locally-central nodes ($\alpha \not\gtrsim \bar{\alpha}$ yet), and then also into globally-central nodes ($\alpha > \bar{\alpha}$). The following example illustrates this phenomenon in a simple star network where the centre node is latent and the peripheral nodes are manifest.

EXAMPLE 4.6 *(Undirected star networks with varying self-loop weights).* Consider an undirected uniform star network given by

$$
A_0 = \begin{bmatrix} l_p I_{n-1} & a_{cp}\mathbf{1}_{n-1} \\ a_{cp}\mathbf{1}_{n-1}^T & l_c \end{bmatrix},
$$

where $\mathbf{1}_{n-1}$ denotes the $(n-1)$-dimensional vector of all ones and the positive constants $l_c$, $l_p$ and $a_{cp}$ are the central self-loop weight, peripheral self-loop weight and the link weight between the centre node and any peripheral node, respectively. The $2k$-communicabilities of this network are computed analytically in Proposition F.3 (nodes are re-labeled here for conformity with Theorem 4.5). It follows from (F.9) that for any $i \in \{1, \ldots, n-1\}$,

$$
R_n(1) - R_i(1) = l_c^2 - l_p^2 + (n-2)a_{cp}^2. \tag{12}
$$

Therefore, if $l_p \leq l_c$, then $R_n(k) > R_i(k)$ for all $k \geq 1$, i.e., the centre node is the optimal control node at all times. However, when $l_c < l_p$, the network can exhibit different behaviours. From (F.9), we can also see that

$$\lim_{k \to \infty} R_n(k) > \lim_{k \to \infty} R_i(k) \Leftrightarrow \lambda_1 - l_p > a_{cp}. \tag{13}$$

Define $\underline{l}_p = \sqrt{l_c^2 + (n-2)a_{cp}^2}$ and $\bar{l}_p = l_c + (n-2)a_{cp}$. Using (12)–(13) and after some computations, one can see that

$$r(k) = n \text{ for all } k, \qquad\qquad\qquad\qquad\qquad \text{if } l_p \leq \underline{l}_p,$$

$$r(1) = \{1, \ldots, n-1\} \text{ but } r(k) = n \text{ for large enough } k, \quad \text{if } \underline{l}_p < l_p < \bar{l}_p,$$

$$r(k) = \{1, \ldots, n-1\} \text{ for all } k, \qquad\qquad\qquad\quad \text{if } l_p \geq \bar{l}_p.$$

In other words, when the manipulation is weak, the (latent) centre node is the optimal control node at all times. As the manipulation gains strength, scale-heterogeneity emerges, making the (manifest) peripheral nodes the optimal control node at local scales, while the centre node remains still the optimal control node at global scales. Finally, when the manipulation is strong enough, scale-heterogeneity vanishes, leaving the (manifest) peripheral nodes as the optimal control nodes at all scales. Notice that with the terminology of Theorem 4.5,

$$E = \left[ \begin{array}{cc} I_{n-1} & 0 \\ 0 & 0 \end{array} \right], \quad n_1 = n-1, \quad \alpha = \underline{l}_p, \quad \text{and} \quad \overline{\alpha} = \bar{l}_p. \qquad \square$$

A fair concern, however, exists regarding the minimum size of the manipulation needed to make the TVCS all-manifest. If this is excessively high, the prescribed approach may be infeasible in practice. Nevertheless, among networks of various size and structure, random manipulations with norm of about 10% of the norm of $A$ are on average sufficient (Fig. 4). Here, we see that the largest manipulations are needed for manifest subnetworks of about 10% the total size of the network. This is because when the size of the manifest subnetwork is extremely small, manipulations are focused on this small subset of nodes and thus more efficient, while with extremely large manifest subnetworks, the majority of the nodes are accessible for control and there is little restriction on the TVCS.

Finally, Fig. 4 also shows the comparison, in terms of controllability, of the manipulation-based approach against the alternative approach of selecting an optimal TVCS with the additional constraint that control nodes must be manifest (without any manipulation of the dynamics), which results in a sub-optimal all-manifest TVCS. For the comparison to be fair, we normalize each network by its spectral radius (largest magnitude of its eigenvalues), and then compare the optimal value of their TVCS (equation (5)). We see that the amount of relative advantage produced by manifest subnetwork manipulation is comparable to the relative size of the manipulation, except for medium-sized manifest subnetworks ($5 \sim 20\%$ of nodes), where the manipulation advantage is about two times its size.

## 5. Case study: TVCS in synthetic and real networks

Here, we discuss the benefits of TVCS and its relation to network structure for several examples of synthetic and real networks. We start with the classical deterministic examples of undirected line, ring
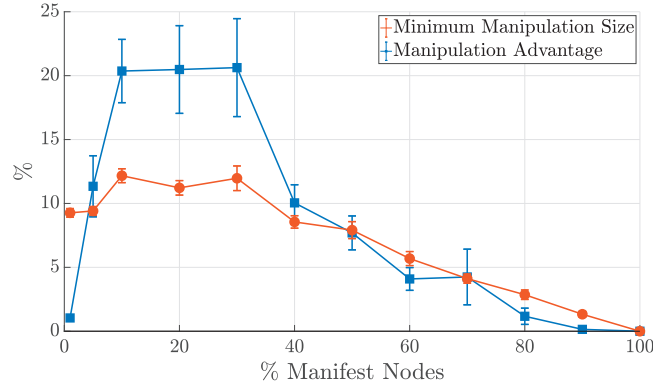
FIG. 4. Manipulation of manifest subnetworks in order to obtain an all-manifest optimal TVCS. The horizontal axis represents the percentage of manifest nodes in the network. In red, we show the minimum size of manipulation needed for the optimal TVCS to only include manifest nodes, relative to the size of the initial adjacency matrix (both measured by induced matrix 2-norm). In blue, we depict the optimal (i.e., maximal) value of $\mathrm{tr}(\mathcal{W}_K)$ for the case where the minimal manifest manipulation is applied, relative to the maximal value of $\mathrm{tr}(\mathcal{W}_K)$ subject to the constraint that all the control nodes are manifest (the former is with manipulation and without constraints on the control nodes, while the latter has no manipulation but control node constraints). Results are for $10^3$ random networks of logarithmically-uniform sizes in $[10^1, 10^3]$ but otherwise similar to Fig. 3. Markers (circles/squares) represent average values and error bars represent standard error of the mean (s.e.m). In both cases, the overall adjacency matrix is normalized by its spectral radius for fairness of comparison. We see that medium-sized manifest subnetworks ($5 \sim 20\%$) are the hardest yet most fruitful to manipulate.
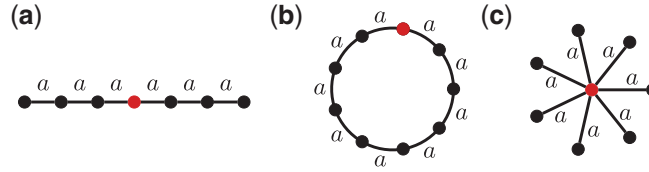


FIG. 5. Simple networks with closed-form $2k$-communicabilities. (a) A line network, (b) a ring network and (c) a star network. All networks are undirected and have homogeneous edge weights $a$. The $2k$-communicabilities of these networks are analytically computed (cf. Appendix F), concluding that all networks belong to class $\mathcal{I}$, with the optimal control node depicted in red in each case (the optimal control node is arbitrary in a ring network due to its symmetry).

and star networks (Fig. 5). Due to their simple structure, the $2k$-communicabilities of these networks can be analytically computed in closed form (cf. Appendix F). Using these results, it follows that for the line and star networks, the optimal control node is always the centre node (or any of the two centre nodes if a line has even number of nodes), while the optimal control node is arbitrary in a ring network. Notice that in all cases, it is the *homogeneity* of these networks that results in a single node having the greatest centrality at all scales (cf. Example 4.6 for non-homogeneous star networks that have scale-heterogeneous central nodes and thus belong to class $\mathcal{V}$).

Next, we analyse the role of TVCS in three classes of probabilistic complex networks that are widely used to capture the behaviour of various dynamical networks. These include the Erdös-Rényi (ER) random networks, Barabási–Albert (BA) scale-free networks, and Watts-Strogatz (WS) small-world networks. Each network has its own characteristic properties, and these properties lead to different behaviours under TVCS. The average $\chi$-values of these networks are computed for various values of $n$ and network parameters (Fig. 6). For ER networks, $\chi$ is in general small, and decays with $n$. This is because ER
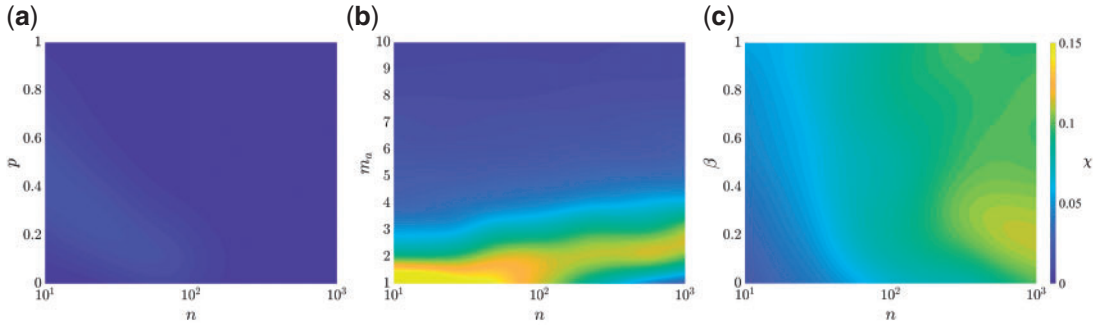
FIG. 6. The average $\chi$-value for (a) ER, (b) BA and (c) WS probabilistic networks. The horizontal axis determines the size of the network $n$ in all cases, while the vertical axis determines the values of the corresponding parameters for each network: edge probability $p$ for ER, growth (link attachment) rate $m_a$ for BA, and rewiring probability $\beta$ for WS. After constructing the unweighted connectivity according to each algorithm (ER, BA, or WS), standard uniformly random weights are assigned to each edge, which is then converted to $A$ using transmission method (cf. Appendix A). For each value of $n$ and network parameter over a coarse mesh ($\sim$ 100 points), 100 networks are generated and the average of their $\chi$-value is computed, which is then smoothly interpolated over a fine mesh (MATLAB `csaps`).

networks, especially when $n$ is large, are extremely homogeneous. This homogeneity is further increased during the transmission method, leading to a network matrix $A$ that is extremely insensitive to the choice of control nodes.

The connectivity structure of BA networks, in contrast, is extremely inhomogeneous, with one (sometimes 2) highly central nodes and a hierarchy down to peripheral leafs. As one would expect, this implies a small $\chi$-value since the centre node has the highest centrality at all scales (Fig. A.2). However, when the connectivity matrix is transformed to $A$ using the transmission method, the incoming links to all nodes are made uniform (adding up to 1). This in turns make the centrality levels of all the nodes comparable, leading to high $\chi$-values observed (notice that the underlying connectivity structures are still highly inhomogeneous, distinguishing them from the homogeneous ER networks). Notice that as the growth rate $m_a$ is increased, smaller networks tend towards complete graphs and high $\chi$ values *shift* to larger $n$.

As our last class of probabilistic networks, WS networks have the broadest range of size-parameter values with significant $\chi$. As one would expect, $\chi$ is low near $\beta = 0, 1$, corresponding to regular ring lattice and ER networks, respectively. For $\beta \sim 0.2$, there is a sufficiently high probability of having multiple nodes that are close to many rewired links (increasing their centrality), yet there is a low probability that these nodes, and the nodes close to them, are rewired all alike, resulting in heterogeneous central nodes and high $\chi$-values. This heterogeneity is increased with $n$ as larger networks have more possibilities of rewiring every edge.

Finally, we used the tools and concepts introduced so far to analyse TVCS in several real-world dynamical networks (Table 1). These networks are chosen from a wide range of application domains, from neuronal networks to transportation and social networks. According to the type of dynamics evolving over each network, we have used either the transmission or induction method to obtain its dynamical adjacency matrix from its static connectivity (the "$C \to A$" column, cf. Appendix A).

We have computed the $\chi$-value for each network using a variable time horizon $K \leq 50$, with the results ranging from 0 to more than 30% for different networks. These large variations even within each category signify both the potential benefits of TVCS and the possibility of its redundancy, a contrast that has been pivotal to our discussion. In particular, four facts about these results worth highlighting.

