

PAPER

A practical guide to methodological considerations in the controllability of structural brain networks

To cite this article: Teresa M Karrer *et al* 2020 *J. Neural Eng.* **17** 026031

View the [article online](#) for updates and enhancements.

Recent citations

- [Benjamin Chiêm *et al*](#)
- [Modeling brain, symptom, and behavior in the winds of change](#)
David M. Lydon-Staley *et al*
- [Gene coexpression patterns predict opiate-induced brain-state transitions](#)
Julia K. Brynildsen *et al*



The Department of Bioengineering at the University of Pittsburgh Swanson School of Engineering invites applications from accomplished individuals with a PhD or equivalent degree in bioengineering, biomedical engineering, or closely related disciplines for an open-rank, tenured/tenure-stream faculty position. We wish to recruit an individual with strong research accomplishments in Translational Bioengineering (i.e., leveraging basic science and engineering knowledge to develop innovative, translatable solutions impacting clinical practice and healthcare), with preference given to research focus on neuro-technologies, imaging, cardiovascular devices, and biomimetic and biorobotic design. It is expected that this individual will complement our current strengths in biomechanics, bioimaging, molecular, cellular, and systems engineering, medical product engineering, neural engineering, and tissue engineering and regenerative medicine. In addition, candidates must be committed to contributing to high quality education of a diverse student body at both the undergraduate and graduate levels.

[CLICK HERE FOR FURTHER DETAILS](#)

To ensure full consideration, applications must be received by June 30, 2019. However, applications will be reviewed as they are received. Early submission is highly encouraged.



PAPER

A practical guide to methodological considerations in the controllability of structural brain networks

RECEIVED
20 August 2019REVISED
30 December 2019ACCEPTED FOR PUBLICATION
22 January 2020PUBLISHED
9 April 2020Teresa M Karrer^{1,2}, Jason Z Kim^{2,13}, Jennifer Stiso^{3,13}, Ari E Kahn³, Fabio Pasqualetti⁴, Ute Habel^{1,5,6} and Danielle S Bassett^{2,7,8,9,10,11,12}¹ Faculty of Medicine, Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen, Germany² Department of Bioengineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104, United States of America³ Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States of America⁴ Department of Mechanical Engineering, University of California, Riverside, CA 92521, United States of America⁵ JARA—Translational Brain Medicine, Aachen, Germany⁶ Institute of Neuroscience and Medicine, JARA-Institute Brain Structure Function Relationship (INM 10), Research Center Jülich, Jülich, Germany⁷ Department of Physics and Astronomy, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104, United States of America⁸ Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States of America⁹ Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States of America¹⁰ Department of Electrical and Systems Engineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104, United States of America¹¹ Santa Fe Institute, Santa Fe, NM 87501, United States of America¹² Author to whom any correspondence should be addressed.¹³ These two authors contributed equally.E-mail: dsb@seas.upenn.edu**Keywords:** network neuroscience, control theory, structural connectivity, diffusion imagingSupplementary material for this article is available [online](#)**Abstract**

Objective. Predicting how the brain can be driven to specific states by means of internal or external control requires a fundamental understanding of the relationship between neural connectivity and activity. Network control theory is a powerful tool from the physical and engineering sciences that can provide insights regarding that relationship; it formalizes the study of how the dynamics of a complex system can arise from its underlying structure of interconnected units. **Approach.** Given the recent use of network control theory in neuroscience, it is now timely to offer a practical guide to methodological considerations in the controllability of structural brain networks. Here we provide a systematic overview of the framework, examine the impact of modeling choices on frequently studied control metrics, and suggest potentially useful theoretical extensions. We ground our discussions, numerical demonstrations, and theoretical advances in a dataset of high-resolution diffusion imaging with 730 diffusion directions acquired over approximately 1 h of scanning from ten healthy young adults. **Main results.** Following a didactic introduction of the theory, we probe how a selection of modeling choices affects four common statistics: average controllability, modal controllability, minimum control energy, and optimal control energy. Next, we extend the current state-of-the-art in two ways: first, by developing an alternative measure of structural connectivity that accounts for radial propagation of activity through abutting tissue, and second, by defining a complementary metric quantifying the complexity of the energy landscape of a system. We close with specific modeling recommendations and a discussion of methodological constraints. **Significance.** Our hope is that this accessible account will inspire the neuroimaging community to more fully exploit the potential of network control theory in tackling pressing questions in cognitive, developmental, and clinical neuroscience.

1. Introduction

The brain is a complex system of interconnected units that dynamically transitions through diverse activation states supporting cognitive function [1]. Understanding the mechanisms and processes that give rise to these trajectories through state space is crucial for intervening in disease to restore cognitive functioning [2]. One relevant factor enabling such rich neural dynamics is the network architecture of the underlying structural substrate [3–5]. Yet, the exact mechanisms by which the physical architecture of the brain both supports and constrains its function remain largely unknown [6–8].

Recent advances in network control theory offer a formal means to study how the temporal dynamics of a complex system emerges from its underlying network structure [9, 10]. Applying this theory to the brain requires that one first builds a network model in which brain regions (nodes) are anatomically connected to one another (edges) [11, 12]. The state of the brain network system is then reflected in the pattern of neurophysiological activity across network nodes, and state trajectories represent the temporal sequence of brain states that the system traverses [13, 14]. With definitions of the network and its state in hand, we can consider the problem of network controllability, which in essence amounts to asking how the system can be driven to specific target states by means of internal or external control input [15]. In the context of the brain, such input can intuitively take the form of electrical stimulation [16–20], task modulation [21–23], or other perturbations from the world or from different portions of the body [24, 25]. Practically, network control theory and its associated toolkit enables us to study the general role of brain regions in controlling neural dynamics in diverse scales and species [26–29], and in both health [30, 31] and disease [32, 33] or injury [34]. For instance, we could ask whether the functional alterations often observed in diseased brain networks can be explained by structural network differences. Moreover, the approach can be used to determine the patterns of input required to induce specific state transitions necessary for behavior [17, 21, 34, 35]. For instance, we could address the question of which brain regions we should target with neurofeedback to elicit specific changes in a large-scale brain network supporting memory performance.

