

# Role of graph architecture in controlling dynamical networks with applications to neural systems

Jason Z. Kim<sup>1</sup>, Jonathan M. Soffer<sup>1</sup>, Ari E. Kahn<sup>2,3</sup>, Jean M. Vettel<sup>1,4,5</sup>, Fabio Pasqualetti<sup>6</sup>  
and Danielle S. Bassett<sup>1,7\*</sup>

**Networked systems display complex patterns of interactions between components. In physical networks, these interactions often occur along structural connections that link components in a hard-wired connection topology, supporting a variety of system-wide dynamical behaviours such as synchronization. Although descriptions of these behaviours are important, they are only a first step towards understanding and harnessing the relationship between network topology and system behaviour. Here, we use linear network control theory to derive accurate closed-form expressions that relate the connectivity of a subset of structural connections (those linking driver nodes to non-driver nodes) to the minimum energy required to control networked systems. To illustrate the utility of the mathematics, we apply this approach to high-resolution connectomes recently reconstructed from *Drosophila*, mouse, and human brains. We use these principles to suggest an advantage of the human brain in supporting diverse network dynamics with small energetic costs while remaining robust to perturbations, and to perform clinically accessible targeted manipulation of the brain's control performance by removing single edges in the network. Generally, our results ground the expectation of a control system's behaviour in its network architecture, and directly inspire new directions in network analysis and design via distributed control.**

Network systems are composed of interconnected units that interact with each other on diverse temporal and spatial scales<sup>1</sup>. The exact patterns of interconnections between these units can take on many different forms that dictate how the system functions<sup>2</sup>. Indeed, specific features of network topology—such as small worldness<sup>3</sup> and modularity<sup>4</sup>—can improve efficiency and robustness. Yet, exact mechanisms driving the relationship between structure and function remain elusive, hampering the analysis, modification, and control of interconnected complex systems. The relationship between interconnection architecture and dynamics is particularly important in biological systems such as the brain<sup>5</sup>, where it is thought to support optimal information processing at cellular<sup>6</sup> and regional<sup>7,8</sup> levels. Understanding structure–function relationships in this system could inform personalized therapeutics<sup>9</sup> including more targeted treatments for drug-resistant epilepsy to make the epileptic state energetically unfavourable to maintain<sup>10,11</sup>, especially due to the development of multi-site stimulation tools<sup>12,13</sup> that allow for exponentially increasing stimulation configurations.

Existing paradigms seeking to explain how a complex network topology drives observable dynamics have advantages and disadvantages. Efforts in nonlinear dynamics define basins of attraction and perturbations driving a system between basins<sup>14,15</sup>. Efforts in network science define graph metrics and report statistical correlations with observed functions such as attention<sup>16</sup> and learning<sup>17,18</sup>. Neither approach offers comprehensive analytical solutions explaining mechanisms of control. A promising paradigm that meets these challenges is linear network control theory<sup>19,20</sup>, which assumes that the state of a system at a given time is a function of the previous state, the structural network linking system units, and injected control energy. From this paradigm, one can identify

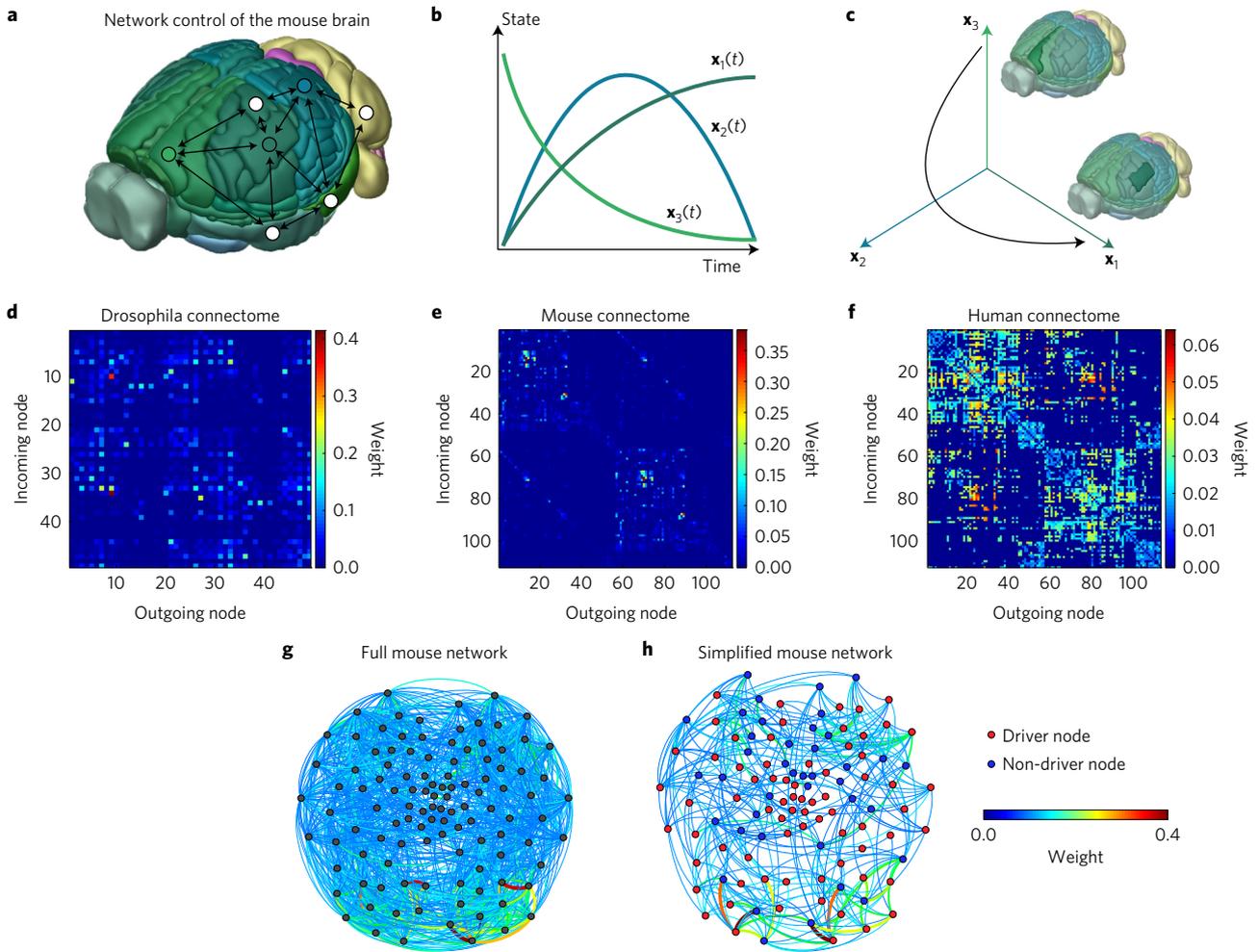
driver nodes<sup>21,22</sup> capable of influencing the system along diverse trajectories, and optimal inputs that move the system from one state to another with minimal cost. This latter formulation has proven useful in understanding the human brain, where control points enable diverse cognitive strategies<sup>23,24</sup>, facilitate efficient intrinsic activation<sup>25</sup>, and inform optimal targets for brain stimulation<sup>26</sup>.