TABLE 1 *Characteristics of the real-world networks studied in the article*

| Category | Name | $n$ | $|E|$ | Directed | $C \rightarrow A$ | $\chi$(%) | $r(0) =$ $r(K-1)$ | Dominance of $r(0)$ ($\times 10^{-3}$) | ref. |
|---|---|---|---|---|---|---|---|---|---|
| Neuronal | BCTNet fMRI | 638 | 37 250 | N | T | 1.8 | N | N/A | [47] |
| | Cocomac | 58 | 1078 | Y | T | 5.5 | N | N/A | [48] |
| | BCTNet Cat | 95 | 2126 | Y | T | 1.9 | N | N/A | [47] |
| | C. elegans | 306 | 2345 | Y | T | 0 | Y | 0 | [49] |
| Transportation | air500 | 500 | 5960 | N | T | 22.4 | N | N/A | [50] |
| | airUS | 1858 | 28 236 | Y | T | 0 | Y | 0 | [51] |
| | airGlobal | 7976 | 30 501 | Y | T | 0 | Y | 0 | [51] |
| | Chicago | 1467 | 2596 | N | T | 0 | Y | 0 | [52, 53] |
| Gene Regulatory | *E. coli* | 4053 | 127 544 | N | T | 0 | Y | 0 | [54] |
| PPI | Yeast | 2361 | 13 828 | N | T | 0 | Y | 0 | [55] |
| | Stelzl | 1706 | 6207 | Y | T | 0 | Y | 0 | [56] |
| | Figeys | 2239 | 6452 | Y | T | 0 | Y | 0 | [57] |
| | Vidal | 3133 | 12 875 | N | T | 0 | Y | 0 | [58] |
| Power | WesternUS | 4941 | 13 188 | N | T | 33.7 | Y | 0 | [49] |
| Food | Florida | 128 | 2106 | Y | T | 34.6 | N | N/A | [59] |
| | LRL | 183 | 2494 | Y | T | 27.3 | N | N/A | [60] |
| Social | Facebook group | 4039 | 176 468 | N | I | 0.4 | N | N/A | [61] |
| | E-mail | 1005 | 25 571 | Y | I | 0 | Y | 40.5 | [62, 63] |
| | Southern Women | 18 | 278 | N | I | 0 | Y | 1.6 | [64] |
| | UCI P2P | 1899 | 20 296 | Y | I | 0 | Y | 5.5 | [65] |
| | UCI Forum | 899 | 142 760 | N | I | 0 | Y | 2.8 | [66] |
| | Freeman's EIES | 48 | 830 | Y | I | 0 | Y | 1.4 | [67] |
| | Dolphins | 62 | 318 | N | I | 0 | Y | 0.7 | [68] |
| Trust | Physicians | 241 | 1098 | Y | I | 8.8 | N | N/A | [69] |
| | Org. Consult Advice | 46 | 879 | Y | I | 0 | Y | 0.1 | [70] |
| | Org. Consult Value | 46 | 858 | Y | I | 0 | Y | 1.2 | [70] |
| | Org. R&D Advice | 77 | 2228 | Y | I | $6 \times 10^{-3}$ | N | N/A | [70] |
| | Org. R&D Aware | 77 | 2326 | Y | I | 0 | Y | 0.3 | [70] |

For each network, we have reported the number of nodes $n$, number of edges $|\mathcal{E}|$ (with each bidirectional edge counted twice), whether the network is directed, the method used for obtaining dynamical adjacency matrix $A$ from static connectivity $C$ ($A \rightarrow C$), the $\chi$ value (equation (6)), and whether the most local and global central nodes coincide ($r(0) = r(K-1)$). Since the value of $\chi$ is a function of $K$, we have chosen the value of $K \leq 50$ that has the largest $\chi$ for each network. Detailed descriptions of these datasets are provided in Appendix H.

(i) As measured by tr($\mathcal{W}_K$), the majority of networks tested do not benefit from TVCS, but a few do so significantly. (ii) Despite coming from various domains, the networks that do significantly benefit from TVCS share *scale-heterogeneity* as their common qualitative property (cf. Section 4.1). (iii) Networks with inductive $C \rightarrow A$ transformation benefit significantly less from TVCS than those with transmission

$C \rightarrow A$ transformation. (iv) Significantly higher values of $\chi$ are expected for all networks if using $\lambda_{\min}(\mathcal{W}_K)$ or similar measures for controllability, cf. Appendix B.

In the last column, we have also indicated whether the most local and most global central nodes coincide in each network. Recall that this is a sufficient but not necessary condition for a network to be in class $\mathcal{V}$ (Theorem 4.3 and Fig. 3). Though only sufficient, this simple metric can correctly classify class members of $\mathcal{V}$ from $\mathcal{I}$ among these networks, except for the WesternUS power network, for which $r(0) = r(K-1)$ only marginally holds (the dominance of $r(0)$ is 0) (cf. Fig. 3(b)).

## 6. Conclusions and discussion

Despite the breadth and depth of existing literature on the controllability of complex networks and control scheduling, the significant potential of TVCS has been greatly overlooked. This work strives to explore the advantages of TVCS in linear dynamical networks and obtaining theoretical and computational relationships between these advantages and network structure. Using Gramian-based measures of controllability, we showed that TVCS can significantly enhance the controllability of *many but not all* synthetic and real networks. This motivated the pursuit of identifying properties based on network structure that explain when, why and by how much TVCS is beneficial.

Using the newly introduced notion of $2k$-communicability, we showed that the scale-heterogeneity of central nodes in a network is the main cause and correlate of TVCS advantages. If a network has several distinct central nodes at different scales, the optimal TVCS involves starting the control from the most global central nodes and gradually moving towards most local ones as the time horizon is approached. If, on the other hand, a single node acquires the highest centrality at all scales, optimal TVCS prescribes the sole control of this node over the entire horizon, leading to optimality of TICS.

A striking finding that defied our expectations is the effect of network dynamics, beyond its raw connectivity structure, on TVCS. Here, we differentiated between the raw connectivity structure of a network (obtained using specific field knowledge and measure the *relative* strength of nodal connections) and its dynamical adjacency matrix which determines the evolution of network state over time. Depending on the nature of network state, we proposed two methods, transmission and induction, for obtaining the dynamical adjacency matrix from static connectivity. The effects of these methods, however, is noteworthy on the benefits of TVCS, even though the underlying network connectivity is the same (Table 1 and Fig. A.2). While the transmission method significantly enhances the merit of TVCS, the induction method depresses it (both compared with raw connectivity). We believe the reason for the former is the additional *homogeneity* that the transmission method introduces among the nodes, while the latter is due to the conversion from continuous to discrete-time dynamics, which enables long-distance connections even over small sampling times (due to the fact that interactions occur over infinitesimal intervals in continuous time) (cf. Section A and Fig. A.1). These results suggest that controllability of network dynamics is not only a function of its structural connectivity, but also greatly relies on the type of dynamics evolving over the network, an aspect that has received little attention in the existing literature and warrants future research.

Our discussion so far applies to networks with and without self-loops alike. However, it follows from the results in Section 4 that self-loops play an important role in TVCS. This is because (i) the self-loop of each node directly adds to its $2k$-communicability for all $k$, and (ii) the self-loop of each node also contributes indirectly to the $2k$-communicability of its neighbours less than $k-1$ hops away. As a result, the self-loop of any node has the largest effect on its own $2k$-communicability for all $k$, but also a lesser effect on the $2k$ communicability of all other nodes in the network. This latter effect becomes smaller and limited to higher $k$ for more distant nodes. A clear demonstration of the effects of self-loops can be

seen in Example 4.6, where as the self-loops of the peripheral nodes get stronger, they gradually become the central nodes in the network, first at local scales (small $k$) and eventually at all scales.

Further, the focus of our discussion has so far been on single input networks where one node is controlled at a time, in order to enhance the simplicity and clarity of concepts. Nevertheless, our results have straightforward generalizations to multiple-input networks (cf. Appendix E). If $m$ denotes the number of control inputs, the optimal TVCS involves applying these control inputs to the $m$ nodes with the highest centralities at the appropriate scale at every time instance (i.e., the $m$ nodes with the largest $R_i(K - 1 - k)$ have to be controlled at every time instance $k$). It is clear that the additional flexibility due to the additional inputs makes $\mathcal{V}$ larger, i.e., more networks have $\chi > 0$. Nevertheless, this additional flexibility also makes TICS significantly more efficient. Therefore, it is not immediately clear whether this enlargement of $\mathcal{V}$ also entails larger $\chi$ for networks with the same size and sparsity. In fact, increasing $m$ reduces average $\chi$ for all the classes of ER, BA and WS networks (Fig. E.1), suggesting that the additional flexibility is more advantageous for TICS than TVCS.

Regardless of the number of inputs (1 or more), an important implicit assumption of TVCS is that this number is limited, i.e., no more than $m$ nodes can be controlled at every time instance. This may at first seem over-conservative since TVCS requires, by its essence, the installation of actuators at all (or many) nodes of the network. Therefore, one might wonder why limit the control to only $m$ nodes at every time instance when all the nodes are ready for actuation. The answer lies within the practical limitations of actuators. For ideal actuators, distributing the control energy over as many nodes as possible is indeed optimal. However, this is not possible in many scenarios, including when (i) actuators exhibit nonlinear *dead-zone* behaviours, so that each one requires a sizable activation energy. In many applications ranging from distributed industrial processes to opinion dynamics in social networks, nodes cannot be actuated with arbitrarily small amounts of control energy. If $E_{\min}$ is the minimum activation energy of any actuator, at least $mE_{\min}$ is required for actuation of $m$ nodes at a time. Thus, when $E_{\min}$ is sizable and $n$ is large, simultaneous actuation of all nodes ($m = n$) requires a significant amount of control energy which is often infeasible (notice that the dead-zone behaviour of actuators does not violate the linearity assumption in (1) as one can replace $u$ with $v = \phi(u)$, where $\phi$ denotes dead-zone nonlinearity); (ii) actuators are geographically disperse so that precise coordination becomes difficult or time-consuming. A familiar example of this is the social opinion dynamics in pre-election times during political campaigns, where rallies and speeches by candidates act as control inputs to the network. Even though all nodes may be actuatable, at most one node can be actuated at every time; (iii) simultaneous control of proximal nodes results in actuator interference. This is the case in many biological networks. In neuronal networks, for instance, common control technologies such as transcranial magnetic stimulation (TMS) do not allow for simultaneous actuation of all cortical areas due, in part, to electromagnetic interference between multiple sources of actuation (note that TVCS is still possible by installation and sequential activation of multiple coils at different locations); and when (iv) actuators are controlled via communication channels with limited capacity, so that only a small number of devices can be simultaneously operated. This may be the case in industrial applications where large numbers of geographically distributed actuators are remotely (and centrally) controlled over shared communication channels with limited bandwidth. In all these scenarios, TVCS has the potential to significantly enhance network controllability, conditioned on the scale-heterogeneity of the central nodes in the network.

Although the dynamics of all real networks have some degrees of nonlinearity, the analysis of linear(ized) dynamics is a standard first step in analysis of dynamical properties of complex networks [2–9, 14–17]. This is mainly due to the fact that stability and controllability of linearized dynamics of a nonlinear network implies the same properties *locally* for the original nonlinear dynamics, making linear dynamics a powerful tool in analysing many dynamical properties that are in general intractable for

nonlinear dynamics. The local validity of linearization, however, is a main limitation of this work, particularly in networks where the change of state is significant relative to the size of the domain over which the linearization is valid. For these networks, whether the nonlinearity enhances or decreases the benefits TVCS with respect to its linearization is in general dependent on the type of nonlinearity. However, for saturation nonlinearities, being perhaps the most widespread, we expect TVCS to be more beneficial than linear counterparts. This is because in TICS all the control input is injected through a fixed node, requiring the state of that node to potentially undergo large over- and undershoots in order to convey sufficient input to the rest of the network. Saturation clearly prevents this from happening, further limiting the scope of TICS. The generalization of this work to nonlinear dynamics with saturation and linear *time-varying* dynamics (namely, $A(k)$ instead of $A$ in equation (1)) is a warranted next step for future exploration of the role of TVCS in general nonlinear networks.

## Funding

## Appendix

### A. Obtaining dynamical adjacency matrix from static connectivity

A standard starting point for the analysis of network dynamics of the form (1) is the assumption that the network adjacency matrix $A$ is known. While this is a valid assumption (as the construction of $A$ is itself the subject of vast research in network identification and corresponding field sciences), care should be taken in how one interprets raw network connectivity matrices. Usually, the network structure is described not by its dynamic adjacency matrix $A$ (which determines the evolution of network *state* according to (1)) but rather by its static connectivity matrix $C$ (our implicit assumption is that each node has a well-defined state that evolves over time through network dynamics, so our discussion is not applicable to completely static networks). While for any $i,j \in \mathcal{N}$, $a_{ij}$ describes the impact of $x_j$ on $x_i$ over one time step (relative to $x_j$), $c_{ij}$ often describes the strength of the link $(i,j)$ in arbitrary units (e.g., number of synapses between two neurons, capacity of high-voltage lines between two generators, or number of seats on a flight). In particular, multiplying $C$ by a positive constant results in an equivalent description of the network structure, yet multiplying $A$ by a constant significantly alters network dynamics. Here, we outline two methods for obtaining $A$ from $C$, and describe example domains where each method seems more relevant. Consider an arbitrary link $(i,j) \in \mathcal{E}$.

- **Transmission:** This method applies to dynamical networks where at each time step, the value of the state of node $i$ is itself affected (reduced) as a result of interaction with neighbour node $j$. Here, the state of each node corresponds to a physical quantity that is *transmitted* to its neighbours in order to affect their states. Neuronal, transportation, food, gene regulatory, protein–protein interaction and power networks are all examples of this type of interaction. If the sampling time is chosen long enough such that "current" state of a node is completely diffused through the network until the next time step, we can obtain $A$ from $C$ using
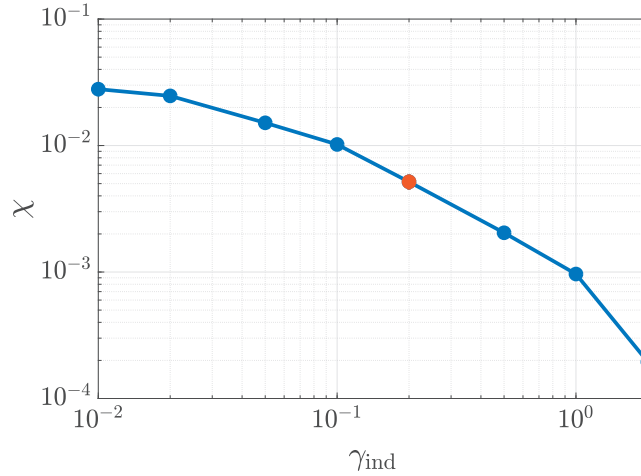
$$A = CD_{C,\text{in}}^{-1},$$

FIG. A.1. The average value of $\chi$ for the induction method and varying values of $\gamma_{\text{ind}}$ (corresponding to varying discretization step sizes $T_s$). Each point represents the average value of $\chi$ for 50 realizations of ER networks with $n = 100$ and $p = 0.2$ and vertical bars (when visible) show one standard error of the mean (s.e.m.). For each network, the value of $K \leq 10^3$ that gives the largest value of $\chi$ is chosen. The average value of $\chi$ drops with $\gamma_{\text{ind}}$, showing the effect of discretization on $\chi$ and the merit of TVCS. The red point corresponds to $\gamma_{\text{ind}} = 0.2$ used throughout this work.

where $D_{C,\text{in}}$ is the *augmented in-degree matrix* of $C$ (a diagonal matrix with the sum of the columns of $C$ on its diagonal, except where the sum of a column of $C$ is zero, in which case the corresponding diagonal element of $D_{C,\text{in}}$ is 1). This means that over each time step, $x_i$ is transmitted to the in-neighbours of node $i$ proportionally to their connectivity strength, if $i$ has any in-neighbours, and preserved otherwise.