Network control theory offers three primary advantages over traditional approaches to the study of brain network function. First, the multi-modal nature of the theoretical framework explicitly enforces a simultaneous study of brain structure and function, in contrast to approaches that characterize each separately and then assess statistical covariance. Second, network control theory exceeds the often purely descriptive approach of network science [36–38] by building a generative model parameterized by both a network's spatial features and its temporal features

[39]. The model then offers predictions of the brain's response to both endogenous and exogenous input signals. In the case of the former, the model could hypothetically prove useful in understanding how the brain enacts cognitive control to reach task-relevant cognitive states [21, 23, 30]. In the context of the latter, the model could similarly prove useful in informing neuromodulation for the treatment of neurological and psychiatric disorders [39]. Third, initial studies applying network control theory in neuroscience demonstrate that network controllability is a useful marker of brain dynamics.

Network control theory has been used to address a broad range of neuroscientific questions. Relevant studies have quantified the capacity of different brain regions to alter whole-brain dynamics [30], demonstrated that this capacity grows with development [40], and found that controllability is associated with executive functioning [23]. Moreover, the theory has been applied to data collected during invasive neuromodulation regimens to predict response to electrical stimulation in practice [17, 20] and in theory [16]. A part of these initial applications is particularly promising since they explicitly link predictions of network control theory to independently measured neurobiological variables [17, 20, 23]. One of these studies suggests that controllability is associated with cognitive performance by showing that network control theory predicts individual differences in executive function as well as brain activation during a working memory task [23]. Another study relates network control theory to the reconfiguration of functional interactions and to transitions towards better memory encoding brain states, both induced by intracranial stimulation in epilepsy patients [20]. More recent work shows that network control theory can successfully predict activity changes elicited by grid stimulation in epilepsy patients, although the correlation with the true brain state was small in magnitude but statistically significant [17]. These three examples illustrate the utility of network control theory in understanding how spatial characteristics of the brain give rise to its complex function.

In light of the promising applicability of network control theory in neuroscience, we wish to provide a systematic overview of how the framework can be used to study the controllability of neural dynamics. This primer is constructed so as to offer neuroscientists some basic intuitions regarding the foundational concepts, and to guide them through the necessary prerequisites and considerations. For a more technical introduction that nevertheless remains heavily motivated by neuroscience, we refer the interested reader to [41, 42]; and for further information about the underlying mathematics (which remains agnostic to the application domain), we refer the reader to [10, 43]. Because the application of network control theory can be formulated in several ways, we systematically probe how diverging theoretical assumptions and pos-

sible modeling choices influence controllability metrics and the estimated energy of state transitions. For example, we consider discrete and continuous time systems, methods for system stabilization, the time horizon for control, and the set of control nodes. We complement these studies with specific recommendations for best practices, which depend in no small part upon the nature of the neuroscientific question being investigated. To further stimulate research in this exciting field, we suggest a few useful extensions of the theoretical framework, such as alternative estimates of structural connectivity and a complementary metric that quantifies the complexity of the energy landscape.

2. Theoretical framework

2.1. Network control theory

The core of the theoretical framework is the structural network of neurons (or larger neural units) in the brain that allows the activity of a brain region to diffuse and change the activity of connected brain regions (figure 1(A)). Here, we introduce a mathematical model that describes the natural dynamics of a complex linear system (figure 1(B)). Formally, the temporal evolution of network activity is modeled as a linear function of its connectivity:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t), \quad (1)$$

where $\mathbf{x}(t)$ is a vector of size $N \times 1$ that represents the state of the system. Here we operationalize the system's state to reflect the magnitude of the neurophysiological activity of the N brain regions at a single point in time. Over time, $\mathbf{x}(t)$ denotes the state trajectory, which is the temporal sequence of states or activity patterns that is traversed by the system. The adjacency matrix $\mathbf{A}$ is of size $N \times N$, and denotes the relationships between the system elements. Here, we operationalize that relation as the structural connectivity between each pair of brain regions.

Next, we extend this model to account for controlled dynamics, which occur when the brain is induced to deviate from its natural trajectory by the injection of internal or external input signals (figure 1(C)). In this case, the temporal dynamics of a system additionally depends on the control energy injected into a set of nodes across time

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_\kappa \mathbf{u}_\kappa(t). \quad (2)$$

Here, $\mathbf{B}_\kappa$ is a matrix of size $N \times m$ that denotes the set of m control nodes or brain regions into which we wish to inject inputs. For the remainder of the paper, we construct this matrix to represent the independent control of m brain regions using m indicator vectors, each having set only the i th element to 1, corresponding to a control node. If we control all brain regions, $\mathbf{B}_\kappa$ corresponds to the $N \times N$ identity matrix with ones on the diagonal and zeros elsewhere. If we control only a single brain region i , $\mathbf{B}_\kappa$ reduces to a single $N \times 1$ vector with a one in the i th element

and zeros elsewhere. The term $\mathbf{u}_\kappa(t)$ is a vector of m functions of size $m \times 1$ denoting the control input, which is the amount of input injected into each of the m control nodes at each time point t . Over time, $\mathbf{u}_\kappa(t)$ denotes the injected control input over time.