Although practical tools exist, basic intuitions about the network properties that enhance control have remained elusive. Here, we address this challenge by formulating a linear control problem on the bipartite subgraph linking driver nodes to non-driver nodes, which provides excellent estimates of the control of the full network. Our results include analytical derivations of expressions relating a network's minimum control energy to its connectivity, an intuitive geometric representation to visualize this relationship, and rules for modifying edges to alter control energy in a predictable manner. Although our mathematical contributions are applicable to any complex network system whose dynamics can be approximated by a linear model, we illustrate their utility in the context of networks estimated from mouse<sup>27,28</sup>, *Drosophila*<sup>29</sup> and human brains (Fig. 1d–f). Our results offer fundamental insights into the patterns of connections between brain regions that directly impact their minimum control energy, providing a link between the structure and function of neural systems, and informing potential clinical interventions. An extension of this framework to non-bipartite graphs with corresponding results can be found in the Supplementary Methods and results.

## Network topology and controllability

We consider a network represented by the directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V} = \{1, \dots, n\}$  and  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  are the sets of network vertices

<sup>1</sup>Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. <sup>2</sup>Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. <sup>3</sup>US Army Research Laboratory, Aberdeen, Maryland 21001, USA. <sup>4</sup>Human Research & Engineering Directorate, US Army Research Laboratory, Aberdeen, Maryland 21001, USA. <sup>5</sup>Department of Psychological and Brain Sciences, University of California, Santa Barbara, California 93106, USA. <sup>6</sup>Department of Mechanical Engineering, University of California, Riverside, California 92521, USA. <sup>7</sup>Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. \*e-mail: dsb@seas.upenn.edu



**Figure 1 | Network control of the drosophila, mouse, and human connectomes.** **a**, A representation of the mouse brain via the Allen Mouse Brain Atlas, with a superimposed simplified network. Each brain region is represented as a vertex, and the connections between regions are represented as directed edges. **b**, Example trajectories of state over time for three brain regions, where the state represents the level of activity in each region. **c**, A state-space representation of activity on the mouse connectome over time, where each point on the black line represents the brain state at a point in time. **d–f**, Connectomes represented as  $n \times n$  adjacency matrices where each  $i, j$ th element of the adjacency matrix represents the strength of the connection from node  $j$  to node  $i$  for Drosophila (**d**), mouse (**e**), and human (**f**). **g**, The mouse connectome represented as a graph with vertices as brain regions, and edges coloured by their weight, or the magnitude of the relevant element of the adjacency matrix. **h**, Simplified graph representation: a bipartite subgraph containing edges linking driver vertices (red) to non-driver vertices (blue).

and edges, respectively. Let  $a_{ij} \in \mathbb{R}$  be the weight associated with the edge  $(i, j) \in \mathcal{E}$ , and let  $A = [a_{ij}]$  be the weighted adjacency matrix of  $\mathcal{G}$ . We associate a real value (state) with each node, collect the nodes' states into a vector (network state), and define the map  $\mathbf{x} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n$  to describe the evolution (dynamics) of the network state over time (Fig. 1a–c). We assume that a subset of  $N$  nodes, called drivers, is independently manipulated by external controls and, without loss of generality, we reorder the network nodes such that the  $N$  drivers come first. Thus, the network dynamics read as

$$\begin{bmatrix} \dot{\mathbf{x}}_d \\ \dot{\mathbf{x}}_{nd} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_d \\ \mathbf{x}_{nd} \end{bmatrix} + \begin{bmatrix} I_N \\ 0 \end{bmatrix} \mathbf{u} \quad (1)$$

where  $\mathbf{x}_d$  and  $\mathbf{x}_{nd}$  are the state vectors of the driver and non-driver nodes,  $A_{11} \in \mathbb{R}^{N \times N}$ ,  $M = n - N$ ,  $A_{12} \in \mathbb{R}^{N \times M}$ ,  $A_{21} \in \mathbb{R}^{M \times N}$ ,  $A_{22} \in \mathbb{R}^{M \times M}$ ,  $I_N$  is the  $N$ -dimensional identity matrix, and  $\mathbf{u} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^N$  is the control input.

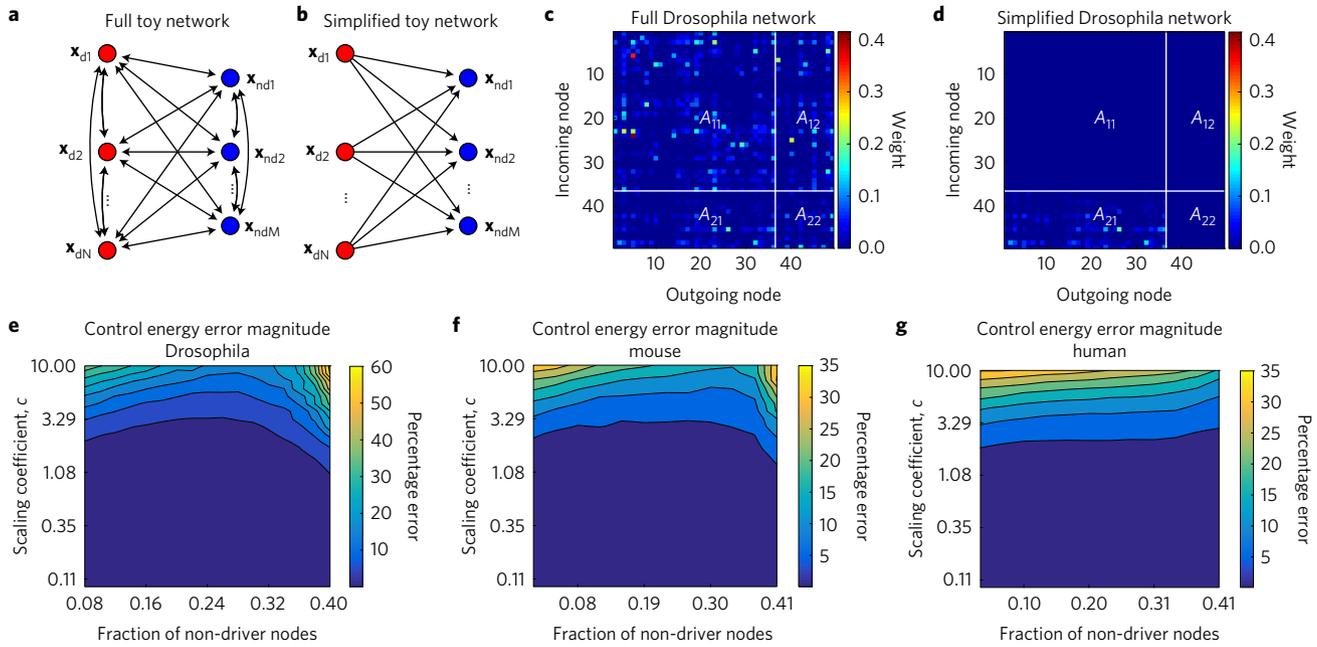
We will use the word controllable to refer to networks that are point-to-point controllable at time  $T \in \mathbb{R}_{\geq 0}$  if, for any pair of states  $\mathbf{x}_d^*$  and  $\mathbf{x}_{nd}^*$ , there exists a control input  $\mathbf{u}$  for the dynamics equation (1) such that  $\mathbf{x}_d(T) = \mathbf{x}_d^*$  and  $\mathbf{x}_{nd}(T) = \mathbf{x}_{nd}^*$ . For a detailed

discussion and rigorous conditions for the controllability of a system with linear dynamics, see ref. 30. We define the energy of  $\mathbf{u}$  as

$$E(\mathbf{u}) = \sum_{i=1}^N \underbrace{\int_0^T u_i(t)^2 dt}_{E_i}$$

where  $u_i$  is the  $i$ th component of  $\mathbf{u}$ . The energy of  $u_i$  can be thought of as a quadratic cost that penalizes large control inputs.