- **Induction:** This method is appropriate for networks in which nodal states are not physical quantities and thus do not reduce as a result of network interactions. Opinion or epidemic dynamics evolving over social and/or trust networks have such properties. Here, in order to compute $A$ from $C$, we start from the underlying continuous-time dynamics $\dot{x} = (-\alpha I + C)x$ where $\alpha > 0$ is chosen such that $-\alpha I + C$ is stable (Hurwitz), and then discretize it to obtain (1), where

$$A = e^{(-\alpha I + C)T_s},$$

and $T_s$ is the sampling time [24, eq. (4.17)]. From the expansion of matrix exponential ($e^M = I + M + \frac{M^2}{2} + \frac{M^3}{3!} + \cdots$), we see that $A$ does not inherit the sparsity pattern of $C$ (and $G$) since nodes interact in continuous time. However, if $\|(-\alpha I + C)^2 T_s^2/2\| \ll \|(-\alpha I + C)T_s\|$, then the sparsity pattern of $C$ is almost preserved in $A$. Therefore, in this work we use $T_s = \gamma_{\text{ind}}/\|\alpha I + C\|$ for the induction method with $\gamma_{\text{ind}} = 0.2$ unless otherwise stated. Further, Fig. A.1 shows the effect of $\gamma_{\text{ind}}$ on the value of $\chi$ when using the induction method. As expected, the larger $\gamma_{\text{ind}}$, the larger $T_s$, the closer $A$ gets to $\lim_{k \to \infty} A^k$, the more similar $2k$-communicabilities for different $k$ become, and the smaller $\chi$ becomes.

Unless otherwise stated, we use the transmission method in this work. Nevertheless, it is to be noted that the method used for obtaining $A$ from $C$ can have profound effects on network controllability and should thus be chosen carefully. Figure A.2 illustrates this concept by showing the mean $\chi$-value of ER, BA and WS networks for a number of different choices for this transformation.
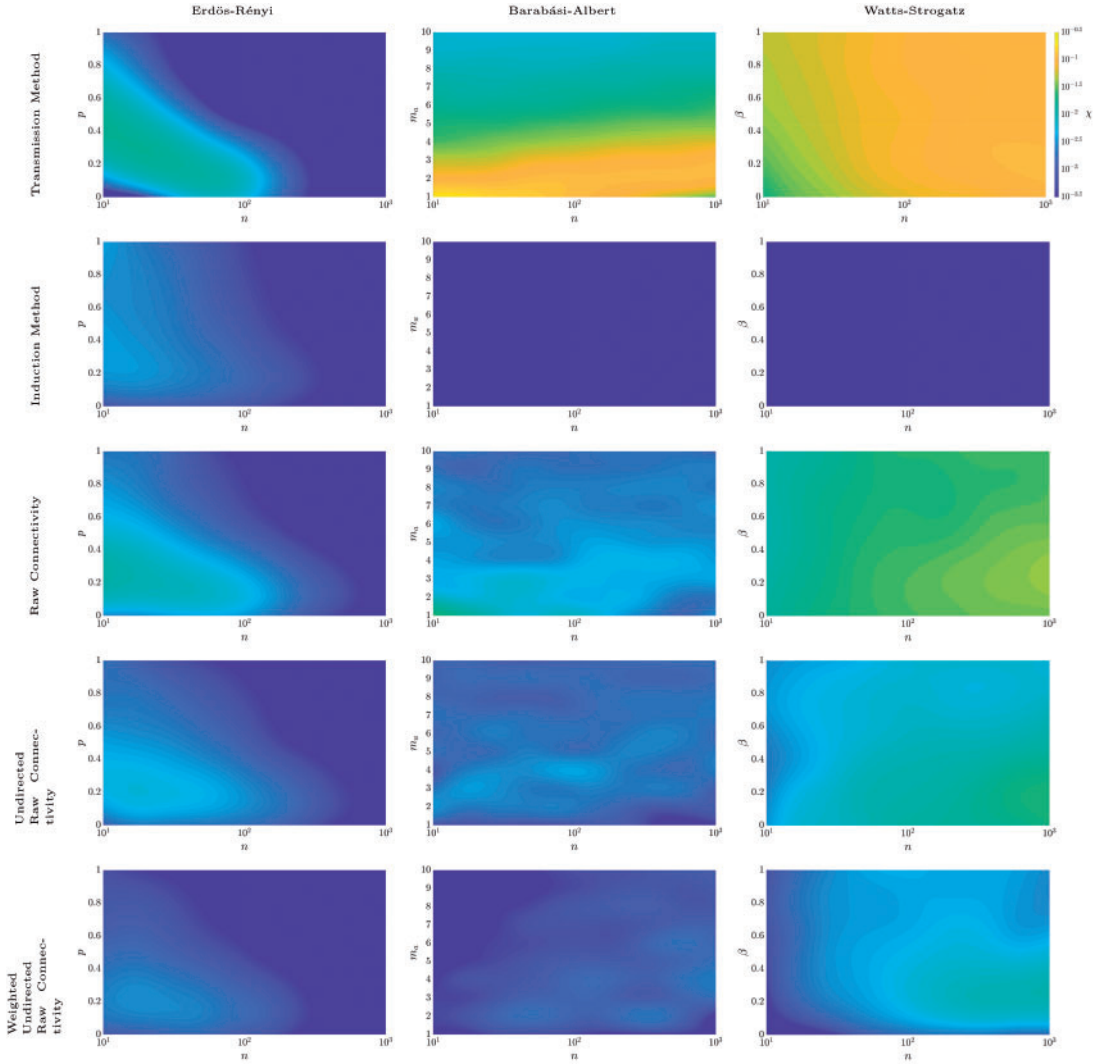
FIG. A.2. Average value of $\chi$ for different methods of obtaining dynamical adjacency matrix $A$ from static connectivity $C$. The plots show the effect of these methods on TVCS. The details on how to obtain the plots are similar to Fig. 6 in the main text. All matrices are normalized by their spectral radius for uniformity and comparison. The plots show a sizable enhancement (respectively depression) of $\chi$ by the transmission (respectively induction) method compared with raw connectivity, except for Erdös-Rényi networks whose $\chi$ maintains a robust pattern irrespective of the method of obtaining the dynamic adjacency matrix $A$ from the raw static connectivity $C$.

## B.  Comparison between Gramian-based measures of controllability

In this section, we first derive and elaborate on the relationship between the eigenvalues of the Gramian and control energy. Then, we discuss the different Gramian-based measures of controllability and their respective properties.

Assume that $\mathcal{W}_K$ is invertible (the network dynamics (1) is controllable). Then, for any $x_f \in \mathbb{R}^n$, among the (usually infinitely many) choices of $\{u(k)\}_{k=0}^{K-1}$ that take the network from $x(0) = 0$ to $x(K) = x_f$, the one that has the smallest energy is given by [24, Theorem 6.1]

$$u^*(k) = b(k)^T (A^T)^{K-1-k} \mathcal{W}_K^{-1} x_f, \quad k \in \{0, \dots, K-1\}.$$

Similar expression holds for arbitrary $x_0$, but it is customary to evaluate control energy starting from the network's unforced equilibrium $x = 0$. It is immediate to verify that this gives the minimal energy $\sum_{k=0}^{K-1} u^{*2}(k) = x_f^T \mathcal{W}_K^{-1} x_f$. Therefore, the unit-energy reachability set is given by

$$\{x_f \in \mathbb{R}^n \mid x_f^T \mathcal{W}_K^{-1} x_f \leq 1\}.$$

Since $\mathcal{W}_K^{-1}$ is positive definite, this is a hyper-ellipsoid in $\mathbb{R}^n$, with axes aligned with the eigenvectors of $\mathcal{W}_K$. Let $(\lambda_i, v_i)$ be an eigen-pair of $\mathcal{W}_K$ and $x_f = cv_i$. Then,

$$x_f^T \mathcal{W}_K^{-1} x_f \leq 1 \Leftrightarrow c^2 \lambda_i^{-1} \leq 1 \Leftrightarrow |c| \leq \lambda_i^{1/2},$$

showing that the axis lengths of this hyper-ellipsoid are given by the square roots of the eigenvalues of $\mathcal{W}_K$. Intuitively, the "larger" the reachability hyper-ellipsoid, the "more controllable" the network dynamics (equation (1)) are. To quantify how large the hyper-ellipsoid is, several measures based on the eigenvalues of $\mathcal{W}_K$ have been proposed in the literature [6, 8, 71]. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ denote the eigenvalues of $\mathcal{W}_K$. The most widely used Gramian-based measures are

- $\text{tr}(\mathcal{W}_K) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$,
- $\text{tr}(\mathcal{W}_K^{-1})^{-1} = (\lambda_1^{-1} + \lambda_2^{-1} + \cdots + \lambda_n^{-1})^{-1}$,
- $\det(\mathcal{W}_K) = \lambda_1 \lambda_2 \cdots \lambda_n$,
- $\lambda_{\min}(\mathcal{W}_K) = \lambda_n$.

It is clear from these relationships that all these measures, except for $\text{tr}(\mathcal{W}_K)$, approach 0 if $\lambda_n \to 0$. This property, i.e., the behaviour of a measure as $\lambda_n \to 0$, is the most critical difference between $\text{tr}(\mathcal{W}_K)$ and the other three measures. For the rest of this discussion, let $f_c(\cdot)$ be any of $\text{tr}((\cdot)^{-1})^{-1}$, $\det(\cdot)$, or $\lambda_{\min}(\cdot)$. Since the network is (Kalman-) controllable if and only if $\lambda_n > 0$, having $f_c(\mathcal{W}_K) > 0$ guarantees network controllability while $\text{tr}(\mathcal{W}_K) > 0$ does not. This is a major disadvantage of $\text{tr}(\mathcal{W}_K)$ for small networks, where controllability in all directions in state space is both achievable and desirable. As the size of the network grows, however, $\lambda_n$ typically decays exponentially fast to zero [6], irrespective of network structure. This exponential decay of worst-case controllability is even evident in the example network of Fig. 1(a) comprising of only $n = 5$ nodes.

Computationally, this means that $\lambda_n$ (and in turn $f_c(\mathcal{W}_K)$) can quickly drop below machine precision as $n$ grows. In fact, for $K = 10$ and double-precision arithmetics, this happens for $n \sim 15$, making the TVCS (equation (5)) with $f = f_c$ numerically infeasible (as it involves the comparison of $f_c(\mathcal{W}_K)$ for different $\{b_k\}_{k=0}^{K-1}$, which may be zero up to machine accuracy). Further, notice that the computational complexity of TVCS for $f = f_c$ grows as $n^K$ due to the NP-hardness of TVCS, enforcing the use of sub-optimal greedy algorithms even if machine precision was not a concern (see [16] and the references therein for details).

In addition to the computational aspects of TVCS, the exponential decay of $\lambda_n$ also has theoretical implications for the choice of $f$. When using $f = f_c$, TVCS seeks to assign the control nodes $\{\iota_k\}_{k=0}^{K-1}$ such that controllability is maintained in all directions in the state space, with special emphasis on the hardest-to-reach directions. The use of $\text{tr}(\mathcal{W}_K)$, on the other hand, involves maximizing the average of Gramian eigenvalues, which usually strengthens the largest eigenvalues and spares the few smallest ones. In large networks, the latter is in general more realistic as controllability is hardly needed in all $n$ directions of the state space. As discussed in detail in [72], this seems to be the case in the resting-state structural brain networks: this article shows that $\text{tr}(\mathcal{W}_K)$ is maximized by controlling specific brain regions that have long been identified as the structural 'core' or 'hubs' of the cerebral cortex, while the Gramian is itself close to singular.

Further, due to the same strong dependence of $f_c(\mathcal{W}_K)$ but not $\text{tr}(\mathcal{W}_K)$ on $\lambda_n$, we often observe that $\text{tr}(\mathcal{W}_K)$ is significantly less sensitive to the choice of the control nodes $\{\iota_k\}_{k=0}^{K-1}$, leading to orders of magnitude smaller $\chi$ than that of $f_c(\mathcal{W}_K)$ (Fig. 1(b)). This means that $\mathcal{V}$ is only a small subclass of networks that benefit from TVCS measured by $f_c$. This also has a clear interpretation, since maintaining controllability in all directions in the state space requires a broader distribution of the control nodes that facilitates the reach of the control action $\{u(k)\}_{k=0}^{K-1}$ to all the nodes in the network.

Finally, we highlight the need for development and analysis of measures that are neither strongly reliant on the least controllable directions (such as $f_c(\mathcal{W}_K)$) nor mainly ignore them (such as $\text{tr}(\mathcal{W}_K)$). Two such candidates are:

- $\text{tr}(C^T \mathcal{W}_K C)$ where $C$ is a matrix (or vector) with columns that point towards some particular directions of interest in the state space. This measure is a modular set function similar to $\text{tr}(\mathcal{W}_K)$ [8], but the extensions of the notion of $2k$-communicability and the relationship between class $\mathcal{I}/\mathcal{V}$ networks and scale-heterogeneity are unclear;

- appropriate approximations of $\log(f_c(\mathcal{W}_K))$. While computing the exact value of $\log(f_c(\mathcal{W}_K))$ is subject to the same issues as $f_c(\mathcal{W}_K)$ itself, approximations can be used that provide a *mitigation* of the effects of the smallest eigenvalues of $\mathcal{W}_K$. In the case of $f_c(\cdot) = \det(\cdot)$, e.g., various algorithms have been proposed to approximate log det of large matrices, see e.g., [73–79]. These algorithms, however, are predominantly designed with the aim of reducing the computational complexity of determinant calculation and not mitigation of the effects of its high condition number, and often rely on assumptions (such as sparsity or knowledge of lower and upper bounds on matrix eigenvalues) that do not apply to $\mathcal{W}_K$. Thus, development of *appropriate* approximations of $\log(f_c(\mathcal{W}_K))$ constitutes a warranted direction for future research.