For the interested reader, we wish to provide a few mathematical intuitions that might facilitate a deeper understanding of the presented concepts. By equation (1), the structure of the network determines its dynamic evolution over time. Mathematically, the structural connectivity matrix $\mathbf{A}$ serves as linear operator that maps each state, $\mathbf{x}$, to the rate of change from that state, $\dot{\mathbf{x}}$. This linear transformation can be described in terms of the evolutionary modes of the system consisting of the N eigenvectors of $\mathbf{A}$ and their associated eigenvalues (figure 1(D)). Each eigenvector of $\mathbf{A}$ can be imagined as an axis of the linear transformation which remains invariant over time. Thus, the eigenvectors reflect directions in the state space along which the system independently moves, each characterized by a specific pattern of brain region activity. Each eigenvalue, in turn, determines the rate of growth or decay along its associated eigenvector; that is, each eigenvalue determines how slow or fast the system grows or decays in the direction defined by the eigenvector. Thus, the eigenvalues control the temporal persistence of the set of supported modes of activity.

Especially for the interpretation of results, it is important to keep in mind that the dynamic model is relatively simple and relies on the assumptions of linearity, time invariance, and freedom from noise. Linearity implies that the system evolves linearly over time which is not an accurate reflection of extended dynamics in most neural processes. However, it has been shown that non-linear dynamics can be locally approximated by linear dynamics [44, 45]. Time invariance implies that the system's response does not depend on the time point because both the structural network $\mathbf{A}$ and the control set $\mathbf{B}_\kappa$ are constant over time. This assumption likely holds true for short time scales but could be challenged by long-term structural reorganization, which has been observed across development and adulthood [40, 46, 47]. Freedom from noise implies that all properties of signal propagation are accounted for deterministically by the model. Yet, noise is a feature of neural signals at both small [48–50] and large time scales [51, 52]. Nevertheless, it is customary and reasonable when first developing a mathematical model of a complex system to consider the salient features of the model that do not depend on noise [42, 53, 54].

2.2. Prerequisites

The core of the dynamic model is the network structure that enables activity changes within a particular brain region to diffuse and induce state changes in connected regions. Thus, the first step is to build a structural connectivity network by defining the weighted adjacency matrix $\mathbf{A}$ (figure 1(A)). The