In the context of the brain, we approximate the interactions between brain regions as linear, time-invariant dynamics, where a stronger structural connection between two regions represents a stronger dynamic interaction (for empirical motivation, see refs 23,31,32). We specifically study the empirical inter-areal meso-scale connectomes of the mouse (112 brain regions, example schematic in Fig. 1g,h) from the Allen Brain Institute, the Drosophila (49 brain regions)<sup>29</sup>, and a set of human connectomes (116 brain regions) interconnected by white matter tracts (for empirical details regarding connectivity estimates, see Supplementary Results XA).



**Figure 2 | The simplified network representation offers a reasonable prediction for the full network’s control energy.** **a**, Graphical representation of a non-simplified network of  $N$  drivers (red) and  $M$  non-drivers (blue), with directed connections between all nodes present. **b**, Graphical representation of a simplified first-order network containing only first-order connections from drivers  $\rightarrow$  non-drivers. **c,d**, As an example, we show the adjacency matrix for the *Drosophila* connectome segmented into driver  $\rightarrow$  driver  $A_{11}$ , driver  $\rightarrow$  non-driver  $A_{21}$ , non-driver  $\rightarrow$  driver  $A_{12}$ , and non-driver  $\rightarrow$  non-driver  $A_{22}$  sections for a non-simplified network as per equation (1), with randomly designated driver and non-driver nodes (**c**), and the corresponding simplified network as per equation (2) (**d**). **e–g**, Percentage error contour plots of the total control energy for simplified versus non-simplified networks as a function of the fraction of non-driver nodes and matrix scale given by  $c = \|\lambda_{\max}\|$ . For each combination of parameters, the median error magnitude to drive the networks from initial states  $\mathbf{x}_d = 0, \mathbf{x}_{nd} = 0$  to 1,000 random final states  $\mathbf{x}_{nd}^* \in (-1, 1)^M, \mathbf{x}_d^* \in (-1, 1)^N$  along 1,000 corresponding random selections of non-drivers is shown. Each contour represents a 5% interval for the *Drosophila* (**e**), mouse (**f**), and human (**g**) connectome.

**Predicting control energy**

We seek an accurate, tractable relationship between the energy required to drive a network to a specific state and its connectivity. We begin with the original, non-simplified network (Fig. 2a) involving edges between all nodes, and consider dynamics along the simplified network (Fig. 2b) involving only edges from the driver to the non-driver nodes (for a conceptual schematic of the full and simplified *Drosophila* connectome, see Fig. 2c,d). We then derive an approximation of the minimum control energy (Supplementary Lemma X.2–X.4) by assuming that  $\mathbf{x}_d(0) = 0, \mathbf{x}_{nd}(0) = 0$  (Assumption 1), and  $A_{11} = 0, A_{12} = 0,$  and  $A_{22} = 0$  (Assumption 2) in equation (1), which reads as

$$E(\mathbf{u}) = 12 \left( \mathbf{x}_{nd}^* - \frac{1}{2} A_{21} \mathbf{x}_d^* \right)^T \left( A_{21} A_{21}^T \right)^{-1} \left( \mathbf{x}_{nd}^* - \frac{1}{2} A_{21} \mathbf{x}_d^* \right) + \mathbf{x}_d^{*T} \mathbf{x}_d^* \quad (2)$$

We make Assumption 1 because we are interested in the change in brain state through control, and consider initial conditions  $\mathbf{x}_d(0) = 0, \mathbf{x}_{nd}(0) = 0$  to be a neutral baseline. Because equation (2) involves only edges from driver to non-driver nodes, we call equation (2) a first-order approximation to the minimum control energy of the non-simplified network equation (1). Importantly, this approximation requires at least as many driver nodes as non-driver nodes for  $A_{21} A_{21}^T$  to be invertible (that is,  $N \geq M$ ). To assess the accuracy of our expression, we look to classic results in the mathematical theory of systems and control<sup>30</sup>, where the spectral properties of the reachability Gramian  $W_R(0, T) = \int_0^T e^{A t} B B^T e^{A^T t} dt$  quantify the minimum amount of energy (Supplementary Section XI A 2) to control the non-simplified network equation (1).

In these brain networks, we observe that the first-order energy approximation is accurate across a range of parameters, which are the magnitude of the adjacency matrix (given by the magnitude of

the largest eigenvalue,  $c = \|\lambda_{\max}\|$ , after multiplying  $A$  by a constant scalar), and the fraction  $d$  of nodes selected as non-driver nodes (Fig. 2e–g). The error remains below approximately 5% for scaling  $c < 1.5$  and non-driver fraction  $d < 0.4$  (Fig. 2e–g). In this paper, we will use these connectomes scaled such that  $c = \|\lambda_{\max}\| = 1$ , and non-driver fraction  $d \leq 0.4$ , to ensure generalizability of our findings to the non-simplified versions of these same networks.

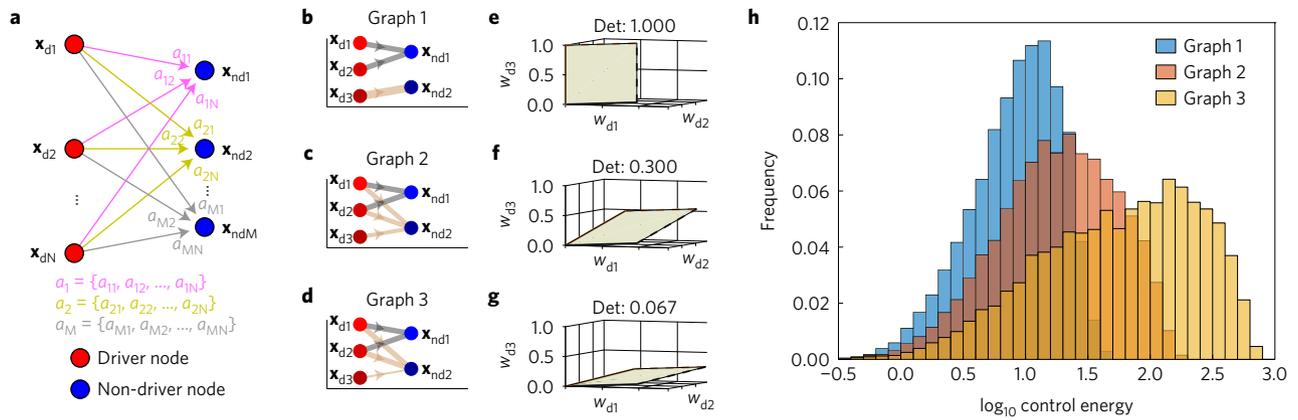
**Determinant of the driver-to-non-driver network**