## C. Relationships between $2k$-communicability, degree and eigenvector centrality

The notion of $2k$-communicability introduced in this article has close connections with the degree and eigenvector centrality in the limit cases of $k = 1$ and $k \to \infty$, respectively. Recall that the out-degree centrality and 2-communicability of a node $i \in \mathcal{N}$ are defined as, respectively,

$$d_i^{\text{out}} = \sum_{j=1}^{n} a_{ji},$$

$$R_i(1) = \sum_{j=1}^{n} a_{ji}^2.$$

Therefore, if the network is unweighted (i.e., all the edges have the same weight), then $R_i(1) \propto d_i^{\text{out}}$, so 2-communicability and out-degree centrality result in the same ranking of the nodes (in particular, $r(1)$ is the node with the largest out-degree). As edge weights become more heterogenous, these two rankings become less correlated, with 2-communicability putting more emphasis on stronger weights.

A similar relation exists between $\infty$-communicability and left eigenvector centrality, as we show next. Let $v_1, u_1 \in \mathbb{R}^n$ be the right and left Perron-Frobenius eigenvectors of $A$, respectively, normalized such that $v_1^T v_1 = u_1^T v_1 = 1$ (notice that $u_1$ has unit inner product with $v_1$ but does *not* in general have unit length). Since the network is by assumption strongly connected and aperiodic, we have

$$\lim_{k \to \infty} \left( \frac{1}{\rho(A)} A \right)^k = v_1 u_1^T. \tag{C.1}$$

Thus for any $i \in \mathcal{N}$,

$$\lim_{k \to \infty} \left( \frac{1}{\rho(A)} \right)^{2k} R_i(k) = \lim_{k \to \infty} \left( \frac{1}{\rho(A)} \right)^{2k} \left( (A^k)^T A^k \right)_{ii} = (u_1 v_1^T v_1 u_1^T)_{ii} = u_{1,i}^2.$$

Given that dividing $R_i(k)$ by $\rho(A)^{2k}$ for all $i$ does not change the ranking of nodes, we define $R_i(\infty) = u_{1,i}^2$ for all $i$. Since squaring non-negative numbers preserves their order, nodal rankings based on $\infty$-communicability and left eigenvector centrality are identical.

## D. Nodal dominance

Among the networks where the nodes with the greatest $R_i(1)$ and $R_i(\infty)$ coincide (i.e., $r(0) = r(\infty)$), there is a higher chance (than in general) that any network belongs to class $\mathcal{I}$. However, about half of these networks still belong to class $\mathcal{V}$, meaning that there exists $1 < k < \infty$ such that $r(k) \neq r(0)$. To assess the importance of this time-variation of optimal control nodes, we define the *dominance* of the node $r(0)$ (over the rest of the network) as follows. Let $r'(0)$ be the index of the node with the second largest $R_i(1)$ (largest after removing $r(0)$). Similarly, let $r'(\infty)$ be the index of the second largest $R_i(\infty)$. We define

$$\text{Dominance of } r(0) = \min \left\{ \frac{R_{r(0)}(0) - R_{r'(0)}(0)}{R_{r(0)}(0)}, \frac{R_{r(0)}(\infty) - R_{r'(\infty)}(\infty)}{R_{r(0)}(\infty)} \right\}.$$

A small dominance indicates that another node has very similar value $R_i(0)$ *or* $R_i(\infty)$ to $r(0)$, while a large dominance is an indication of a large gap between $R_{r(0)}(k)$ and the next largest $R_i(k)$ for both $k = 0$ *and* $k \to \infty$.

## E. Networks with multiple inputs

Consider a multiple-input network, namely, a network in which $m \geq 1$ nodes are controlled at every time step. Let $\iota_k^1, \ldots, \iota_k^m \in \mathcal{N}$ denote the indices of the control nodes at every time $k$, and $\boldsymbol{\iota}_k = \{\iota_k^1, \ldots, \iota_k^m\}$. Then, the corresponding TICS and TVCS are defined as

$$\max_{\boldsymbol{\iota}_0, \ldots, \boldsymbol{\iota}_{K-1} \in \mathcal{N}} \quad f(\mathcal{W}_K) \tag{E.1a}$$

$$\text{s.t.} \qquad \boldsymbol{\iota}_0 = \cdots = \boldsymbol{\iota}_{K-1} \tag{E.1b}$$

and

$$\max_{\iota_0,\ldots,\iota_{K-1}\in\mathcal{N}} f(\mathcal{W}_K),$$ (E.2)

respectively. Accordingly, a multiple-input network is said to belong to class $\mathcal{I}$ if the solution of (E.2) satisfies (E.1b), and to class $\mathcal{V}$ otherwise.

Clearly, for a multiple-input network to belong to class $\mathcal{V}$, any of the first $m$ largest of $\{R_i(k)\}_{i=1}^n$ should change over time, which is often implied by (but does not imply) a change in $r(k)$. Therefore, the condition of Theorem 4.3 can still be used as a tight proxy for networks in $\mathcal{V}$, but is too conservative and can be relaxed as follows: assume that $A$ is irreducible, aperiodic and diagonalizable. Let $\{r_j^d \in \mathbb{R}^n \mid j \in \mathcal{J}^d\}$ be the set of nodes with the $m$ highest 2-communicabilities, where the index set $\mathcal{J}^d$ accounts for different choices of rankings if there are nodes with equal 2-communicabilities. Similarly, let $\{r_j^c \in \mathbb{R}^n \mid j \in \mathcal{J}^c\}$ be the set of nodes with the $m$ highest centralities. Then, if $r_{j_1}^d \neq r_{j_2}^c$ for all $(j_1, j_2) \in \mathcal{J}^d \times \mathcal{J}^c$, the network belongs to class $\mathcal{V}$ when $K$ is sufficiently large. The proof of this statement is a straightforward generalization of the proof of Theorem 4.3 and thus omitted.

Similarly, the three conditions in Theorem 4.4 can be generalized to undirected multiple-input networks as follows (with similar proofs as the proof of Theorem 4.4):

*(i)* For all $i \in \{1, \ldots, m\}$,

$$\frac{1 - w_{i1}}{\sum_{\ell \leq m+1, \ell \neq i+1} w_{\ell 1}} \leq \frac{|\lambda_1| - |\lambda_2|}{|\lambda_1| - |\lambda_n|}.$$

This condition can be simplified, at the expense of being more conservative, to $\frac{1-w_{i1}}{iw_{i1}} \leq \frac{|\lambda_1|-|\lambda_2|}{|\lambda_1|-|\lambda_n|}$, for all $i \in \{1, \ldots, m\}$,

*(ii)* for all $i \in \{1, \ldots, m\}$, $w_{i2} = 1 - w_{i1}$,

*(iii)* the network has three or fewer nonzero eigenvalues with different absolute values and

$$R_1(1) \geq R_2(1) \geq \cdots \geq R_m(1) \geq R_i(1),$$

for all $i \in \{m + 1, \ldots, n\}$.

Finally, Fig. E.1 illustrates the effect of $m$ on $\chi$-values of ER, BA and WS networks discussed in the main text.

## F. 2k-Communicabilities of simple networks

As mentioned in the main text, cf. Fig. 5, the simple structure of homogeneous undirected line, ring and star networks allows us to compute their 2$k$-communicabilities analytically in closed form, as derived in the following. Throughout, $\mathbb{Z}$ denotes the set of integers and for $a, b \in \mathbb{Z}$, $a \mid b$ denotes that $a$ divides $b$.

PROPOSITION F.1 *(2k-communicabilities of line networks).* Consider a line network of $n$ nodes with uniform link weights $a$ (and no self-loops). Then, for $i \in \mathcal{N}$ and $k \in \mathbb{N}$,

$$R_i(k) = a^{2k} \sum_{p\in\mathcal{I}}\left[\binom{2k}{k+p(n+1)} - \binom{2k}{k+p(n+1)-i}\right],$$ (F.1)
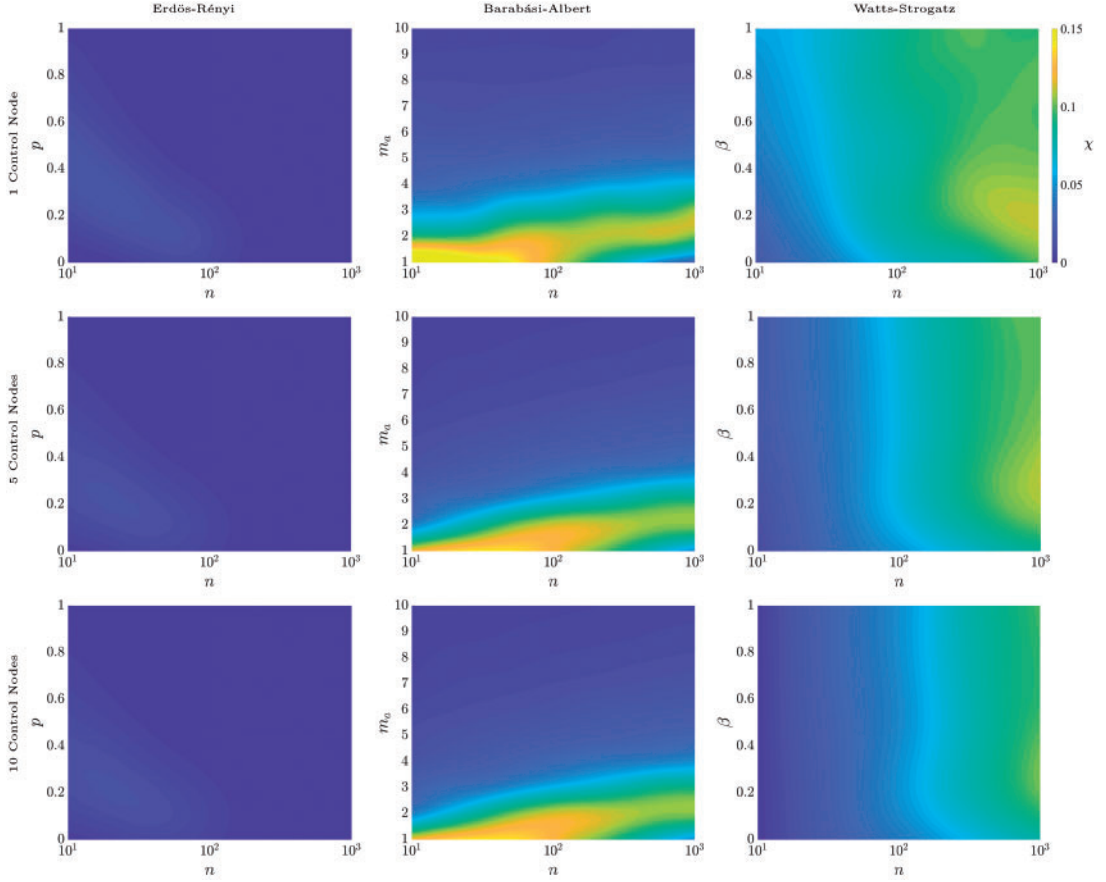
FIG. E.1. Average value of $\chi$ for networks with increasing number of inputs. The plots show the effect of multiple inputs on TVCS. The details on how to obtain the plots are similar to Fig. 6 in the main text. The dynamic adjacency matrix $A$ is obtained from the raw static connectivity $C$ using the transmission method in all cases. These plots show a slight depression in the benefit of TVCS as the number of control nodes grows, despite the fact that networks with more control nodes have a higher probability of belonging to $\mathcal{V}$ (namely, having $\chi > 0$).

where $\mathcal{I} = \{-\lceil \frac{k}{n+1} \rceil, \ldots, \lceil \frac{k}{n+1} \rceil\}$ and $\binom{n}{k} \triangleq 0$ if $k \notin \{0, \ldots, n\}$. In particular, if $i \leq \lceil \frac{n}{2} \rceil$ and $k \leq \lceil \frac{n}{2} \rceil - 1$,

$$R_i(k) = a^{2k} \left[ \binom{2k}{k} - \binom{2k}{k-i} \right].$$

(F.2)

*Proof.* From [80, Lemma 1.77], we have

$$\lambda_j = 2a \cos \frac{j\pi}{n+1} \quad \text{and} \quad w_{ij} \propto \sin^2 \frac{ij\pi}{n+1},$$

for $i,j \in \{1,\ldots,n\}$ where $\propto$ accounts for normalization. In order to normalize the eigenvectors, we use the identities $\sin^2 \alpha = \frac{1}{2}(1 - \cos 2\alpha)$ and

$$\sum_{j=1}^{n} \cos \frac{2sj\pi}{n+1} = -1 \quad \text{for all } s \nmid n+1, \tag{F.3}$$

to get $w_{ij} = \frac{2}{n+1} \sin^2 \frac{ij\pi}{n+1}$ for all $i,j \in \{1,\ldots,n\}$ (one can show (F.3) by multiplying and dividing the LHS by $\sin \frac{s\pi}{n+1}$ and using the identity $2\sin\alpha\cos\beta = \sin(\alpha+\beta) + \sin(\alpha-\beta)$ for each term). Thus, by substitution, we have $R_i(k) = \frac{2a^{2k}}{n+1} \sum_{j=1}^{n} \tau_{ijk}^2$ where

$$\tau_{ijk} = 2^k \sin \frac{ij\pi}{n+1} \cos^k \frac{j\pi}{n+1}.$$

By using the identity $2\sin\alpha\cos\beta = \sin(\alpha+\beta) + \sin(\alpha-\beta)$, $k$ times and collecting terms, we get

$$\tau_{ijk} = \sum_{\ell=0}^{k} \binom{k}{\ell} \sin \frac{(i+k-2\ell)j\pi}{n+1}.$$

Hence, by squaring $\tau_{ijk}$ and substituting it in $R_i(k)$, and using the identity $2\sin\alpha\sin\beta = \cos(\alpha-\beta) - \cos(\alpha+\beta)$, we get

$$R_i(k) = \frac{a^{2k}}{n+1} \sum_{\ell,r=0}^{k} \binom{k}{\ell}\binom{k}{r} \Big[ \sum_{j=1}^{n} \cos \frac{2(\ell-r)j\pi}{n+1} - \sum_{j=1}^{n} \cos \frac{2(i+k-\ell-r)j\pi}{n+1} \Big]. \tag{F.4}$$