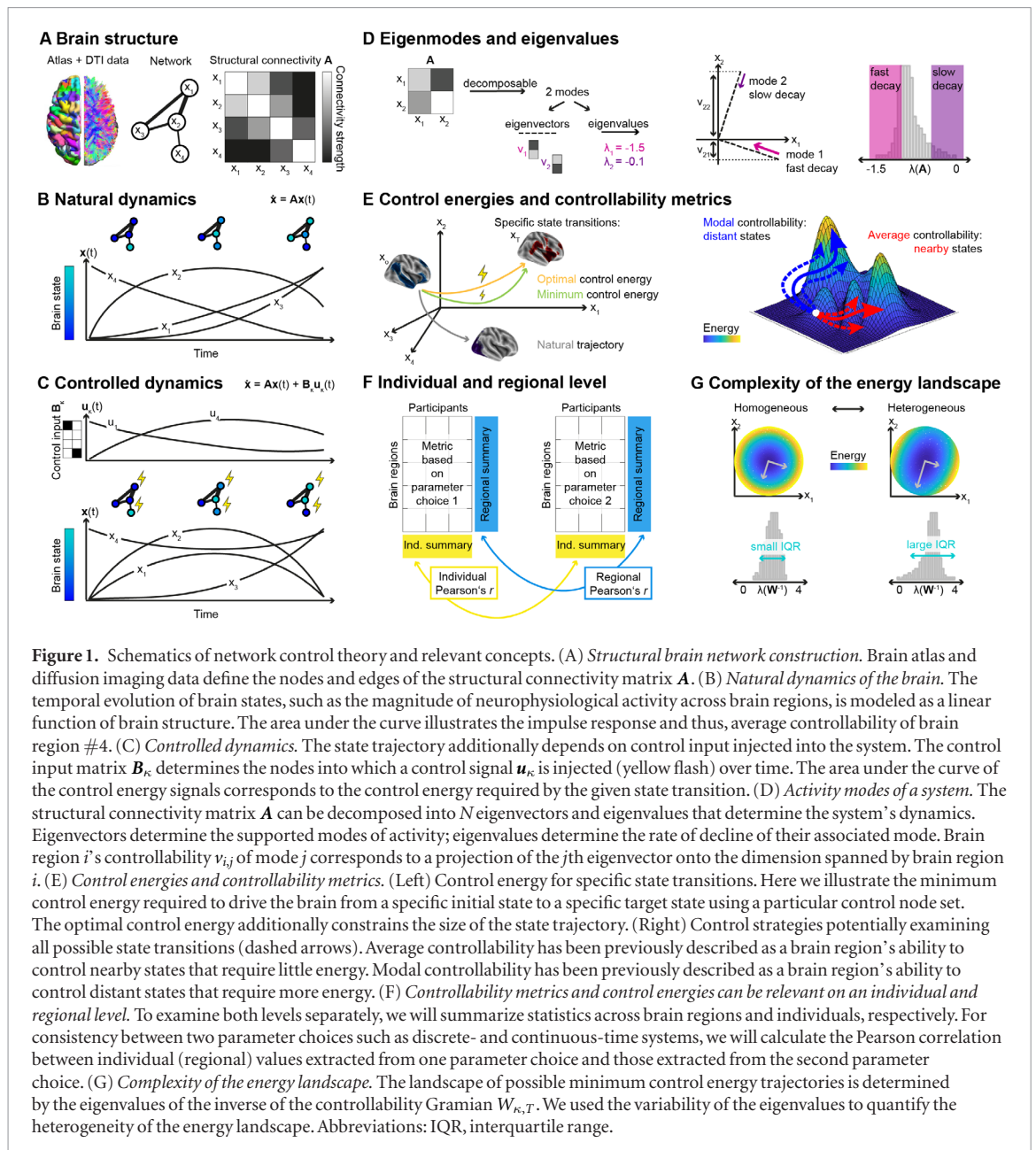


Figure 1. Schematics of network control theory and relevant concepts. (A) *Structural brain network construction.* Brain atlas and diffusion imaging data define the nodes and edges of the structural connectivity matrix A . (B) *Natural dynamics of the brain.* The temporal evolution of brain states, such as the magnitude of neurophysiological activity across brain regions, is modeled as a linear function of brain structure. The area under the curve illustrates the impulse response and thus, average controllability of brain region #4. (C) *Controlled dynamics.* The state trajectory additionally depends on control input injected into the system. The control input matrix B_k determines the nodes into which a control signal u_k is injected (yellow flash) over time. The area under the curve of the control energy signals corresponds to the control energy required by the given state transition. (D) *Activity modes of a system.* The structural connectivity matrix A can be decomposed into N eigenvectors and eigenvalues that determine the system's dynamics. Eigenvectors determine the supported modes of activity; eigenvalues determine the rate of decline of their associated mode. Brain region i 's controllability v_{ij} of mode j corresponds to a projection of the j th eigenvector onto the dimension spanned by brain region i . (E) *Control energies and controllability metrics.* (Left) Control energy for specific state transitions. Here we illustrate the minimum control energy required to drive the brain from a specific initial state to a specific target state using a particular control node set. The optimal control energy additionally constrains the size of the state trajectory. (Right) Control strategies potentially examining all possible state transitions (dashed arrows). Average controllability has been previously described as a brain region's ability to control nearby states that require little energy. Modal controllability has been previously described as a brain region's ability to control distant states that require more energy. (F) *Controllability metrics and control energies can be relevant on an individual and regional level.* To examine both levels separately, we will summarize statistics across brain regions and individuals, respectively. For consistency between two parameter choices such as discrete- and continuous-time systems, we will calculate the Pearson correlation between individual (regional) values extracted from one parameter choice and those extracted from the second parameter choice. (G) *Complexity of the energy landscape.* The landscape of possible minimum control energy trajectories is determined by the eigenvalues of the inverse of the controllability Gramian $W_{k,T}$. We used the variability of the eigenvalues to quantify the heterogeneity of the energy landscape. Abbreviations: IQR, interquartile range.

structural network of human and nonhuman animals can be modeled using a range of spatial scales of neural units and physical links between them [12]. Here, we focus on the construction of the human connectome which requires (i) a brain parcellation that defines the N nodes of the network and (ii) diffusion imaging data that define the strength of structural connectivity A_{ij} between two brain regions i and j . Because self-loops are difficult to resolve using diffusion MRI techniques, we set the diagonal of A to zero. As a further practical note, the sparse nature of human connectomes typically does not require any thresholding of the matrix.

The next step is to pick a time-system that best reflects the neural dynamics under study. Here, we consider two options: discrete and continuous. A discrete-time system assumes that the system evolves in discrete time steps whereas a continuous-time system models continuously changing dynamics. Neural processes

can often not be clearly assigned to one of these categories. Depending on the choice, the modeled dynamics can differ substantially because of their distinct mathematical implementation. More concretely, discrete-time dynamics rely on difference equations whereas continuous-time systems are based on differential equations. Note that we exclusively present formulas for continuous-time systems in the main text; the discrete-time versions can be found in the supplementary formulas (stacks.iop.org/JNE/17/026031/mmedia).

The third step is to choose a method to stabilize the system to avoid its infinite growth over time. Because extremely large brain states are neurobiologically implausible, we normalize the system such that it either approaches the largest supported mode of activity or goes to zero over time:

$$A_{\text{norm}} = \frac{A}{\lambda(A)_{\text{max}} + c} - I. \quad (3)$$

Here, $\mathbf{I}$ denotes the identity matrix of size $N \times N$, and $\lambda(\mathbf{A})_{\max}$ denotes the largest positive eigenvalue of the system. For non-negative matrices, the largest positive eigenvalue is guaranteed to have the same or larger magnitude than all other eigenvalues due to the Perron–Frobenius theorem. Note that this normalization sets the diagonal of $\mathbf{A}$ to -1 . These decaying internal dynamics within each brain region are necessary for the stabilization of the system. To normalize the system, we must specify the parameter c , which determines the rate of stabilization of the system. If $c = 0$, the largest mode of activity is stable and all other modes decay; thus, the system approaches the largest mode over time. If $c > 0$, such as the commonly used choice $c = 1$, all modes decay; thus, the system goes to zero over time. As will become clear in the next section, the latter variant can be especially useful for the computation of average controllability in infinite time as well as modal controllability due to its mathematical definition.