After deriving a closed-form approximation for the minimal energy to control a network, we seek a physical interpretation of the mathematical features that predict the control energy. We let  $Q = A_{21} A_{21}^T$ , and write equation (2) as

$$E(\mathbf{u}) = 12 \frac{\mathbf{v}_1^T \text{adj}(Q) \mathbf{v}_1}{\det(Q)} + \mathbf{v}_2^T \mathbf{v}_2 \quad (3)$$

where  $\mathbf{v}_1 = \mathbf{x}_{nd}^* - (1/2) A_{21} \mathbf{x}_d^*$  and  $\mathbf{v}_2 = \mathbf{x}_d^*$ , and  $\text{adj}(Q)$  is the adjugate matrix of  $Q$ . We notice that the determinant of  $Q$  acts as a scaling factor for the total energy. This insight is useful because of the geometric interpretation of a Gram matrix determinant. Specifically, let  $\mathbf{a}_i \in \mathbb{R}^{1 \times N}$  be the  $i$ th row of  $A_{21}$  (which we will call the weight vector), representing weights from all  $N$  drivers to the  $i$ th non-driver node (Fig. 3a). Then, the determinant of the Gram matrix  $Q$  is equal to the squared volume of the parallelotope formed by all  $\mathbf{a}_i$ .

To gain an intuition for these results, we show a simple system with three drivers and two non-drivers with varying network topologies in Fig. 3b–d, and their corresponding geometric parallelotopes in Fig. 3e–g with weight vector  $\mathbf{a}_1$  in grey and  $\mathbf{a}_2$  in tan. We also compute the distribution of control energy required to drive each network from initial states  $\mathbf{x}_d = 0, \mathbf{x}_{nd} = 0$  to 10,000 random final states  $\mathbf{x}_{nd}^* \in (-1, 1)^M, \mathbf{x}_d^* \in (-1, 1)^N$  in Fig. 3h. As the



**Figure 3 | Geometric interpretation of simplified, first-order networks with corresponding control energies and trajectories.** **a**, Graph representation of a simplified first-order network containing connections from  $N$  driver nodes in red to  $M$  non-driver nodes in blue. The edges connecting all driver nodes to the  $i$ th non-driver corresponding to the  $i$ th row of  $A_{21}$  are shown in different colours. **b–d**, Graph representations of a network with driver nodes in red, non-driver nodes in blue, weight distribution into non-driver 1 in grey, and weight distribution into non-driver 2 in tan, for dissimilarly distributed weights (**b**), for somewhat similarly distributed weights (**c**), and for very similarly distributed weights (**d**). **e–g**, Geometric representations of the parallelotope formed by the two vectors of weight distributions into non-drivers 1 and 2, with the volume shaded in beige for dissimilarly distributed weights (**e**), for somewhat similarly distributed weights (**f**), and for very similarly distributed weights (**g**). **h**, Base-10 log distribution of control energy required to bring each graph to 10,000 random final states  $\mathbf{x}_{nd}^* \in (-1, 1)^M$ ,  $\mathbf{x}_d^* \in (-1, 1)^N$ .

non-drivers  $x_{nd1}$ ,  $x_{nd2}$  become more similarly connected, the total area of the parallelotope (and corresponding Gram determinant) decreases (Fig. 3e–g), and the control energy increases (Fig. 3h). We note that this determinant relationship persists for any number of nodes where  $N > M$ . We conclude that the similarity between weight vectors generally scales the control energy through  $\det(Q)$ , allowing us to analyse and modify the connectivity of a network with respect to its control energy.

### Identifying energetically favourable control nodes

Here, we further explore the idea of ‘similarity’ between connections  $\mathbf{a}_i$ , to quantify the impact of each individual non-driver on the control energy.

**Topological contributors to control energy.** Our analysis is rooted in the intuition that the edge weights  $\mathbf{a}_i$  that maximize the parallelotope volume, thereby facilitating network control, are large in magnitude and orthogonal to each other. Let  $\lambda_i$  and  $\mathbf{e}_i$  be the eigenvalues and eigenvectors of the matrix  $Q$  in equation (3). We derive in Supplementary Lemma X.6 the equivalent, alternative control energy expression

$$E(\mathbf{u}) = 12 \left( \frac{\sum_{i=1}^M w_i c_i^2}{\sum_{i=1}^M w_i} \right) \left( \sum_{k=1}^M \frac{1}{\|\mathbf{a}_k\|^2 \sin(\theta_k)^2} \right) + \mathbf{v}_2^T \mathbf{v}_2 \quad (4)$$

where  $w_i = \prod_{j \neq i} \lambda_j$ ,  $c_i = \mathbf{e}_i^T \mathbf{v}_1$ , and  $\theta_k$  is the angle formed between  $\mathbf{a}_k$  and the parallelotope formed by  $\mathbf{a}_{j \neq k}$ . We also derive in Supplementary Lemma X.7 the average control energy to reach all random final states drawn uniformly from  $-1$  to  $1$ ,  $\mathbf{x}_{nd}^* \in (-1, 1)^M$ ,  $\mathbf{x}_d^* \in (-1, 1)^N$ , as

$$\mathbb{E}[E(\mathbf{u})] = \frac{1}{3}N + M + 4 \left( \sum_{k=1}^M \frac{1}{\|\mathbf{a}_k\|^2 \sin(\theta_k)^2} \right) \quad (5)$$

For  $N$  drivers and  $M$  non-drivers, we can visualize the  $M$  weight vectors  $\mathbf{a}_k$  as forming a parallelotope in an  $N$ -dimensional space. The variable  $\theta_k$  then represents the angle formed between  $\mathbf{a}_k$  and the parallelotope formed by the remaining  $M - 1$  vectors  $\mathbf{a}_{j \neq k}$ . An example with  $N = 3$ ,  $M = 2$  is shown in Fig. 3e–g, where  $\theta_1 = \theta_2$  is the angle between the tan and grey vectors.