However, by (F.3), the two sums in (F.4) cancel each other unless $\ell - r \,|\, n+1$ or $i+k-\ell-r \,|\, n+1$ (the cases where both of these happen need not be excluded since they automatically cancel). Thus,

$$R_i(k) = a^{2k} \Big[ \sum_{\mathcal{I}_1} \binom{k}{\ell}\binom{k}{r} - \sum_{\mathcal{I}_2} \binom{k}{\ell}\binom{k}{r} \Big], \tag{F.5}$$

where

$$\mathcal{I}_1 = \{(\ell,r) \in \{0,\ldots,k\}^2 \mid \ell - r \,|\, n+1\},$$
$$\mathcal{I}_2 = \{(\ell,r) \in \{0,\ldots,k\}^2 \mid i+k-\ell-r \,|\, n+1\}.$$

Defining $p = \frac{n+1}{\ell-r}$ in the first and $p = \frac{i+k-\ell-r}{n+1}$ in the second sum in (F.5), we get

$$R_i(k) = a^{2k} \sum_{p\in\mathcal{I}} \Big[ \sum_{\ell=0}^{k} \binom{k}{\ell}\binom{k}{\ell-p(n+1)} - \sum_{\ell=0}^{k} \binom{k}{\ell}\binom{k}{\ell+p(n+1)-i} \Big], \tag{F.6}$$

where we have used the identity $\binom{k}{s} = \binom{k}{k-s}$. Equation (F.1) then follows by applying the formula $\sum_{\ell=0}^{k} \binom{k}{\ell}\binom{k}{\ell\pm s} = \binom{2k}{k\pm s}$ [81, Eq. 6.69-70] to each of the two sums in (F.6). To get (F.2), note that if $i \leq \lceil \frac{n}{2} \rceil$ and $k \leq \lceil \frac{n}{2} \rceil - 1$, then the only nonzero term in (F.1) is the one corresponding to $p = 0$. $\qquad\square$

According to this result, in the case of no self-loops, the value of $R_i(k)$ increases with $i$ until $i = \lceil \frac{n}{2} \rceil$ (i.e., the middle node) for $k \le \lceil \frac{n}{2} \rceil - 1$ (this can be observed from the expression (F.2)). For general $k$, it can be shown that the value of the sum in (F.1) for $R_i(k)$ is strongly dominated by the summand corresponding to the index $p = 0$, which increases with $i$ until $i = \lceil \frac{n}{2} \rceil$ and decreases afterwards. Thus, **the optimal control node corresponds always to (one of) the centre node(s)**, i.e., $b^*(k) = e_{\lceil \frac{n}{2} \rceil}$ for all $k$. If nodes have uniform self-loops (i.e., self-loops all with the same weight), $R_i(k)$ can no longer be computed analytically but simulations show the exact same behaviour;

PROPOSITION F.2 (*2k-communicabilities of ring networks*). Consider a ring network of $n$ nodes and uniform link weights $a$ (with no self-loops). Then, for $i \in \mathcal{N}$ and $k \in \mathbb{N}$,

$$R_i(k) = \frac{(2a)^{2k}}{n} \left[ 1 + 2 \sum_{j=1}^{\lceil \frac{n}{2} \rceil - 1} \cos^{2k} \left( \frac{2j\pi}{n} \right) + \delta_n^{\mathrm{E}} \right], \tag{F.7}$$

where $\delta_n^{\mathrm{E}}$ equals one if $n$ is even and zero otherwise.

*Proof.* From [80, Lemma 1.77], we have $\lambda_j = 2a \cos \frac{2j\pi}{n}$ and (after normalization of eigenvectors),

$$w_{ij} = \begin{cases} \frac{2}{n} \cos^2 \frac{2(i-1)j\pi}{n} & \text{if } 1 \le j < \frac{n}{2}, \\ \frac{2}{n} \sin^2 \frac{2(i-1)(n-j)\pi}{n} & \text{if } \frac{n}{2} < j < n, \\ \frac{1}{n} & \text{if } j = n, \text{ or } n \in \mathbb{Z} \text{ even and } j = \frac{n}{2}, \end{cases}$$

for $i, j \in \{1, \ldots, n\}$. Note that to normalize the eigenvectors, we follow a similar procedure to the one described in the proof of Lemma F.1 (setting $s = 2i$ and substituting $n$ by $n - 1$ in (F.3)). The result then follows by substituting these expressions in $R_i(k)$. □

We can infer from the preceding result that without self-loops, the value of $R_i(k)$ is independent of $i$ (as shown by (F.7)) for a uniform ring network, so **the optimal control node is arbitrary for all** $k$. Similar result can be proved analytically if the nodes have uniform self-loops.

PROPOSITION F.3 (*2k-communicabilities of star networks*). Consider a star network given by

$$A = \begin{bmatrix} l_c & a^T \\ a & l_p I_{n-1} \end{bmatrix}, \tag{F.8}$$

where $a \in \mathbb{R}^{n-1}$ contains the link weights between the centre node and peripheral nodes. Then

$$R_1(k) = \frac{(\lambda_1 - l_p)^2}{(\lambda_1 - l_p)^2 + \|a\|^2} \lambda_1^{2k} + \frac{(l_p - \lambda_2)^2}{(l_p - \lambda_2)^2 + \|a\|^2} \lambda_2^{2k},$$

$$R_i(k) = \frac{a_{i-1}^2}{(\lambda_1 - l_p)^2 + \|a\|^2} \lambda_1^{2k} + \frac{a_{i-1}^2}{(l_p - \lambda_2)^2 + \|a\|^2} \lambda_2^{2k} + \frac{\|a\|^2 - a_{i-1}^2}{\|a\|^2} l_p^{2k}, \tag{F.9}$$

for all $k \in \mathbb{N} \cup \{0\}$ and $i \in \{2, \ldots, n\}$, where

$$\lambda_{1,2} = \frac{l_c + l_p \pm \sqrt{(l_c - l_p)^2 + 4\|a\|^2}}{2}. \tag{F.10}$$

*Proof.* Using the formula

$$\left| \begin{array}{cc} P & Q \\ R & S \end{array} \right| = (P - 1)|S| + |S - RQ|,$$

for scalar $P$, row vector $Q$, column vector $R$, and square matrix $S$, and some algebra, we get $|sI_n - A| = \left(s^2 - (l_c + l_p)s + l_c l_p - \|a\|^2\right)(s - l_p)^{n-2}$, so the eigenvalues of $A$ are given by

$$\lambda_{3,\ldots,n} = l_p, \tag{F.11}$$

and $\lambda_{1,2}$ in (F.10). Note that we may or may not have $|\lambda_1| \geq \cdots \geq |\lambda_n|$ as the order depends on the values of the parameter. By solving $(A - \lambda_j I_n)v_j = 0$ for $j = 1, 2$, and then using the orthogonality of eigenvectors, we get

$$v_{1,2} \propto \left[ \begin{array}{c} \lambda_{1,2} - l_p \\ a \end{array} \right], \quad (v_j)_1 = 0 \quad \forall j \in \{3, \ldots, n\}, \tag{F.12}$$

where $\propto$ accounts for normalization. The result then follows by substituting (F.10)-(F.12) into $R_i(1) = \sum_j v_{ij}^2 \lambda_j^2$ separately for $i = 1$ and $i \geq 2$, and simplifying. $\qquad \square$

Using this result, if all self-loop weights are the same ($l_c = l_p$ in (F.8)), then $R_1(1) > R_i(1)$ for all $i \geq 2$ from (12). Therefore Theorem 4.4(iii) implies that *the centre node is the optimal control node at all times*.

## G. Additional lemmas and proofs

In this section, we formulate and prove a number of lemmas that underlie the main results of this article and also provide the proofs of the main results presented in the main text. Throughout, $\mathbb{C}$ denotes the set of complex numbers and for $M \in \mathbb{C}^{n \times n}$, $\overline{M}$ and $M^*$ denote its complex conjugate and complex conjugate transpose, respectively, and $M^{-*} = (M^*)^{-1}$. Further, for $\lambda \in \mathbb{R}^n$ and $\ell \in \mathbb{N} \cup \{0\}$, $\lambda^\ell \triangleq [\lambda_1^\ell \; \cdots \; \lambda_n^\ell]^T$ and $|\lambda| \triangleq [|\lambda_1| \; \cdots \; |\lambda_n|]^T$.

***Proof of Theorem 4.3.*** Define

$$U = V^{-\star}.$$

Notice that the columns of $U$ are the left eigenvectors of $A$, with the same order as in $\Lambda$ and $V$. Since for any $k$,

$$(A^k)^T A^k = (A^k)^* A^k = (V \Lambda^k U^*)^* V \Lambda^k U^*,$$

it follows that for any $i$ and $k$,

$$R_i(k) = \left( (A^k)^T A^k \right)_{ii} = (V\Lambda^k U_{i,:}^*)^* V\Lambda^k U_{i,:}^* = \|V\Lambda^k U_{i,:}^*\|_2^2,$$

where $U_{i,:}$ denotes the $i$th row of $U$. For simplicity, define $c^{(i,k)} = V\Lambda^k U_{i,:}^* \in \mathbb{C}^n$. It is straightforward to check that

$$c_\ell^{(i,k)} = \sum_{j=1}^n v_{\ell j}\overline{u}_{ij}\lambda_j^k, \qquad \ell \in \{1,\dots,n\}$$

so

$$R_i(k) = \sum_{\ell=1}^n \left| c_\ell^{(i,k)} \right|^2 = \sum_{\ell=1}^n c_\ell^{(i,k)}\overline{c_\ell^{(i,k)}} = \sum_{\ell=1}^n \sum_{j=1}^n \sum_{m=1}^n v_{\ell j}\overline{v}_{\ell m}\overline{u}_{ij}u_{im}\lambda_j^k\overline{\lambda}_m^k$$

$$= \sum_{j,m=1}^n \underbrace{\left( \sum_{\ell=1}^n v_{\ell j}\overline{v}_{\ell m}\overline{u}_{ij}u_{im} \right)}_{\beta_i^{(j,m)}} \lambda_j^k\overline{\lambda}_m^k.$$

Dividing both sides by $\lambda_1^{2k}$ and taking the limits as $k \to \infty$, we see that for all $i$, $\beta_i^{(1,1)} = u_{i1}^2 = R_i(\infty)$ (notice that $u_{i1} \in \mathbb{R}$ for all $i$ since $\lambda_1 \in \mathbb{R}_{>0}$ according to the Perron-Frobenius Theorem [21, Fact 4.11.4]). Choose $r(1) \in \arg\max_i R_i(1)$ and $r(\infty) \in \arg\max_i R_i(\infty)$. The network belongs to class $\mathcal{V}$ if for some $k > 1$,

$$R_{r(\infty)}(k) > R_{r(1)}(k) \Leftrightarrow R_{r(\infty)}(\infty)\lambda_1^{2k} + \sum_{(j,m)\neq(1,1)} \beta_{r(\infty)}^{(j,m)}\lambda_j^k\overline{\lambda}_m^k > R_{r(1)}(\infty)\lambda_1^{2k} + \sum_{(j,m)\neq(1,1)} \beta_{r(1)}^{(j,m)}\lambda_j^k\overline{\lambda}_m^k$$

$$\Leftrightarrow \left[ R_{r(\infty)}(\infty) - R_{r(1)}(\infty) \right]\lambda_1^{2k} > \sum_{(j,m)\neq(1,1)} \left[ \beta_{r(1)}^{(j,m)} - \beta_{r(\infty)}^{(j,m)} \right]\lambda_j^k\overline{\lambda}_m^k$$

$$\overset{(a)}{\Leftarrow} \left[ R_{r(\infty)}(\infty) - R_{r(1)}(\infty) \right]\lambda_1^{2k} > \lambda_1^k|\lambda_2|^k \left| \sum_{(j,m)\neq(1,1)} \beta_{r(1)}^{(j,m)} - \beta_{r(\infty)}^{(j,m)} \right|$$

$$\Leftrightarrow \left[ R_{r(\infty)}(\infty) - R_{r(1)}(\infty) \right]\lambda_1^{2k} > \lambda_1^k|\lambda_2|^k \sum_{(j,m)\neq(1,1)} \left| \beta_{r(1)}^{(j,m)} \right| + \left| \beta_{r(\infty)}^{(j,m)} \right|$$

$$\Leftarrow \left[ R_{r(\infty)}(\infty) - R_{r(1)}(\infty) \right]\lambda_1^k > |\lambda_2|^k \cdot 2 \max_{i\in\{1,\dots,n\}} \sum_{j,m=1}^n \left| \beta_i^{(j,m)} \right|,$$

where in (a) we have used the fact that $|\lambda_j \bar{\lambda}_m| \leq \lambda_1 |\lambda_2|$ for any $(j, m) \neq (1, 1)$. Now, using the definition of $\beta_i^{(j,m)}$,

$$
\begin{aligned}
\sum_{j,m=1}^{n} \left| \beta_i^{(j,m)} \right| &\leq \sum_{j,m=1}^{n} \sum_{\ell=1}^{n} |v_{\ell j}| |v_{\ell m}| |u_{ij}| |u_{im}| \\
&= \sum_{j,m=1}^{n} |u_{ij}| |u_{im}| \left( \sum_{\ell=1}^{n} |v_{\ell j}| |v_{\ell m}| \right) \\
&\stackrel{(b)}{\leq} \sum_{j,m=1}^{n} |u_{ij}| |u_{im}| \underbrace{\|V_{:,j}\|_2}_{1} \underbrace{\|V_{:,m}\|_2}_{1} \\
&= \|U_{i,:}\|_1^2 \\
&\leq \|U\|_\infty^2,
\end{aligned}
$$

where (b) follows from the Cauchy-Schwarz inequality. Thus,

$$
\begin{aligned}
R_{r(\infty)}(k) > R_{r(1)}(k) &\Leftarrow \left[ R_{r(\infty)}(\infty) - R_{r(1)}(\infty) \right] \lambda_1^k > |\lambda_2|^k \cdot 2\|U\|_\infty^2 \\
&\Leftrightarrow k > \frac{\log 2\|U\|_\infty^2 - \log \left[ R_{r(\infty)}(\infty) - R_{r(1)}(\infty) \right]}{\log \lambda_1 - \log |\lambda_2|}.
\end{aligned}
$$

Therefore, the result follows by choosing $K > \bar{K}$, where $\bar{K} = \lceil \frac{\log 2\|U\|_\infty^2 - \log \left[ R_{r(\infty)}(\infty) - R_{r(1)}(\infty) \right]}{\log \lambda_1 - \log |\lambda_2|} \rceil$.          $\square$

The following lemma will be useful in the proof of Theorem 4.4.