2.3. Optimal control energy

Optimal control energy could intuitively be described as the internal cognitive control or external stimulation effort required to drive a system from one state to another. For instance, a transition from a resting state to a simple visual attention task would likely require less effort compared to a complex working memory task. Note that the effort is not only based on energy costs but also the length of the state transition. To quantify the degree of controllability of a network, we consider an optimal control problem to steer the network from a specific initial state $\mathbf{x}(0) = \mathbf{x}_0$ to a specific target state $\mathbf{x}(T) = \mathbf{x}_T$ over the time horizon T while minimizing a combination of both the length of the state trajectory and the required control energy [34, 55, 56]. Formally, we consider the problem

$$\mathbf{u}(t)_{\kappa}^* = \underset{\mathbf{u}_{\kappa}}{\operatorname{argmin}} J(\mathbf{u}_{\kappa}) = \underset{\mathbf{u}_{\kappa}}{\operatorname{argmin}} \int_0^T ((\mathbf{x}_T - \mathbf{x}(t))^{\top} (\mathbf{x}_T - \mathbf{x}(t)) + \rho \mathbf{u}_{\kappa}(t)^{\top} \mathbf{u}_{\kappa}(t)) dt, \quad (4)$$

where the parameter ρ determines the relative weighting between the costs associated with the length of the state trajectory and input energy. We use the cost function $J(\mathbf{u}(t)_{\kappa}^*)$ to find the unique optimal control input $\mathbf{u}(t)_{\kappa}^*$ which allows us to calculate the *optimal control energy* (figure 1(E)) required by a single brain region i (figure 1(C)):

$$E_i^* = \int_0^T \|\mathbf{u}_i^*(t)\|_2^2 dt, \quad (5)$$

and in total

$$E^* = \sum_{i=1}^m E_i^* = \int_0^T \mathbf{u}_{\kappa}^*(t)^{\top} \mathbf{u}_{\kappa}^*(t) dt. \quad (6)$$

To calculate optimal control energy, we must specify an initial brain state $\mathbf{x}_0$ and a target brain state $\mathbf{x}_T$ by assigning each brain region an initial and target activity

level. For this purpose, we can extract regional activity values directly from functional neuroimaging data such as electrocorticography or magnetic resonance imaging [17, 57, 58], or we can use model-based estimates of task-related activation such as β values from a general linear model [59]. Another option to generate realistic brain states lies in the meta-analysis of large functional neuroimaging databases such as Neurosynth [60] and BrainMap [61]. Additionally, we must specify the control set $\mathbf{B}_{\kappa}$, a set of brain regions into which we wish to inject signals. Theoretically, this choice can vary from controlling a single region to controlling the full brain. The choice of small- to medium-sized control sets, however, can lead to large numerical instabilities that accumulate and bias the results. Thus, we recommend to ensure that the system reached the desired target state. To reduce the numerical error of the calculation, we can also define a relaxed control set $\mathbf{B}_{\kappa}$ by allowing large control input to control regions and small, random inputs to all other brain regions [17]. We must also specify the time horizon T over which the control input is effective. For pragmatic reasons such as the potential translation to real external brain stimulation, the time horizon is usually set to finite time. Empirical evidence on the timeline of the state transition under study could further guide this modelling choice. However, it should be noted that the time horizon is measured in arbitrary units even if brain states are defined by functional imaging data. Finally, we must specify the time step dt to numerically approximate a continuous-time system. Because the numerical simulations converge on truly continuous dynamics as dt approaches zero, the time step dt should be set sufficiently small [34].

The cost function J is motivated by the fact that biological systems might constrain the features of the traversed states, such as their type, diversity, or magnitude. Transitioning through states not too far away from the target state is supposed to avoid extremely large and thus neurobiologically implausible brain state transitions. In the case where no specific assumptions are made on the relative importance of the two constraints and where both the distance and energy values are of a comparable scale, an equal weighting of $\rho = 1$ is a reasonable choice. Depending on our neurobiological assumptions, we can also define alternative cost functions and potentially restrict them to a subset of brain regions [21].

2.4. Minimum control energy

A specific and commonly used subform of optimal control energy is the minimum control energy, which could also intuitively be described as the internal cognitive control or external stimulation effort required to drive a system from one state to another. In contrast to optimal control energy, the effort is only based on energy costs. Minimum control energy is obtained by letting $\rho \rightarrow \infty$ in (4), so that the cost function J accounts only for the energy of

the control input to steer the network from an initial state $\mathbf{x}(0) = \mathbf{x}_0$ to a target state $\mathbf{x}(T) = \mathbf{x}_T$. Thus, we call this metric *minimum control energy* (figure 1(E)). To compute the minimum control energy for a given network, it is convenient to define the controllability Gramian as

$$\mathbf{W}_{\kappa,T} = \int_0^T e^{\mathbf{A}t} \mathbf{B}_{\kappa} \mathbf{B}_{\kappa}^{\top} e^{\mathbf{A}^{\top}t} dt. \quad (7)$$

The eigenvalues of $\mathbf{W}_{\kappa,T}$ can be used to answer several questions regarding the controllability of a network. First, if the smallest eigenvalue of $\mathbf{W}_{\kappa,T}$ is zero, then the network is not controllable. That is, there exist final states $\mathbf{x}_T$ that cannot be reached by any control input, independent of its energy. Second, the magnitude of the smallest eigenvalue of $\mathbf{W}_{\kappa,T}$ is inversely proportional to the largest energy needed to reach a final state. That is, there exists a final state $\mathbf{x}_T$ that can be reached only using inputs whose energy is at least proportional to the inverse of the smallest eigenvalue of $\mathbf{W}_{\kappa,T}$. The foundational papers [30, 62] have shown that brain networks are controllable from any single region; that is, the smallest eigenvalue of $\mathbf{W}_{\kappa,T}$ is greater than zero. However, brain networks require very large control energy; that is, the smallest eigenvalue of $\mathbf{W}_{\kappa,T}$ can be extremely small. It should also be noted that the computation of the smallest eigenvalue of $\mathbf{W}_{\kappa,T}$ tends to be numerically difficult, which motivates the next metric.