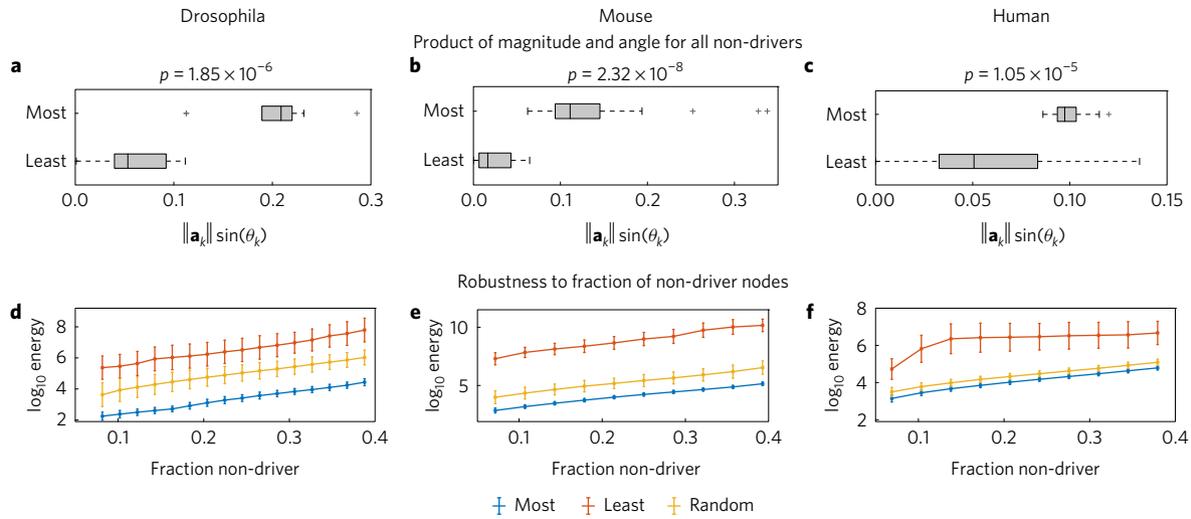
Here, we have segregated the control energy into a task-based  $(\sum_{i=1}^M w_i c_i^2 / \sum_{i=1}^M w_i)$  and topology-based  $(\sum_{k=1}^M (1 / \|\mathbf{a}_k\|^2 \sin(\theta_k)^2))$  term (equation (4)), where the average minimum control energy depends linearly on the topology-based term (equation (5)). This segregation allows us to analyse the topology separate from the specific control task, and shows that each non-driver additively contributes to the total control energy minimally when  $\|\mathbf{a}_i\|$  and  $\sin(\theta_i)$  are large.

**Energetically favourable driver–non-driver sets.** To support this discussion, we used the expression in equation (4) to find the selections of  $M$  non-drivers that minimized and maximized this topology term (see Supplementary Results X B), which we define as the energetically most favourable and energetically least favourable selections, respectively. We show example distributions of each weight vector’s magnitude  $\|\mathbf{a}_k\|$  times angle  $\sin(\theta_k)$  (Fig. 4a–c) between these selections in *Drosophila*, mouse, and human for non-driver fraction 0.2. We observe that the energetically least favourable selections have significantly weaker magnitudes and angles than the most favourable selections.

Next, we demonstrate the utility and robustness of these topological features for control by computing the minimum control energy along the non-simplified networks using the driver and non-driver designations from the simplified networks in equation (4) for a range of non-driver fractions. For each non-driver fraction and species, we computed the control energy to bring the energetically most and least favourable non-driver selections, and 2,000 random non-driver selections to a corresponding set of 2,000 random final states  $\mathbf{x}_{nd}^* \in (-1, 1)^M$ ,  $\mathbf{x}_d^* \in (-1, 1)^N$  (Fig. 4d–f). Across all three species, the most favourable selections require around 0.5–1 orders of magnitude less control energy than the random selections, and 2.5–4 orders of magnitude less control energy than the least favourable selections. This difference indicates an energetic advantage for some configurations of drivers and non-drivers over others.

### Complex brain networks are energetically favourable

Given the relationship between a network’s connectivity and minimum control energy in equation (4), we seek to understand if brain networks are organized along energetically favourable principles. Fundamentally, we ask how well a network’s specific



**Figure 4 | Topological characteristics and energetic performance of networks with energetically favourable and unfavourable topologies. a–c,** Box plots of each non-driver weight vector’s magnitude and angle product ( $\|\mathbf{a}_k\| \sin(\theta_k)$ ) between the energetically most and least favourable networks in the Drosophila (a), mouse (b), and human (c) connectomes, for a non-driver fraction of 0.2 and p-values from a two-sample t-test. **d–f,** Mean and standard deviations of the base-10 log of the minimum control energies required to bring the system to 2,000 random final states  $\mathbf{x}_{nd}^* \in (-1, 1)^M$ ,  $\mathbf{x}_d^* \in (-1, 1)^N$  for each of a range of non-driver fractions for the energetically most favourable, least favourable, and random networks for the Drosophila (d), mouse (e), and human (f).

set of connectivity features  $\|\mathbf{a}_k\|$  and  $\sin(\theta_k)$  combine to minimize the topology-dependent energy term  $\sum_{k=1}^M (1/\|\mathbf{a}_k\|^2 \sin^2(\theta_k))$ . In networks that are not designed along these energetic principles, we expect to see no particular relationship between  $\|\mathbf{a}_k\|$  and  $\sin(\theta_k)$ . In networks that minimize the topology-dependent energy term, we expect a compensatory effect, where non-drivers with small angles have large magnitudes, and vice versa.

To explore the relationship between  $\|\mathbf{a}_k\|$  and  $\sin(\theta_k)$  in brain networks, we selected 10,000 random permutations of non-drivers in each of the Drosophila, mouse, and 10 human connectomes, at non-driver fraction  $d$ . For each permutation, we calculated  $\|\mathbf{a}_k\|$  and  $\sin(\theta_k)$  for every non-driver. Then, we averaged  $\|\mathbf{a}_k\|$  and  $\sin(\theta_k)$  for each non-driver across all permutations, giving us an averaged magnitude  $\|\mathbf{a}_k\|$  and  $\sin(\theta_k)$  for each brain region in each network. Finally, we plotted the averaged  $\sin(\theta_k)$  versus  $\|\mathbf{a}_k\|$  for all brain regions in each network for  $d = 0.2$  (Fig. 5a–c). We find little relationship between the averaged  $\|\mathbf{a}_k\|$  and  $\sin(\theta_k)$  in the Drosophila (Spearman  $\rho = -0.25$ ,  $p = 0.0748$ ), a moderate negative relationship in the mouse ( $\rho = -0.36$ ,  $p = 0.000125$ ), and a strong negative relationship in the human ( $\rho = -0.73$ ,  $p \approx 0$ ). This ordering holds for a wide range of non-driver fractions (Fig. 5d). We graphically demonstrate how this negative  $\sin(\theta_k)$  versus  $\|\mathbf{a}_k\|$  relation might arise in networks, using a simple five-node network with two communities of three and two strongly interconnected sets of nodes (Fig. 5e–g), which has a strong negative relationship (Fig. 5h).

### Network manipulation to facilitate control

Here, we consider network modifications that lead to lower control energies. We focus on the effects of edge deletion since it is often useful in the study of biological systems such as brain<sup>33</sup>, metabolic<sup>34</sup>, and gene regulatory<sup>35</sup> networks. Specifically, we quantify the effect of modifying each edge weight on the determinant in Supplementary Lemma X.5 as

$$\frac{\partial}{\partial A_{21}} \det(Q) = 2 \det(Q) (Q^{-1} A_{21}) \quad (6)$$

and compute the decrease in control energy as a result of deleting edges that maximally increase the determinant.