LEMMA G.1  Let $W \in \mathbb{R}^{n \times n}$ be a doubly-stochastic matrix and $\gamma \in \mathbb{R}_{\geq 0}^n$ be such that $\gamma_1 \geq \cdots \geq \gamma_n$. If $\frac{1-w_{11}}{w_{11}} \leq \frac{\gamma_1 - \gamma_2}{\gamma_1 - \gamma_n}$, then $1 \in \arg\max_{1 \leq i \leq n} (W\gamma)_i$.

*Proof.*  Note that we have

$$
\begin{aligned}
\frac{1 - w_{11}}{w_{11}} \leq \frac{\gamma_1 - \gamma_2}{\gamma_1 - \gamma_n} &\Leftrightarrow (\gamma_1 - \gamma_2) w_{11} \geq (\gamma_1 - \gamma_n)(1 - w_{11}) \\
&\Leftrightarrow \gamma_n + w_{11}(\gamma_1 - \gamma_n) \geq \gamma_2 + (1 - w_{11})(\gamma_1 - \gamma_2) \\
&\Rightarrow \forall i \geq 2 \quad \gamma_n + w_{11}(\gamma_1 - \gamma_n) \geq \gamma_2 + w_{i1}(\gamma_1 - \gamma_2),
\end{aligned}
$$

where the last implication is because $w_{i1} \leq 1 - w_{11}$ for all $i \in \{1, \ldots, n\}$. The last inequality can be equivalently expressed, for any $i \in \{2, \ldots, n\}$, as

$$
w_{11} \gamma_1 + (1 - w_{11}) \gamma_n \geq w_{i1} \gamma_1 + (1 - w_{i1}) \gamma_2,
$$

which, given that $\gamma_n \leq \gamma_j \leq \gamma_2$ for all $j \in \{2, \ldots, n\}$, implies

$$
w_{11} \gamma_1 + \sum_{j=2}^{n} w_{1j} \gamma_j \geq w_{i1} \gamma_1 + \sum_{j=2}^{n} w_{ij} \gamma_j,
$$

for any $i \in \{2, \ldots, n\}$. This can be equivalently written as

$$\sum_{j=1}^{n} w_{1j}\gamma_j \geq \sum_{j=1}^{n} w_{ij}\gamma_j \Leftrightarrow (W\gamma)_1 \geq (W\gamma)_i,$$

completing the proof. □

***Proof of Theorem 4.4.*** For convenience, let $\lambda = [\lambda_1 \cdots \lambda_n]^T$. After some algebraic manipulations, one can show that

$$R_i(k) = (A^{2k})_{ii} = \sum_{j=1}^{n} v_{ij}^2 \lambda_j^{2k} = (W\lambda^{2k})_i. \tag{G.1}$$

The assumption that node 1 has the largest eigenvector centrality is equivalent to the largest element of the first column of $W$ being $w_{11}$, i.e.,

$$w_{11} = \max_{1 \leq i \leq n} w_{i1}, \tag{G.2}$$

or, also equivalently, $r(\infty) = 1$. This can always be realized by a permutation of the rows of $W$ achieved by relabelling the node with the largest centrality as node 1 (note that relabelling the nodes only permutes the rows of $W$ and not its columns. The order of its columns is arbitrary and corresponds to the order of the diagonal elements of $\Lambda$).

The claim of the theorem is trivial in all cases for $k = 0$. Under condition *(i)*, for $k = 1$, we have

$$\frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 - \lambda_n^2} = \frac{|\lambda_1| - |\lambda_2|}{|\lambda_1| - |\lambda_n|} \frac{|\lambda_1| + |\lambda_2|}{|\lambda_1| + |\lambda_n|} \geq \frac{|\lambda_1| - |\lambda_2|}{|\lambda_1| - |\lambda_n|} \geq \frac{1 - w_{11}}{w_{11}}.$$

For $k \geq 2$, using the above inequality, we have

$$\frac{\lambda_1^{2k} - \lambda_2^{2k}}{\lambda_1^{2k} - \lambda_n^{2k}} = \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 - \lambda_n^2} \frac{\lambda_1^{2k-2} + \cdots + \lambda_2^{2k-2}}{\lambda_1^{2k-2} + \cdots + \lambda_n^{2k-2}} \geq \frac{1 - w_{11}}{w_{11}}.$$

Thus, the result follows from Lemma G.1.

Under condition *(ii)*, for any $k \geq 1$,

$$1 \in \arg\max_{1 \leq i \leq n} R_i(k) \Leftrightarrow \sum_{j=1}^{n} w_{1j}\lambda_j^{2k} \geq \sum_{j=1}^{n} w_{ij}\lambda_j^{2k}$$

$$\Leftrightarrow w_{11}\lambda_1^{2k} + (1 - w_{11})\lambda_2^{2k} \geq \sum_{j=1}^{n} w_{ij}\lambda_j^{2k}$$

$$\Leftarrow w_{11}\lambda_1^{2k} + (1 - w_{11})\lambda_2^{2k} \geq w_{i1}\lambda_1^{2k} + (1 - w_{i1})\lambda_2^{2k}$$

$$\Leftrightarrow (w_{11} - w_{i1})(\lambda_1^{2k} - \lambda_2^{2k}) \geq 0,$$

where the last inequality is always true (cf. equation (G.2)).

Finally, under condition *(iii)*, first consider the case when $|\lambda_1| > |\lambda_2|$. By contradiction, assume $R_i(k) > R_1(k)$ for some $i \in \{2, \ldots, n\}$ and $k \geq 2$. Since $|\lambda_1| > |\lambda_i|$ for all $i \in \{2, \ldots, n\}$, there exists a sufficiently large $\bar{k}$ where $R_1(\bar{k}) > R_i(\bar{k})$ (recall our node labelling convention in (G.2)). Note that it is not necessary for $\bar{k}$ to be less than $K$. Thus, $R_1$ and $R_i$ swap orders at least 2 times. However, since $A$ has (at most) three distinct nonzero eigenvalues, [82, Theorem 1] implies that $R_1$ and $R_i$ can swap orders at most once, which is a contradiction. On the other hand, if $|\lambda_1| = |\lambda_2|$, then each $R_i$ is essentially the sum of at most two distinct exponential functions and thus, using [82, Theorem 1] again, the order of all $R_i$'s remains unchanged for all $k$, yielding the result. $\qquad\square$

***Proof of Theorem 4.5.*** We first prove the first part of the theorem for general (not necessarily symmetric) $A_0$ and $E$. Recall that for $k \in \{0, \ldots, K-1\}$

$$r(k) = \arg\max_{i \in \mathcal{N}} R_i(k) = \arg\max_{i \in \mathcal{N}} \left( ((A + \alpha E)^k)^T (A + \alpha E)^k \right)_{ii}$$

$$\overset{(a)}{=} \arg\max_{i \in \mathcal{N}} \left( ((\alpha^{-1} A + E)^k)^T (\alpha^{-1} A + E)^k \right)_{ii},$$

where (a) holds because the maximizer of a set is invariant to the scaling of all the elements of the set by a constant. Using $\lim_{\alpha \to \infty} \alpha^{-1} A + E = E$ and the continuity of polynomials, we get

$$\lim_{\alpha \to \infty} R_i(k) = \tilde{R}_i(k),$$

where $\tilde{R}_i$ denotes the $2k$-communicabilities of a node $i$ in the additive network $E$. Since $E$ is not acyclic, powers of $E$ never vanish, and thus

$$\forall k \in \{0, \ldots, K-1\} \; \exists i \in \{1, \ldots, n_1\} \qquad \tilde{R}_i(k) > 0,$$

while $\tilde{R}_i(k) = 0$ for $i \in \{n_1 + 1, \ldots, n\}$ and all $k$. Therefore, for any $k \in \{0, \ldots, K-1\}$, there exists $\bar{\alpha}_k > 0$ such that

$$r(k) \in \{1, \ldots, n_1\},$$

for $\alpha > \bar{\alpha}_k$. The claim follows by taking $\bar{\alpha} = \max_{k \in \{0, \ldots, K-1\}} \bar{\alpha}_k$.

Now, assume $A_0$ and $E$ are symmetric. As before, let $\lambda = [\lambda_1 \; \cdots \; \lambda_n]^T \in \mathbb{R}^n$ and $V \in \mathbb{R}^{n \times n}$ be the vector of eigenvalues (with $|\lambda_1| \geq \cdots \geq |\lambda_n|$) and the matrix of eigenvectors of $A$, respectively, and $W$ be the element-wise square of $V$. Recall that this gives

$$R_i(k) = (A^{2k})_{ii} = \sum_{j=1}^{n} v_{ij}^2 \lambda_j^{2k} = (W\lambda^{2k})_i.$$

Let $i^* \in \{1, \ldots, n_1\}$ be the node with the greatest eigenvector centrality in $E$ and $\gamma \in \mathbb{R}^n$ be any vector such that $\gamma_1 \geq \cdots \geq \gamma_n \geq 0$. Fix $i \in \{n_1 + 1, \ldots, n\}$ arbitrarily and let $r \leq n_1$ be the rank of $E$. Using the

inequalities

$$\sum_{j=1}^{n} w_{i^*j}\gamma_j \geq w_{i^*1}\gamma_1,$$

$$\sum_{j=1}^{r} w_{ij}\gamma_j \leq \gamma_1 \sum_{j=1}^{r} w_{ij},$$

$$\sum_{j=r+1}^{n} w_{ij}\gamma_j \leq \gamma_{r+1},$$

it follows that $(W\gamma)_{i^*} > (W\gamma)_i$ if

$$w_{i^*1}\gamma_1 > \gamma_1 \sum_{j=1}^{r} w_{ij} + \gamma_{r+1}. \tag{G.3}$$

Note that if (G.3) holds for $\gamma = |\lambda|$, then it holds for $\gamma = \lambda^{2k}$ for any $k \geq 1$. This is because

$$w_{i^*1}\lambda_1^{2k} = |\lambda_1|^{2k-1} \cdot w_{i^*1}|\lambda_1| > |\lambda_1|^{2k-1}\left(|\lambda_1|\sum_{j=1}^{r} w_{ij} + |\lambda|_{r+1}\right) > \lambda_1^{2k}\sum_{j=1}^{r} w_{ij} + \lambda_{r+1}^{2k}.$$

Therefore, our proof strategy is to find $\overline{\alpha}$ such that (G.3) holds for $\gamma = |\lambda|$ if $\alpha > \overline{\alpha}$. To this end, let $\tilde{\lambda} = [\tilde{\lambda}_1 \cdots \tilde{\lambda}_n]^T \in \mathbb{R}^n$ and $\tilde{V} \in \mathbb{R}^{n \times n}$ be the vector of eigenvalues (with $|\tilde{\lambda}_1| \geq \cdots \geq |\tilde{\lambda}_n|$) and the matrix of eigenvectors of $E$, respectively, and $\tilde{W}$ be the element-wise square of $\tilde{V}$. Note that $\tilde{W}$ has the structure

$$\tilde{W} = \begin{bmatrix} \overbrace{\star}^{n_1} & \overbrace{0}^{n-n_1} \\ 0 & \star \end{bmatrix} \begin{matrix} \}n_1 \\ \}n-n_1 \end{matrix}.$$

$$\tag{G.4}$$

In the following, we bound $\lambda$ and $V$ using perturbation theory of eigenvalues and eigenvectors. For simplicity of exposition, we only deal with the case where the $r$ nonzero eigenvalues of $E$ are all distinct (the proof for the general case proceeds along the same lines but is more involved).

To bound the eigenvalues in $\lambda$, let $\pi_A : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ be a permutation that re-orders the eigenvalues of $A$ based on their *signed* value, i.e., $\lambda_{\pi_A(1)} \geq \lambda_{\pi_A(2)} \geq \cdots \geq \lambda_{\pi_A(n)}$. Define $\pi_E$ similarly for $E$ (i.e., such that $\tilde{\lambda}_{\pi_E(1)} \geq \tilde{\lambda}_{\pi_E(2)} \geq \cdots \geq \tilde{\lambda}_{\pi_E(n)}$). By Weyl's Theorem [83, Thm 4.3.1],

$$|\lambda_{\pi_A(j)} - \alpha\tilde{\lambda}_{\pi_E(j)}| \leq \rho(A_0), \tag{G.5}$$

for all $j \in \{1, \ldots, n\}$. We know from the Perron-Frobenius theorem [21, Fact 4.11.4] for non-negative matrices that $\pi_A(1) = \pi_E(1) = 1$. Therefore, (G.5) implies that

$$\alpha\rho(E) - \rho(A_0) \leq \lambda_1 \leq \alpha\rho(E) + \rho(A_0). \tag{G.6a}$$

Moreover, since $E$ has $n - r$ zero eigenvalues, (G.5) implies that $A$ has *at least* $n - r$ eigenvalues with absolute value less than or equal to $\rho(A_0)$, i.e.,

$$|\lambda_{r+1}| \leq \rho(A_0). \tag{G.6b}$$

Next, we bound the eigenvectors in $V$. Define

$$\delta_E = \min\{\tilde{\lambda}_{\pi_E(j)} - \tilde{\lambda}_{\pi_E(j+1)} \mid \tilde{\lambda}_{\pi_E(j)} - \tilde{\lambda}_{\pi_E(j+1)} > 0, j \in \{1, \ldots, n-1\}\}.$$

Using [84, Cor. 1], we have

$$\|v_{\pi_A(j)} - \tilde{v}_{\pi_E(j)}\| \leq \frac{2^{3/2}\|A_0\|}{\alpha\delta_E}, \tag{G.7}$$

for $j \in \pi_E^{-1}(\{1, \ldots, r\})$. To see this, set $\Sigma = \alpha E$ and $\hat{\Sigma} = A_0$ in [84, Cor. 1]. This is the only place where we need the nonzero eigenvalues of $E$ to be distinct. If $E$ has a repeated nonzero eigenvalue, then the corresponding eigenvectors are no longer unique, i.e., one has to study the perturbation of eigenspaces rather than eigenvectors. Therefore, one can no longer use the simplified variant [84, Cor. 1] of the Davis-Kahan Theorem but the original result itself, which provides essentially the same result but is more technically involved.