2.5. Average controllability

Apart from examining specific state transitions, the theoretical framework also allows us to ask questions regarding the general role of brain regions in controlling neural dynamics. A third metric is obtained by measuring the impulse response of a system, which is the ability of a network to amplify and spread control inputs. More concretely, average controllability equals $\text{Trace}(\mathbf{W}_{\kappa,T})$ and quantifies the energy of the impulse response of a system, which describes how a system naturally evolves over time from some initial condition [43]. Starting from an exclusive activation of the specified control regions, we observe the brain's natural response (figure 1(B)). The larger and more variable this natural response, the more states can be reached with low energy input by controlling this specific set of brain regions. Average controllability is intuitively described as the ability of a set of control nodes to drive the system to easily reachable, nearby states such as an activation of the default mode system enabling a resting or relaxed brain state (figures 1(B) and (E)) [30]. The relation of $\text{Trace}(\mathbf{W}_{\kappa,T})$ and the average input energy $\text{Trace}(\mathbf{W}_{\kappa,T}^{-1})$ is further discussed in the Supplementary Methods.

To calculate average controllability, we must specify the time horizon T , which is the time period over which we wish to observe the impulse response of the system. Note that the units of the time horizon depend on the units of $\mathbf{A}$. To observe the complete impulse

response, we often assume infinite time. Furthermore, we must determine the control set $\mathbf{B}_{\kappa}$, which is the set of brain regions into which control input can be injected. Even if the control set can comprise multiple, and even all nodes, average controllability is often examined for individual brain regions to enable comparison to another single-node metric: most commonly, modal controllability.

2.6. Modal controllability

Next, we introduce the metric *modal controllability*, which was previously described as the ability of a single node to drive the system to distant, more difficult-to-reach states such as an activation of higher-level attention networks supporting flexible task switching (figure 1(E)) [30]. The controllability metric is obtained directly from the eigenvalues and eigenvectors of the network weighted adjacency matrix. In particular, we use

$$\phi_i = \sum_{j=1}^N (1 - (e^{\lambda_j(\mathbf{A})})) v_{ij}^2, \quad (8)$$

as a scaled summary of node i 's ability to control all N modes of the network. Note that we adapted this continuous-time version from the discrete-time version defined in [10] (details are provided in the Supplementary Methods). To calculate modal controllability, we are not required to specify any parameters except the symmetric adjacency matrix $\mathbf{A}$. This metric capitalizes on information housed in the modes of $\mathbf{A}$, as summarized in the eigenvalues λ_j and the matrix of normalized eigenvectors $\mathbf{V} = [v_{ij}]$. Entry v_{ij} is a measure of the controllability of mode $\lambda_j(\mathbf{A})$ from node i that geometrically corresponds to projecting node i onto the eigenvector j (figure 1(D)) [43, 63]. According to this heuristic, the larger the magnitude of the projection, the higher the ability of node i to control mode j . The metric summarizes this notion across all modes, and then scales them by their rate of decline as determined by the eigenvalues. This weighting emphasizes especially fast decaying modes which might on average be more difficult to control because the injected control energy only has a short-term impact. We note that the presented definition of modal controllability is specific for symmetric networks, although it can theoretically be extended to directed networks.

2.7. Boundary controllability

Lastly, we introduce a third commonly used controllability metric. Boundary controllability measures the ability of a set of brain regions to couple and decouple trajectories of disjoint brain regions [10, 30]. Intuitively, brain regions displaying a high boundary controllability lie between network communities and play an important role in segregating and integrating information across different cognitive systems [30]. The exact implementation of the metric differs

across the neuroscientific literature [10, 30, 64, 65], among others due to the diversity of available methods to partition a neural system into network communities [66]. Lacking a consistent definition of boundary controllability of brain regions, we leave the detailed study of this particular metric to future work. Nonetheless, we wish to point the interested reader towards several references that describe and apply the metric in more detail [10, 30, 64, 65].

3. Materials and methods

3.1. Construction of structural brain networks

Based on the diffusion imaging data (acquisition and preprocessing procedure are described in the Supplementary Methods), we constructed a structural brain network for each participant. Consistent with previous work [30, 34, 35, 40], we defined nodes of the network as brain regions according to the 234-node Lausanne atlas (excluding brainstem) [67]. For this purpose, the Lausanne parcels were dilated by 4mm so that the parcels reached down into the white matter enough to ensure accurate sampling of underlying fibers. In the process of dilation, some voxels were assigned to two or more regions of interest; to eradicate this redundancy, we assigned each voxel to the mode of its neighbors [68]. After warping the parcellation into the subject's diffusion space, we quantified the edges of the network as total streamline count connecting a pair of brain regions, corrected for their volume. Overall, we constructed a 233×233 sparse, weighted, and undirected adjacency matrix for each participant with the number of interregional streamlines representing structural connectivity.

3.2. Mapping to cognitive systems

To define neurobiologically meaningful brain states, we capitalized on an established functional brain atlas [69]. By clustering the resting state functional magnetic resonance imaging data of 1000 healthy adults, Yeo *et al* identified seven cognitive systems, each consisting of a set of distributed brain regions that are functionally coupled [69]. The functional parcellation comprises visual (VIS), somatomotor (SOM), dorsal attention (DOR), ventral attention (VEN), limbic (LIM), frontoparietal control (FPC), and default mode (DM) systems. To link the functional and anatomical atlases, we mapped each brain region to the cognitive system with the highest spatial overlap as reported previously [21, 47]. More concretely, each Lausanne parcel was assigned to the cognitive system that was most frequently associated with its voxels as defined by the purity index. Subcortical regions were summarized in an eighth, subcortical system (SC).