First, for each species and each of a range of non-driver fractions, we randomly selected 2,000 permutations of non-drivers. For

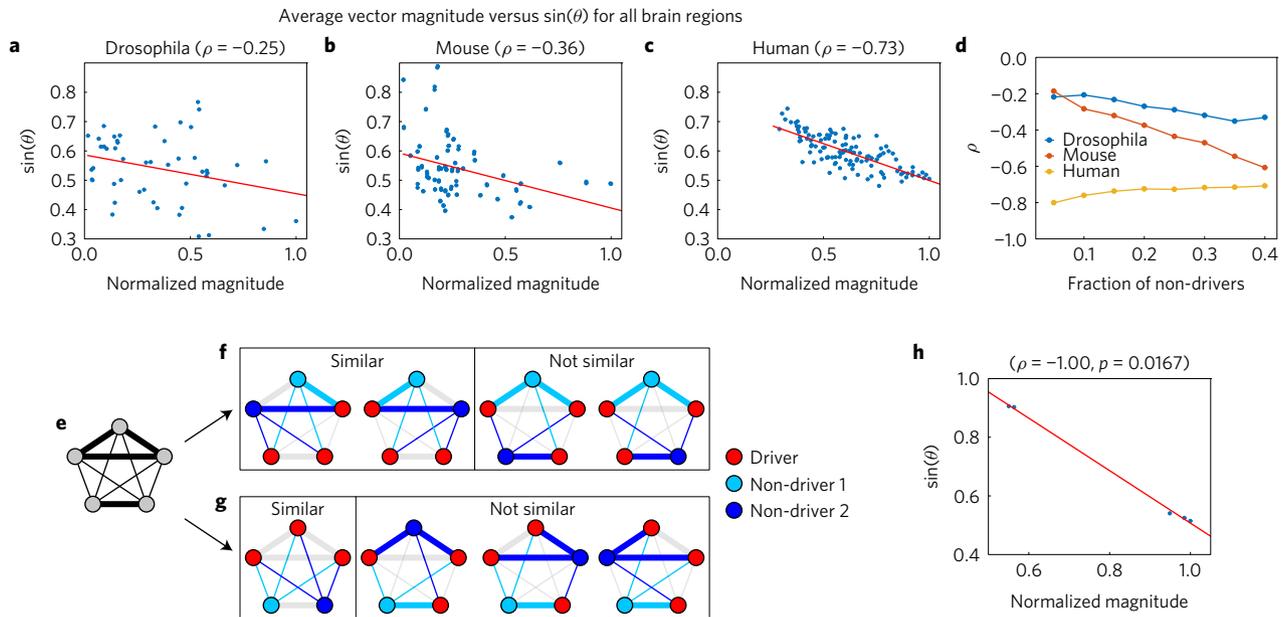
each permutation, we extracted the block matrix  $A_{21}$ , calculated  $2 \det(Q) (Q^{-1} A_{21})$ , and found the element  $a_{ij} \neq 0$  yielding the largest increase in  $\det(Q)$  based on equation (6). We then simulated an edge deletion by setting  $a_{ij} = 0$ , and repeated the process to obtain networks of 1, 2, 3, and 4 deleted edges. Finally, we computed the percentage change in control energy required to bring the non-simplified network from initial states  $\mathbf{x}_{nd}(0) = 0$ ,  $\mathbf{x}_d(0) = 0$ , to final states  $\mathbf{x}_{nd}^* \in (-1, 1)^M$ ,  $\mathbf{x}_d^* \in (-1, 1)^N$  before and after edge deletion (Fig. 6a–d).

As can be seen in Fig. 6a, the removal of one edge can sometimes lead to more than a 10% average reduction in control energy, while the removal of four edges (Fig. 6d) can sometimes lead to more than a 30% reduction. Across most non-driver fractions, the Drosophila experienced greater energy reduction than the mouse, which also experienced greater energy reduction than the human. This corresponds to the previous finding where, because brain networks of these increasingly complex species are already energetically favourably wired, they may not experience as much improvement after modification.

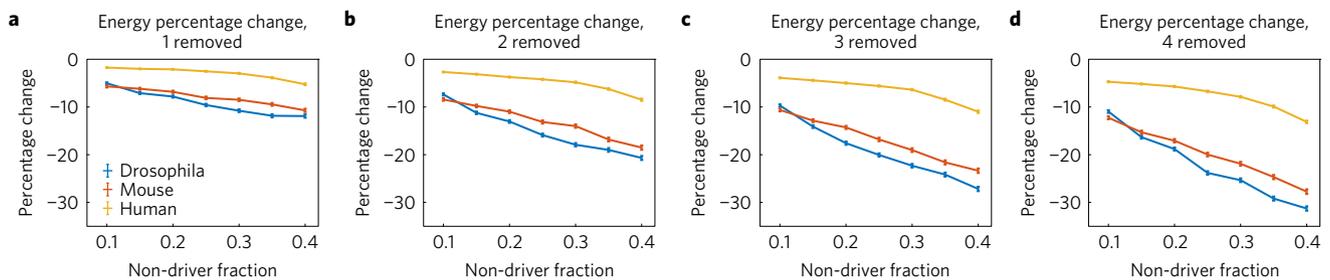
### Contribution and future directions

The control of networked systems is a critical frontier in science, mathematics, and engineering, as it requires a fundamental understanding of the mechanisms that drive network dynamics and subsequently offers the knowledge necessary to intervene in real-world systems to better their outcomes<sup>36</sup>. Although some theoretical predictions exist in nonlinear network systems<sup>15</sup>, the majority of recent advances have been made in the context of linear control<sup>21,22</sup>. Nevertheless, basic intuitions regarding how edge weights impact control have remained elusive. Although spectral analysis of a network’s controllability Gramian<sup>30</sup> yields theoretically useful information about the overall behaviour of the network under control<sup>37</sup>, it is not obvious how specific patterns of connectivity or selections of driver and non-driver nodes contribute to this behaviour. Understanding this relationship is crucial when analysing empirical biological networks such as the brain, where nodes and edges often have known functions<sup>38</sup> that may modulate or influence one other.

A distinct advantage of our approach is the focus on a physically meaningful topological understanding of the principles gov-



**Figure 5 | Energetically favourable organization of topological features in networks.** **a–c**, Average  $\sin(\theta_k)$  versus normalized  $\|\mathbf{x}_k\|$  for each brain region across 10,000 random non-driver selections for a non-driver fraction of 0.2, along with best fit line (red) and corresponding Spearman correlation coefficient in the Drosophila (**a**), mouse (**b**), and human (**c**). **d**, Spearman correlation coefficients in the Drosophila, mouse, and human over 2,000 random non-driver selections for each of a range of non-driver fractions. **e**, Example toy network of five nodes with three strongly interconnected nodes at the top, and two strongly interconnected nodes at the bottom. **f**, Representation of similarity in driver  $\rightarrow$  non-driver connections between Non-Driver 1 (light blue, member of three strongly connected nodes) and all possible selections of Non-Driver 2 (blue). Across all four configurations, Non-Driver 1 has an average of 1.5 strong connections, and 2/4 similarly connected (small angle) configurations. **g**, Similarity in driver  $\rightarrow$  non-driver connections between Non-Driver 1 (light blue, member of two strongly connected nodes) and all selections of Non-Driver 2 (blue). Across all four configurations, Non-Driver 1 has an average of 0.75 strong connections, and 1/4 similarly connected configurations. **h**, Plot of average magnitude versus  $\sin(\theta)$  for the toy network, with Spearman rank correlation coefficient.