Using $\pi_A(1) = \pi_E(1) = 1$ and (G.7), we get

$$|w_{i*1} - \tilde{w}_{i*1}| = |v_{i*1}^2 - \tilde{v}_{i*1}^2| \leq 2||v_{i*1}| - |\tilde{v}_{i*1}|| \tag{G.8}$$
$$\leq 2|v_{i*1} - \tilde{v}_{i*1}| \leq 2\|v_1 - \tilde{v}_1\| \leq \frac{2^{5/2}\|A_0\|}{\alpha\delta_E},$$

which together with $\tilde{w}_{i*1} \geq \frac{1}{n_1}$ gives

$$w_{i*1} \geq \frac{1}{n_1} - \frac{2^{5/2}\|A_0\|}{\alpha\delta_E}. \tag{G.9a}$$

To derive similar bounds on $w_{ij}, j \in \{1, \ldots, r\}$ (recall that we fixed $i \in \{n_1 + 1, \ldots, n\}$ arbitrarily at the beginning of the proof), we need to choose $\alpha > \frac{2\rho(A_0)}{|\tilde{\lambda}_r|}$. This choice of $\alpha$ guarantees that $\pi_A(j) \in \{1, \ldots, r\}$ for all $j \in \pi_E^{-1}(\{1, \ldots, r\})$). Therefore, using (G.7) and (G.4) and following the same steps as in (G.8), we get

$$w_{ij} \leq \frac{2^{5/2}\|A_0\|}{\alpha\delta_E}, \quad j \in \{1, \ldots, r\}. \tag{G.9b}$$

Now, using (G.6) and (G.9), (G.3) holds with $\gamma = |\lambda|$ if

$$\left(\frac{1}{n_1} - \frac{2^{5/2}\|A_0\|}{\alpha\delta_E}\right)\left(\alpha\tilde{\lambda}_1 - \rho(A_0)\right) > \left(\alpha\tilde{\lambda}_1 + \rho(A_0)\right)\frac{r2^{5/2}\|A_0\|}{\alpha\delta_E} + \rho(A_0),$$

which itself holds if $\alpha > \overline{\alpha}$, where

$$\overline{\alpha} \triangleq \max\left\{1, \frac{2\rho(A_0)}{\tilde{\lambda}_r}, \frac{8\|A_0\|}{\delta_E}\left(1 + \frac{\rho(A_0)}{\rho(E)}\right)n_1^2 + 2\frac{\rho(A_0)}{\rho(E)}n_1\right\},$$

completing the proof. □

## H. Description of the analysed real networks

The real networks studied in this work have been acquired from a multitude of sources, which we list here for easier reproduction of our results. All the databases are freely and publicly available.

- **BCTNet fMRI [47]:** This is a human whole-brain functional network. Nodes represent brain areas and edges represent fMRI co-activations. The dataset is available online at https://sites.google.com/site/bctnet/datasets.

- **Cocomac [48]:** This is a macaque whole-brain structural network based on the Felleman and Van Essen atlas. Nodes represent brain areas and edges represent axonal projections (nerve tracts) between them. The dataset is retrieved from http://cocomac.g-node.org/services/axonal_projections.php by entering the specifications detailed in http://cocomac.g-node.org/main/faq.php#connectivity matrix.

- **BCTNet Cat [47]:** This represents the cat structural thalamocortical network. Nodes represent thalamocortical areas and edges represent nerve tracts between them. The dataset is available online at https://sites.google.com/site/bctnet/datasets.

- **C. elegans [49]:** This dataset contains the neural network of *Caenorhabditis elegans* worm (*C. elegans*). Nodes represent individual neurons and edges represent the total number of synapses and gap junctions between any pair of neurons. The dataset of available online at https://toreopsahl.com/datasets/#celegans.

- **air500 [50]:** This is the network of the 500 busiest commercial airports in the United States in 2002. Nodes represent airports and edges represent flights between them. The dataset is available online at https://toreopsahl.com/datasets/#usairports.

- **airUS [51]:** This is the complete US airport network in 2010. Nodes and edges represent airports and flights between them, respectively. The dataset is available online at https://toreopsahl.com/datasets/#usairports.

- **airGlobal [51]:** This dataset contains the global airport network according to OpenFlights.org. Nodes and edges represent airports and flights between them, respectively. The dataset is available online at https://toreopsahl.com/datasets/#usairports.

- **Chicago [52, 53]:** This dataset represents the road transportation network of the Chicago region, USA. Nodes are transport nodes while edges represent connections between them. The dataset is available online as [85].

- ***Escherichia coli* [54]:** This is the probabilistic functional gene network of *E. coli*. Nodes represent genes and edges represent interactions between them. The dataset is available online at http://www.inetbio.org/ecolinet/downloadnetwork.php (the integrated network).

- **Yeast [55]:** This network represents the protein–protein interaction network in the budding yeast. Nodes and edges represent proteins and the interactions among them, respectively. The dataset is available online at http://vlado.fmf.uni-lj.si/pub/networks/data/bio/Yeast/Yeast.htm.

- **Stelzl [56]:** This is a protein–protein interaction network in humans. Nodes and edges represent proteins and the interactions among them, respectively. The dataset is available online as [86].

- **Figeys [57]:** Similar to above, this is a protein–protein interaction network in humans where nodes and edges represent proteins and the interactions among them, respectively. The dataset is available online as [87].

- **Vidal [58]:** Similar to above, this is a protein–protein interaction network in humans where nodes and edges represent proteins and the interactions among them, respectively. The dataset is available online as [88].

- **westernUS [49]:** This dataset describes the high voltage power grid in the Western States of the US. Nodes represent transformers, substations and generators, and the edges represent high-voltage transmission lines. The dataset is available online at https://toreopsahl.com/datasets/#uspowergrid.

- **Florida [59]:** This network describes the food web in the cypress wetlands of South Florida during the wet season. Nodes represent taxa and an edge denotes that a taxon uses another taxon as food. The dataset is available online as [89].

- **LRL [60]:** The networks describes the food web of Little Rock Lake, Wisconsin, USA. Nodes represent autotrophs, herbivores, carnivores and decomposers while links represent food sources. The dataset is available online as [90].

- **Facebook group [61]:** This dataset describes the social interactions among a group of Facebook users. Nodes and edges represent profiles and the connections between them, respectively. The dataset is available online at http://snap.stanford.edu/data/egonets-Facebook.html.

- **E-main [62, 63]:** This datasets contains E-main communications in a research institution. Nodes represent institution members and edges exist between any ordered pair of members if one has sent at least one E-main to the other. The dataset is available online at http://snap.stanford.edu/data/email-Eu-core.html.

- **Southern Women [64]:** This is a social network of 18 Southern women. Nodes are individuals and edges represent mutual attendance at one of the 14 events recorded. The dataset is available online at https://toreopsahl.com/datasets/#southernwomen.

- **UCI P2P [65]:** This dataset describes an online community among the students of the University of California, Irvine. Nodes represent individuals and edges represent at least one message sent between any pair of them. The dataset is available online at https://toreopsahl.com/datasets/#online_social_network.

- **UCI Forum [66]:** This network is based on the same online community as in UCR P2P, but an edge exists between two individuals if they posted on the same topic in a forum. This dataset is also available online at https://toreopsahl.com/datasets/#online_social_network.

- **Freeman's EIES [67]:** This is a network of researchers working on social network analysis. Nodes represent researchers and edges represent personal relationships between them. The dataset is available online at https://toreopsahl.com/datasets/#FreemansEIES (the second dataset in the list).

- **Dolphins [68]:** This is a social network of bottlenose dolphins observed between 1994 and 2001. The nodes are the bottlenose dolphins and edges indicate a frequent association between them. The dataset is available online as [91].

- **Physicians [69]:** This network captures innovation spread among 246 physicians in four towns in Illinois, USA. A node represents a physician and an edge represents that one physician recognizes the other as theor friend or that they turn to them if they need advice or are interested in a discussion. The dataset is available online as [92].

- **Org. Consult Advice & Value [70]:** These are intra-organizational networks between employees of a consulting company. The nodes are individuals, and the edges represent frequency of information or advice requests (Org. Consult Advice) and the value placed on the information or advice received (Org. Consult Value). The datasets are available online at https://toreopsahl.com/datasets/#Cross_Parker.

- **Org. R&D Advice & Aware [70]:** Similar to the networks above, these describe intra-organizational interactions among the members of a research team in a manufacturing company. Nodes represent individuals, and edges represent the extent to which individuals received advice from their peers to accomplish their work (Org. R&D Advice) and employees' awareness of each others' knowledge and skills (Org. R&D Aware). The datasets are available online at https://toreopsahl.com/datasets/#Cross_Parker.

<div align="center">REFERENCES</div>

1. KALMAN, R.E. (1963) Mathematical description of linear dynamical systems. *J. Soc. Ind. Appl. Math. Ser. A Contr.*, **1**, 152–192.
2. LIU, Y. Y., SLOTINE, J. J. & BARABÁSI, A. L. (2011) Controllability of complex networks. *Nature*, **473**, 167–173.
3. COWAN, N. J., CHASTAIN, E. J., VILHENA, D. A., FREUDENBERG, J. S. & BERGSTROM, C. T. (2012) Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PLoS One*, **7**, 1–5.
4. OLSHEVSKY, A. (2014) Minimal controllability problems. *IEEE Trans. Control Netw. Syst.*, **1**, 249–258.
5. YAN, G., REN, J., LAI, Y., LAI, C. & LI, B. (2012) Controlling complex networks: how much energy is needed? *Phys. Rev. Lett.*, **108**, 218703.
6. PASQUALETTI, F., ZAMPIERI, S., & BULLO, F. (2014) Controllability metrics, limitations and algorithms for complex networks. *IEEE Trans. Control Netw. Syst.*, **1**, 40–52.
7. SUMMERS, T. H. & LYGEROS, J. (2014) Optimal sensor and actuator placement in complex dynamical networks. *IFAC World Congress*. Cape Town, South Africa: Elsevier, pp. 3784–3789.
8. SUMMERS, T. H., CORTESI, F. L. & LYGEROS, J. (2016) On submodularity and controllability in complex dynamical networks. *IEEE Trans. Control Netw. Syst.*, **3**, 91–101.
9. TZOUMAS, V., RAHIMIAN, M. A., PAPPAS, G. J., & JADBABAIE, A. (2016) Minimal actuator placement with bounds on control effort. *IEEE Trans. Control Netw. Syst.*, **3**, 67–78.
10. PEQUITO, S., BOGDAN, P. & PAPPAS, G. J. (2015) Minimum number of probes for brain dynamics observability. *IEEE Conference on Decision and Control.*, IEEE, pp. 306–311.
11. BELABBAS, M. A. (2016) Geometric methods for optimal sensor design. *Proc. R. Soc. Lond. Ser. A*, **472**, 20150312.
12. ZHANG, H., AYOUB, R. & SUNDARAM, S. (2017) Sensor selection for Kalman filtering of linear dynamical systems: complexity, limitations and greedy algorithms. *Automatica*, **78**, 202–210.

13. TZOUMAS, V., XUE, Y., PEQUITO, S., BOGDAN, P. & PAPPAS, G. J. (2018) Selecting sensors in biological fractional-order systems. *IEEE Trans. Control Netw. Syst.*, **5**, 709–721.

14. ZHAO, L., ZHANG, W., HU, J., ABATE, A. & TOMLIN, C. J. (2014) On the optimal solutions of the infinite-horizon linear sensor scheduling problem. *IEEE Trans. Autom. Control*, **59**, 2825–2830.

15. JAWAID, S. T. & SMITH, S. L. (2015) Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems. *Automatica*, **61**, 282–288.

16. ZHAO, Y., PASQUALETTI, F. & CORTÉS, J. (2016) Scheduling of control nodes for improved network controllability. *IEEE Conference on Decision and Control*. Las Vegas, NV: IEEE, pp. 1859–1864.

17. HAN, D., WU, J., ZHANG, H. & SHI, L. (2017) Optimal sensor scheduling for multiple linear dynamical systems. *Automatica*, **75**, 260–270.

18. NOZARI, E., PASQUALETTI, F. & CORTÉS, J. (2017) Time-invariant versus time-varying actuator scheduling in complex networks. *American Control Conference*. Seattle, WA: IEEE, pp. 4995–5000.

19. BONACICH, P. (2007) Some unique properties of eigenvector centrality. *Soc. Netw.*, **29**, 555–564.

20. BONACICH, P. (1972) Factoring and weighting approaches to status scores and clique identification. *J. Math. Sociol.*, **2**, 113–120.

21. BERNSTEIN, D. S. (2009) *Matrix Mathematics*, 2nd edn. Princeton, New Jersey, USA: Princeton University Press.

22. ESTRADA, E. & HATANO, N. (2008) Communicability in complex networks. *Phys. Rev. E*, **77**, 036111.

23. KLYMKO, C. (2013) Centrality and communicability measures in complex networks: analysis and algorithms. *Ph.D. Dissertation*, Emory University.

24. CHEN, C.T. (1998) *Linear System Theory and Design*, 3rd edn. New York, NY, USA: Oxford University Press, Inc.

25. BALAKRISHNAN, H. (2016) Control and optimization algorithms for air transportation systems. *Ann. Rev. Control*, **41**, 39–46.