3.3. Probing different modeling choices

We used the structural connectivity matrices of our sample to probe the impact of several modeling choices on average and modal controllability, and

on minimum and optimal control energy. Since network control theory can be utilized to examine controllability differences in both individuals and brain regions, we separately studied the metrics on an individual and regional level (figure 1(F); details are provided in the Supplementary Methods). Note that a part of these results could have been derived from the theory *a priori*. Nonetheless, we consider this illustration as useful for the neuroscience community to gain a better understanding of network control theory and the behavior of control metrics depending on different modeling choices. In each examination of the impact of one specific modeling choice, we systematically varied the modeling choice of interest while keeping all other parameters constant. In the examination of the impact of time system, for instance, we calculated controllability metrics separately for discrete and continuous time-systems while keeping time horizon, normalization parameter, and control set constant. Variable parameter ranges were guided by the common choice and the principles of completeness or convergence (further details are provided in the supplementary methods). Constant modeling choices were guided by the modeling choices most commonly used in the literature [30, 34, 35, 40]. Concretely, we employed a simplified noise-free linear continuous-time and time-invariant network model, stabilized using $c = 1$. When estimating average controllability, we set the time horizon T to infinite time. When estimating control energies, we used $T = 3$, approximated by 1000 time steps i.e. $dt = 0.001$. When the system matrix A is stable, the controllability Gramian equation converges as T approaches infinity. In this case, the Gramian can be computed algebraically by solving the Lyapunov equation.

Motivated by the questions most relevant to each approach, we calculated average and modal controllability for each brain region based on single-node control sets. Control energies, however, were based on full brain control, which could either represent cognitive control exerted by the whole brain internally [70] or external control exerted by complex interventions such as a combination of psychotropic drugs and psychotherapy. We exercise full brain control because this choice (i) has been used in previous work [21, 35, 57], (ii) avoids additional neurobiological assumptions due to control subset selection, and (iii) is numerically tractable. We examined control energies required for the transition between a specific initial and target state of the brain. For the definition of these brain states, we capitalized on previously defined functional systems. For each specific brain state, regions belonging to the activated cognitive system were set to one, whereas all other brain regions were set to zero. Even if such artificial state representations likely do not represent actual neural dynamics, this simplification provides a useful starting point for the systematic examination of modeling choices [21, 34, 35]. To be thorough, we also performed an additional simulation demonstrating

a high consistency between binary and continuously represented initial and target brain states (Sfigure 1). We simulated state transitions from an active default mode system representing the initial brain state to the activation of six different cognitive systems representing the target brain states [69]. Except for the section on full versus partial control, we averaged across the examined state transitions. For optimal control energy, we set the relative energy weight $\rho = 1$. In the restricted set of state transitions we investigated, minimum and optimal control energy yielded highly similar results. To avoid redundancy, we report the results on optimal control energy in the supplementary results (Sfigure 2–7). Nevertheless, we point out deviating results of optimal control energy in the main text. Furthermore, we provide a detailed examination of how control metrics are empirically related to each other in brain networks in the supplementary results (Sfigure 8).

3.4. Construction of spatial adjacency network

In addition to diffusing along white matter fibers, neural signals could potentially also diffuse between spatially adjacent brain regions. In other words, physical contact between two regions can be seen as a form of structural connectivity. To examine this complementary measure of structural connectivity, we generated brain networks, $\mathbf{S}$, based on the amount of shared neighborhood between two brain regions. We defined the edges of $\mathbf{S}$ as the number of face-touching voxels between two parcels of the Lausanne-atlas warped into subject space. In addition to studying each structural matrix separately, we also exploit the combined information of both measures by constructing the matrix $\mathbf{AS}$ as an average of $\mathbf{A}$ and $\mathbf{S}$. Because both diffusion and adjacency measures are expressed in arbitrary units and the actual scaling might impact controllability metrics, we scaled $\mathbf{S}$ and $\mathbf{AS}$ to the range of $\mathbf{A}$. We tested the effects of the structural connectivity types and their binarized version in a repeated measures ANOVA with two within-subject factors. To ensure that the effects of matrix type and binarization were not exclusively based on different edge weight distributions [65], we verified the results by sampling edge weights of $\mathbf{S}$ and $\mathbf{AS}$ from the distribution of $\mathbf{A}$ while preserving their rank order.

3.5. Controllability of fast and slow dynamics

We capitalized on the concept of modal controllability to probe the ability of a brain region to control a specific set of temporal dynamics such as fast and slow modes [17, 71]. Instead of summarizing across all modes that a system supports, we restricted the calculation of modal controllability to a subset of fastest (slowest) modes. We define transient (persistent) modal controllability as the ability of a brain region to control fast (slow) modes. The temporal dynamics of modes are determined by the magnitude of their eigenvalues.

In continuous-time systems, large (small) eigenvalues relate to quickly (slowly) decaying modes. The lack of a formal definition of fast and slow dynamics requires the choice of a threshold that specifies the subset of modes (figure 1(D)). We systematically probed the influence of threshold on a brain region's ability to control different temporal dynamics by calculating transient and persistent modal controllability using the 10%, 20%, 30%, 40%, and 50% fastest and slowest modes. To disentangle these overlapping control tasks, we additionally summarized the ability of each brain region to control a specific interval of modes using the unscaled eigenvector matrix $\mathbf{V}$. For this purpose, we summarized the unscaled eigenvector matrix $\mathbf{V}$ into 10 and 2 intervals respectively; the partitioning is based on the rate of decay of the eigenvectors, that is their associated eigenvalues. This separation enabled a comparison to persistent and transient modal controllability based on a cut-off of 10% and 50%, respectively.