**Figure 6 | Modifying the Drosophila, mouse and human connectomes to decrease the minimum energy required for control.** **a–d**, Means and standard errors of percentage change in control energy before and after deleting edges that maximally increase the determinant based on equation (6) over 2,000 control tasks, with initial states  $\mathbf{x}_{nd}(0) = 0$ ,  $\mathbf{x}_d(0) = 0$ , and random final states  $\mathbf{x}_{nd}^* \in (-1, 1)^M$ ,  $\mathbf{x}_d^* \in (-1, 1)^N$ . Non-drivers were randomly selected for a range of non-driver fractions in the Drosophila, mouse, and human connectomes for 1 deletion (**a**), 2 deletions (**b**), 3 deletions (**c**) and 4 deletions (**d**). Standard errors were computed as  $SE = (s/\sqrt{n})$ , where  $s$  is the sample standard deviation over the 2,000 tasks, and  $n = 2,000$ .

erning network control. We map control behaviour to network topology through a simplified network only involving connections from driver to non-driver nodes. This simplification hard-codes the fact that energy can be transmitted directly from drivers to non-drivers along walks of length unity, and is motivated by recent work demonstrating that relatively sparse network representations of complex biological systems<sup>39,40</sup> can contain much of the information needed to understand the system's structure and dynamics<sup>41,42</sup>. Our results inform our understanding of how much first-order connections contribute to the overall dynamics of our network control systems. Moreover, they inform the development of analytical constraints on the accessible state space of a networked system, particularly informing the set of states within which one might seek to push the brain using stimulation paradigms common in the treatment of neurological disorders and psychiatric disease<sup>43,44</sup>. Although many initial studies have examined unconstrained state

spaces<sup>23,25,26</sup>, understanding viable states and state trajectories is critical for the translation of these ideas into the clinic<sup>45</sup>. Further, by formally quantifying the contribution of the network connectivity to the control energy, we lay the groundwork for the optimization of stimulation sites in neural systems, a problem that has received very little theoretical treatment, and is considered one of the current critical challenges in neuroengineering<sup>46</sup>.

Finally, we make strategic, task-agnostic edge deletions that maximally increase the determinant and observe that, even in an overdetermined, unsimplified system ( $N > M$ ), a single edge deletion could produce a profound improvement in the general controllability of a network. This sensitivity suggests that dynamical networks such as the brain can produce fairly drastic changes in dynamical behaviour given minute changes in physiological topology, consistent with observations of critical dynamics in human and animal neurophysiology<sup>47,48</sup>. Moreover, these results also suggest

that minor, targeted structural changes through concussive injury can lead to drastic changes in overall brain function<sup>49,50</sup>, via altering the controllability landscape of the brain<sup>24</sup>. We further observed that these topological modifications were task-agnostic edge deletions, signifying that even in a linear regime, the presence of an unfavourable edge can have a profoundly negative impact on the controllability of a network. We note that it is natural to perform a similar analysis that takes into account the specific tasks  $\mathbf{v}_1, \mathbf{v}_2$  by taking the derivative of the full energy term  $E_{total}$  with respect to  $A_{21}$ , which would optimize the network topology for a specific task, as studied in more detail in ref. 25.

To achieve the most meaningful comparison between species, we analysed only weighted meso-scale whole brain networks. As such, we did not include binary neuronal connectomes (for example, *C. elegans*), and binary or partial connectomes (for example, macaque). As more connectomes become available, we hope to further explore the role of species complexity in network controllability. Until then, we consider the comparison of energetically favourable connectivity between species to be a preliminary excursion into a nuanced evolutionary phenomena. As demonstrated in the significant percentage change in energy after edge deletion, we emphasize that uncertainty in network connectivity has the potential to yield substantial changes in average control energy. Finally, we note that although methodological limitations prevent us from resolving excitatory versus inhibitory connectivity, all results are directly applicable to networks with signed elements. Further important theoretical considerations and methodological limitations pertinent to our approach, linear model of dynamics, optimality of control trajectories, and empirical data sets are discussed in the Supplementary Information.

In closing, we note that the natural direction in which to take this work will be to use higher-order approximations of this framework found in the supplement to gain intuition for the role of complex network topologies (for example, self-loops, cycles) in controlling networks. Moreover, it would be interesting to apply this reduced framework to random graphs and other well-known benchmarks—both from a mathematical perspective<sup>51</sup> and also in the context of neural systems<sup>52,53</sup>—to better understand the phenotypes present in those graph ensembles. Third and finally, informing the design of new networks with these tools may be particularly useful in neuro-morphic computing<sup>54</sup>, materials science<sup>55</sup>, and other contexts where optimal control of physical systems is of paramount importance.

**Data availability.** The data that support the plots within this paper and other findings of this study are available from the corresponding author on request.

Received 1 February 2017; accepted 18 August 2017;  
published online 25 September 2017