26. CHEN, B. & CHENG, H. H. (2010) A review of the applications of agent technology in traffic and transportation systems. *IEEE Trans. Knowl. Data Eng.*, **11**, 485–497.

27. ZHANG, W., KAMGARPOUR, M., SUN, D. & TOMLIN, C. J. (2012) A hierarchical flight planning framework for air traffic management. *Proceedings of the IEEE*, **100**, 179–194.

28. TEODOROVIC, D. & VUKADINOVIC, K. (2012) Traffic control and transport planning: a fuzzy sets and neural networks approach. *International Series in Intelligent Technologies*. Netherlands: Springer.

29. ALBI, G., PARESCHI, L. & ZANELLA, M. (2014) Boltzmann-type control of opinion consensus through leaders. *Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci.*, **372**, 20140138.

30. QIAN, C., CAO, J., LU, J. & KURTHS, J. (2011) Adaptive bridge control strategy for opinion evolution on social networks. *Chaos*, **21**, 025116.

31. BURSIK, R. J. (1999) The informal control of crime through neighborhood networks. *Sociol. Focus*, **32**, 85–97.

32. PROSKURNIKOV, A. V., MATVEEV, A. S. & CAO, M. (2016) Opinion dynamics in social networks with hostile camps: Consensus vs. polarization. *IEEE Trans. Autom. Control*, **61**, 1524–1536.

33. KITSAK, M., GALLOS, L. K., HAVLIN, S., LILJEROS, F., MUCHNIK, L., STANLEY, H. E. & MAKSE, H. A. (2010) Identification of influential spreaders in complex networks. *Nat. Phys.*, **6**, 888.

34. NOWZARI, C., PRECIADO, V. M. & PAPPAS, G. J. (2016) Analysis and control of epidemics: a survey of spreading processes on complex networks. *IEEE Control Syst.*, **36**, 26–46.

35. SALATHE, M. & JONES, J. H. (2010) Dynamics and control of diseases in networks with community structure. *PLoS Comput. Biol.*, **6**, e1000736.

36. SHAW, L. B. & SCHWARTZ, I. B. (2010) Enhanced vaccine control of epidemics in adaptive networks. *Phys. Rev. E*, **81**, 046120.

37. HUFNAGEL, L., BROCKMANN, D. & GEISEL, T. (2004) Forecast and control of epidemics in a globalized world. *Proc. Natl. Acad. Sci. USA*, **101**, 15124–15129.

38. HUERTA, R. & TSIMRING, L. S. (2002) Contact tracing and epidemics control in social networks. *Phys. Rev. E*, **66**, 056115.

39. CHEN, Y., PAUL, G., HAVLIN, S., LILJEROS, F. & STANLEY, H. E. (2008) Finding a better immunization strategy. *Phys. Rev. Lett.*, **101**, 058701.

40. GODSIL, C. D. & ROYLE, G. F. (2001) *Algebraic Graph Theory*. Graduate Texts in Mathematics, **207**. New York, USA: Springer, 2001.

41. VAN DAAM, E. R. (1996) Graphs with few eigenvalues: an interplay between combinatorics and algebra. *Ph.D. Dissertation*, Tilburg University.

42. KLEINBERG, J. M. (1999) Authoritative sources in a hyperlinked environment. *J. ACM*, **46**, 604–632.

43. REBESCO, J., STEVENSON, I., KOERDING, K., SOLLA, S. & MILLER, L. (2010) Rewiring neural interactions by micro-stimulation. *Front. Syst. Neurosci.*, **4**, 39.

44. GUGGENMOS, D. J., AZIN, M., BARBAY, S., MAHNKEN, J. D., DUNHAM, C., MOHSENI, P. & NUDO, R. J. (2013) Restoration of function after brain damage using a neural prosthesis. *Proc. Natl. Acad. Sci.*, **110**, 21177–21182.

45. LU, T. K., KHALIL, A. S. & COLLINS, J. J. (2009) Next-generation synthetic gene networks. *Nat. Biotechnol.*, **27**, 1139.

46. VECCHIO, D. D. & MURRAY, R. M. (2015) *Biomolecular Feedback Systems*. Princeton, New Jersey, USA: Princeton University Press.

47. RUBINOV, M. & SPORNS, O. (2010) Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, **52**, 1059–1069.

48. BAKKER, R., WACHTLER, T. & DIESMANN, M. (2012) CoCoMac 2.0 and the future of tract-tracing databases. *Front. Neuroinform.*, **6**, 30.

49. WATTS, D. J. & STROGATZ, S. H. (1998) Collective dynamics of 'small-world' networks. *Nature*, **393**, 440–442.

50. COLIZZA, V., PASTOR-SATORRAS, R. & VESPIGNANI, A. (2007) Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nat. Phys.*, **3**, 276–282.

51. OPSAHL, T. (2011) Why anchorage is not (that) important: binary ties and sample selection. Available at: http://wp.me/poFcY-Vw (accessed February 3, 2019).

52. EASH, R. W., CHON, K. S., LEE, Y. J. & BOYCE, D. E. (1983) Equilibrium traffic assignment on an aggregated highway network for sketch planning. *Transp. Res. Rec.*, **994**, 30–37.

53. BOYCE, D. E., CHON, K. S., FERRIS, M. E., LEE, Y. J., LIN, K.-T. & EASH, R. W. (1985) Implementation and evaluation of combined models of urban travel and location on a sketch planning network. *Chicago Area Transportation Study*. Transportation Research Board, pp. xii+169.

54. KIM, H., SHIM, J. E., SHIN, J. & LEE, I. (2015) Ecolinet: a database of cofunctional gene network for *Escherichia coli*. *Database*, **2015**, 1–8.

55. BU, D., ZHAO, Y., CAI, L., XUE, H., ZHU, X., LU, H., ZHANG, J., SUN, S., LING, L., ZHANG, N., LI, G. & CHEN, R. (2003) Topological structure analysis of the protein–protein interaction network in budding yeast. *Nucleic Acids Res.*, **31**, 2443–2450.

56. STELZL, U., WORM, U., LALOWSKI, M., HAENIG, C., BREMBECK, F. H., GOEHLER, H., STROEDICKE, M., ZENKNER, M., SCHOENHERR, A., KOEPPEN, S., TIMM, J., MINTZLAFF, S., ABRAHAM, C., BOCK, N., KIETZMANN, S., GOEDDE, A., TOKSÖZ, E., DROEGE, A., KROBITSCH, S., KORN, B., BIRCHMEIER, W., LEHRACH, H. & WANKER, E. E. (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell*, **122**, 957–968.

57. EWING, R. M., CHU, P., ELISMA, F., LI, H., TAYLOR, P., CLIMIE, S., MCBROOM-CERAJEWSKI, L., ROBINSON, M. D., O'CONNOR, L., LI, M., TAYLOR, R., DHARSEE, M., HO, Y., HEILBUT, A., MOORE, L., ZHANG, S., ORNATSKY, O., BUKHMAN, Y. V., ETHIER, M., SHENG, Y., VASILESCU, J., ABU-FARHA, M., LAMBERT, J. P. P., DUEWEL, H. S., STEWART, I. I., KUEHL, B., HOGUE, K., COLWILL, K., GLADWISH, K., MUSKAT, B., KINACH, R., ADAMS, S. L. L., MORAN, M. F., MORIN, G. B., TOPALOGLOU, T. & FIGEYS, D. (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol. Syst. Biol.*, **3**, 89.

58. RUAL, J., VENKATESAN, K., HAO, T., HIROZANE-KISHIKAWA, T., DRICOT, A., LI, N., BERRIZ, G. F., GIBBONS, F. D., DREZE, M. & AYIVI-GUEDEHOUSSOU, N. (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, **437**, 1173–1178.

59. ULANOWICZ, R. E., HEYMANS, J. J. & EGNOTOVICH, M. S. (2000) Network analysis of trophic dynamics in South Florida ecosystems, FY 99: the graminoid ecosystem. *Annual Report to the United States Geological Service Biological Resources Division Ref. No.[UMCES] CBL 00-0176, Chesapeake Biological Laboratory, University of Maryland*.

**60.** MARTINEZ, N. D., MAGNUSON, J. J., KRATZ, T. & SIERSZEN, M. (1991) Artifacts or attributes? effects of resolution on the Little Rock Lake food web. *Ecol. Monogr.*, **61**, 367–392.

**61.** LESKOVEC, J. & MCAULEY, J. J. (2012) Learning to discover social circles in ego networks. *Advances in Neural Information Processing Systems* (F. Pereira, C. J. C. Burges, L. Bottou & K. Q. Weinberger eds). Curran Associates, Inc., **25**. pp. 539–547.

**62.** YIN, H., BENSON, A. R., LESKOVEC, J. & GLEICH, D. F. (2017) Local higher-order graph clustering. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, pp. 555–564.

**63.** LESKOVEC, J., KLEINBERG, J. & FALOUTSOS, C. (2007) Graph evolution: densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data*, **1**, article number 2.

**64.** DAVIS, A., GARDNER, B. B. & GARDNER, M. R. (1941) *Deep South*. Chicago, IL: University of Chicago Press.

**65.** OPSAHL, T. & PANZARASA, P. (2009) Clustering in weighted networks. *Soc. Networks*, **31**, 155–163.

**66.** OPSAHL, T. (2013) Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Soc. Networks*, **35**, 159–167.

**67.** FREEMAN, S. C. & FREEMAN, L. C. (1979) *The Networkers Network: A Study of the Impact of a New Communications Medium on Sociometric Structure*. Social Sciences Research Reports. California, USA: School of Social Sciences University of California.

**68.** LUSSEAU, D., SCHNEIDER, K., BOISSEAU, O. J., HAASE, P., SLOOTEN, E. & DAWSON, S. M. (2003) The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behav. Ecol. Sociobiol.*, **54**, 396–405.

**69.** COLEMAN, J., KATZ, E. & MENZEL, H. (1957) The diffusion of an innovation among physicians. *Sociometry*, **20**, pp. 253–270.

**70.** CROSS, R. L. & PARKER, A. (2004) *The Hidden Power of Social Networks: Understanding how Work Really Gets Done in Organizations*. Brighton, Massachusetts, USA: Harvard Business School Press.

**71.** MÜLLER, P. C. & WEBER, H. I. (1972) Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems. *Automatica*, **8**, 237–246.

**72.** GU, S., PASQUALETTI, F., CIESLAK, M., TELESFORD, Q. K., YU, A. B., KAHN, A. E., MEDAGLIA, J. D., VETTEL, J. M., MILLER, M. B., GRAFTON, S. T. & BASSETT, D. S. (2015) Controllability of structural brain networks. *Nat. Commun.*, **6**, 8414 EP.

**73.** BAI, Z. & GOLUB, G. H. (1996) Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices. *Ann. Numer. Math.*, **4**, 29–38.

**74.** BARRY, R. P. & PACE, R. K. (1999) Monte Carlo estimates of the log determinant of large sparse matrices. *Linear Algebra Appl.*, **289**, 41–54.

**75.** REUSKEN, A. (2002) Approximation of the determinant of large sparse symmetric positive definite matrices. *SIAM J. Matrix Anal. Appl.*, **23**, 799–818.

**76.** IPSEN, I. C. F. & LEE, D. J. (2011) Determinant approximations. arXiv preprint:1105.0437.

**77.** HAN, I., MALIOUTOV, D. & SHIN, J. (2015) Large-scale log-determinant computation through stochastic chebyshev expansions. *International Conference on Machine Learning*. pp. 908–917.

**78.** BOUTSIDIS, C., DRINEAS, P., KAMBADUR, P., KONTOPOULOU, E. & ZOUZIAS, A. (2017) A randomized algorithm for approximating the log determinant of a symmetric positive definite matrix. *Linear Algebra Appl.*, **533**, 95–117.

**79.** FITZSIMONS, J., CUTAJAR, K., OSBORNE, M., ROBERTS, S. & FILIPPONE, M. (2017) Bayesian inference of log determinants. Conference on Uncertainty in Artificial Intelligence, August 11–15, 2017, Sydney, Australia, arXiv preprint:1704.01445.

**80.** BULLO, F., CORTÉS, J. & MARTÍNEZ, S. (2009) *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton, New Jersey, USA: Princeton University Press, electronically available at http://coordinationbook.info (accessed February 3, 2019).

**81.** GOULD, H. W. (2010) Combinatorial identities: Table I: Intermediate techniques for summing finite series. From the seven unpublished manuscripts of H. W. Gould. Edited and compiled by J. Quaintance.

**82.** TOSSAVAINEN, T. (2006) On the zeros of finite sums of exponential functions. *Gazette Australian Math. Soc.*, **33**, 47–50.

**83.** HORN, R. A. & C. R. JOHNSON (1985) *Matrix Analysis*. Cambridge, UK: Cambridge University Press.

**84.** YU, Y., WANG, T. & SAMWORTH, R. J. (2015) A useful variant of the Davis-Kahan theorem for statisticians. *Biometrika*, **102**, 315–323.

**85.** CHICAGO NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz.de/networks/tntp -ChicagoRegional (accessed February 3, 2019).

**86.** HUMAN PROTEIN (STELZL) NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz. de/networks/maayan-Stelzl (accessed February 3, 2019).

**87.** HUMAN PROTEIN (FIGEYS) NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz. de/networks/maayan-figeys (accessed February 3, 2019).

**88.** HUMAN PROTEIN (VIDAL) NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz. de/networks/maayan-vidal (accessed February 3, 2019).

**89.** FLORIDA ECOSYSTEM WET NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz. de/networks/foodweb-baywet (accessed February 3, 2019).

**90.** LITTLE ROCK LAKE NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz.de/ networks/maayan-foodweb (accessed February 3, 2019).

**91.** DOLPHINS NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz.de/networks/ dolphins (accessed February 3, 2019).

**92.** PHYSICIANS NETWORK DATASET—KONECT. (2017) [Online]. Available: http://konect.uni-koblenz.de/networks/ moreno_innovation (accessed February 3, 2019).