## References

- Newman, M. E. J. *Networks: An Introduction* (Oxford Univ. Press, 2010).
- Newman, M. E. J. The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003).
- Watts, D. J. & Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998).
- Simon, H. The architecture of complexity. *Proc. Am. Phil. Soc.* **10**, 467–482 (1962).
- Bassett, D. S. & Sporns, O. Network neuroscience. *Nat. Neurosci.* **20**, 353–364 (2017).
- Bettencourt, L. M., Stephens, G. J., Ham, M. I. & Gross, G. W. Functional structure of cortical neuronal networks grown *in vitro*. *Phys. Rev. E* **75**, 021915 (2007).
- Bassett, D. S. & Bullmore, E. T. Small-world brain networks revisited. *Neuroscientist* <https://doi.org/10.1177/1073858416667720> (2016).
- Sporns, O. & Betzel, R. F. Modular brain networks. *Annu. Rev. Psychol.* **67**, 613–640 (2016).
- Barabasi, A. L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **12**, 56–68 (2011).
- Ching, S., Brown, E. N. & Kramer, M. A. Distributed control in a mean-field cortical network model: implications for seizure suppression. *Phys. Rev. E* **86**, 021920 (2012).
- Khambhati, A. N., Davis, K. A., Lucas, T. H., Litt, B. & Bassett, D. S. Virtual cortical resection reveals push-pull network control preceding seizure evolution. *Neuron* **91**, 1170–1182 (2016).
- Gonen, T. *et al.* Intra-operative multi-site stimulation: expanding methodology for cortical brain mapping of language functions. *PLoS ONE* **12**, e0180740 (2017).
- Mohanty, S. K. & Lakshminarayanan, V. Optical techniques in optogenetics. *J. Mod. Opt.* **62**, 949–970 (2015).
- Sprott, J. C. & Xiong, A. Classifying and quantifying basins of attraction. *Chaos* **25**, 083101 (2015).
- Cornelius, S. P., Kath, W. L. & Motter, A. E. Realistic control of network dynamics. *Nat. Commun.* **4**, 1942 (2013).
- Shine, J. M., Koyejo, O. & Poldrack, R. A. Temporal metastates are associated with differential patterns of time-resolved connectivity, network topology, and attention. *Proc. Natl Acad. Sci. USA* **113**, 9888–9891 (2016).
- Mantzaris, A. V. *et al.* Dynamic network centrality summarizes learning in the human brain. *J. Complex Netw.* **1**, 83–92 (2013).
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J. & Grafton, S. T. Cross-linked structure of network evolution. *Chaos* **24**, 013112 (2014).
- Kalman, R. E. Mathematical description of linear dynamical systems. *J. SIAM Control Ser. A* **1**, 152–192 (1963).
- Lin, C. T. Structural controllability. *IEEE Trans. Autom. Control* **19**, 201–208 (1974).
- Liu, Y. Y., Slotine, J. J. & Barabasi, A. L. Controllability of complex networks. *Nature* **473**, 167–173 (2011).
- Ruths, J. & Ruths, D. Control profiles of complex networks. *Science* **343**, 1373–1376 (2014).
- Gu, S. *et al.* Controllability of structural brain networks. *Nat. Commun.* **6**, 8414 (2015).
- Gu, S. *et al.* Optimal trajectories of brain state transitions. *Neuroimage* **148**, 305–317 (2017).
- Betzel, R. F., Gu, S., Medaglia, J. D., Pasqualetti, F. & Bassett, D. S. Optimally controlling the human connectome: the role of network topology. *Sci. Rep.* **6**, 30770 (2016).
- Muldoon, S. F. *et al.* Stimulation-based control of dynamic brain networks. *PLoS Comput. Biol.* **12**, e1005076 (2016).
- Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
- Rubinov, M., Ypma, R. J., Watson, C. & Bullmore, E. T. Wiring cost and topological participation of the mouse brain connectome. *Proc. Natl Acad. Sci. USA* **112**, 10032–10037 (2015).
- Shih, C. T. *et al.* Connectomics-based analysis of information flow in the *Drosophila* brain. *Curr. Biol.* **25**, 1249–1258 (2015).
- Kailath, T. *Linear Systems* (Prentice-Hall, 1980).
- Fernandez, G. R. On how network architecture determines the dominant patterns of spontaneous neural activity. *PLoS ONE* **3**, e2148 (2008).
- Honey, C. J. *et al.* Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl Acad. Sci. USA* **106**, 2035–2040 (2009).
- Alstott, J., Breakspear, M., Hagmann, P., Cammoun, L. & Sporns, O. Modeling the impact of lesions in the human brain. *PLoS Comput. Biol.* **5**, e1000408 (2009).
- Aristidou, A. A., San, K.-Y. & Bennett, G. N. Modification of central metabolic pathway in *Escherichia coli* to reduce acetate accumulation by heterologous expression of the *Bacillus subtilis* acetolactate synthase gene. *Biotechnol. Bioeng.* **44**, 944–951 (1994).
- Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355 (2014).
- Motter, A. E. Network control theory. *Chaos* **25**, 097621 (2015).
- Pasqualetti, F., Zampieri, S. & Bullo, F. Controllability metrics, limitations and algorithms for complex networks. *IEEE Trans. Control Netw. Syst.* **1**, 40–52 (2014).
- Lanteaume, L. *et al.* Emotion induction after direct intracerebral stimulations of human amygdala. *Cereb. Cortex* **17**, 1307–1313 (2007).
- Park, H., Niida, A., Miyano, S. & Imoto, S. Sparse overlapping group lasso for integrative multi-omics analysis. *J. Comput. Biol.* **22**, 73–84 (2015).
- Liu, Z., Lin, S., Deng, N., McGovern, D. P. & Piantadosi, S. Sparse inverse covariance estimation with  $L_0$  penalty for network construction with omics data. *J. Comput. Biol.* **23**, 192–202 (2016).
- Clauset, A., Moore, C. & Newman, M. E. Hierarchical structure and the prediction of missing links in networks. *Nature* **453**, 98–101 (2008).
- Zhu, B. & Xia, Y. An information-theoretic model for link prediction in complex networks. *Sci. Rep.* **5**, 13707 (2015).
- Chen, H. I. *et al.* Harnessing plasticity for the treatment of neurosurgical disorders: an overview. *World Neurosurg.* **82**, 648–659 (2014).

44. Chrysikou, E. G. & Hamilton, R. H. Noninvasive brain stimulation in the treatment of aphasia: exploring interhemispheric relationships and their implications for neurorehabilitation. *Restor. Neurol. Neurosci.* **29**, 375–394 (2011).
45. Bassett, D. S., Khambhati, A. N. & Grafton, S. T. Emerging frontiers of neuroengineering: a network science of brain connectivity. *Annu. Rev. Biomed. Eng.* **19**, 327–352 (2017).
46. Johnson, M. D. *et al.* Neuromodulation for brain disorders: challenges and opportunities. *IEEE Trans. Biomed. Eng.* **60**, 610–624 (2013).
47. Rubinov, M., Sporns, O., Thivierge, J. P. & Breakspear, M. Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons. *PLoS Comput. Biol.* **7**, e1002038 (2011).
48. Shew, W. L. *et al.* Adaptation to sensory input tunes visual cortex to criticality. *Nat. Phys.* **11**, 659–663 (2015).
49. Caeyenberghs, K., Verhelst, H., Clemente, A. & Wilson, P. H. Mapping the functional connectome in traumatic brain injury: what can graph metrics tell us? *Neuroimage* **S1053–8119**, 30694–30692 (2016).
50. van der Horn, H. J. *et al.* Altered wiring of the human structural connectome in adults with mild traumatic brain injury. *J. Neurotrauma* **34**, 1035–1044 (2017).
51. Bollobas, B. *Random Graphs* (Academic, 1985).
52. Klimm, F., Bassett, D. S., Carlson, J. M. & Mucha, P. J. Resolving structural variability in network models and the brain. *PLoS Comput. Biol.* **10**, e1003491 (2014).
53. Sizemore, A., Giusti, C. & Bassett, D. S. Classification of weighted networks through mesoscale homological features. *J. Complex Netw.* **5**, 245–273 (2017).
54. Pfeil, T. *et al.* Six networks on a universal neuromorphic computing substrate. *Front Neurosci.* **7**, 11 (2013).
55. Giusti, C., Papadopoulos, L., Owens, E. T., Daniels, K. E. & Bassett, D. S. Topological and geometric measurements of force-chain structure. *Phys. Rev. E* **94**, 032909 (2016).

## Acknowledgements

J.Z.K. acknowledges support from National Institutes of Health T32-EB020087, PD: F. W. Wehrli, and the National Science Foundation Graduate Research Fellowship No. DGE-1321851. J.M.S. and D.S.B. acknowledge support from the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan Foundation, the US Army Research Laboratory and the US Army Research Office through contract numbers W911NF-10-2-0022 and W911NF-14-1-0679, the National Institute of Health (2-R01-DC-009209-11, 1R01HD086888-01, R01-MH107235, R01-MH107703, R01MH109520, 1R01NS099348 R21-M MH-106799, and T32-EB020087), the Office of Naval Research, and the National Science Foundation (BCS-1441502, CAREER PHY-1554488, BCS-1631550, and CNS-1626008). A.E.K. and J.M.V. acknowledge support from the US Army Research Laboratory contract number W911NF-10-2-0022. F.P. acknowledges support from the National Science Foundation (BCS-1430280 and BCS 1631112). The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

## Author contributions

J.Z.K., D.S.B. and F.P. wrote and revised the bulk of the manuscript. J.Z.K. developed the mathematical framework and analysed the data with feedback from F.P. and D.S.B. J.M.S. collected the human diffusion data, and A.E.K. processed the data to produce structural connectivity matrices with support from J.M.V.

## Additional information

Supplementary information is available in the [online version of the paper](#). Reprints and permissions information is available online at [www.nature.com/reprints](http://www.nature.com/reprints). Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to D.S.B.

## Competing financial interests

The authors declare no competing financial interests.