3.6. Definition of complexity of the energy landscape

The control trajectories from any initial state to any target state span the energy landscape of a dynamic system. A homogeneous energy landscape could intuitively indicate that different brain states can be reached by similar control efforts. This characteristic could be either advantageous because it enables access to a diversity of brain states or disadvantageous because unhealthy states are among those that are easily accessed. The heterogeneity of the minimum control energy landscape is determined by the eigenvalues of the inverse of the controllability Gramian [43]. Note that we assume independent control from all brain regions because the inverse of the Gramian is often ill-conditioned for small control sets [30]. We capitalized on the variability of the eigenvalues to quantify the complexity of the minimum control energy landscape of a brain network, that is how the magnitude of the minimum control energy varies across all possible state transitions (figure 1(G)). To account for the observed skewness of the distribution, we adopted the interquartile range as a measure of variability. Formally, we define the complexity of the energy landscape as the difference between the 75th and 25th percentile of the eigenvalue distribution of the inverted controllability Gramian

$$C_{\kappa,T} = P_{75}(\lambda_{\mathbf{W}_{\kappa,T}^{-1}}) - P_{25}(\lambda_{\mathbf{W}_{\kappa,T}^{-1}}). \quad (9)$$

We calculated the complexity of the energy landscape for each participant based on an infinite-time controllability Gramian. Then, we tested the complexity of the energy landscape of the brain network against three null models preserving distinct network characteristics. The topological null model preserved degree and strength distribution by iteratively switching connections between randomly selected edge pairs and subsequently associating the connections with the

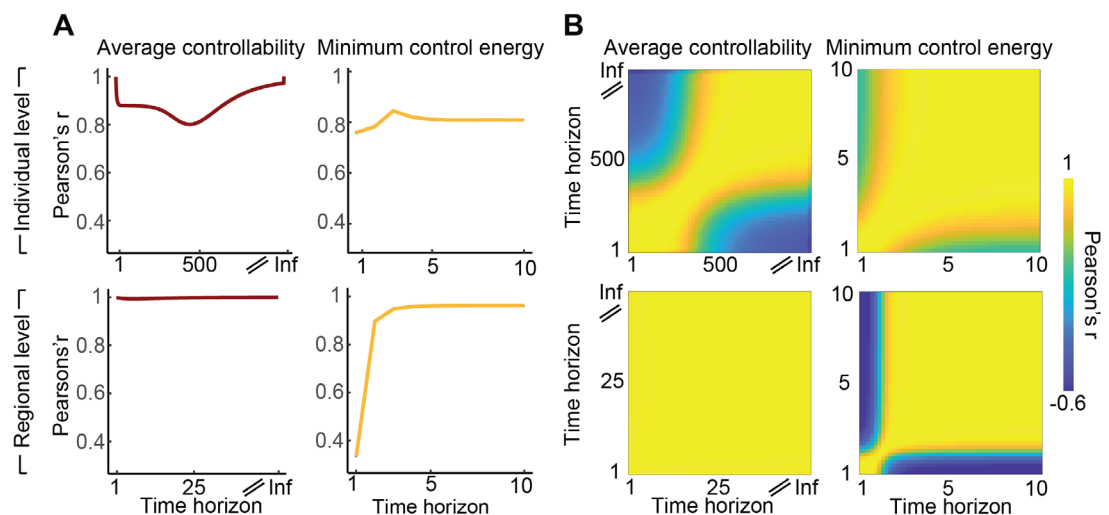


Figure 2. Consistency of metrics across time. (A) Consistency of average controllability and minimum control energy across time systems. Pearson correlation coefficient between a given metric estimated for discrete- versus continuous-time systems across a range of time horizons T . (Left) Average controllability; (right) minimum control energy. (B) Consistency of average controllability and minimum control energy in a continuous-time system for any two choices of time horizon. Heat maps depict correlation matrix of different time horizons. Each heat map entry corresponds to the Pearson correlation of a metric based on two different time horizon choices. (Top) Pearson correlation of individual metrics averaged across brain regions. (Bottom) Pearson correlation of regional metrics averaged across participants. From these results, we deduce that average controllability and minimum control energy differ qualitatively for discrete- versus continuous-time systems when comparing estimates from short time horizons versus from longer time horizons.

empirically observed edge weights [72]. The spatial null model preserved the relationship between Euclidean distance on the edge weights by adding the initially removed distance effects to the randomly rewired graph [73]. The combined null model preserved both the strength distribution and spatial embedding of the brain networks by approximating the observed strength distributions and effects of Euclidean distance on the edge weights [73]. Overall, we generated 1000 random instantiations of each null model.

4. Results

In the application of network control theory, we can rely on different neurobiological assumptions that are reflected in our modeling decisions. We begin with an examination of the impact of different modeling choices, before investigating several proposed model extensions.

4.1. Consistency across time systems

When examining how the brain's architecture gives rise to its complex dynamics by means of network control theory, one of the first modeling steps represents the type of the dynamic model. We can either assume that the neural dynamics evolve in discrete time steps or continuously. In light of potentially distinct dynamics of discrete- and continuous-time systems, we initially examined the consistency of minimum control energy, average controllability, and modal controllability across time systems. For this purpose, we calculated the Pearson correlation of each metric between discrete- and continuous-time systems, separately summarized across brain regions and individuals,

and—if applicable—for different time horizons T (figure 2(A)). Average controllability showed a high consistency across time systems (individual level: $r_{min} = 0.80$, $p = 5 \times 10^{-3}$; regional level $r_{min} = 0.99$, $p = 5 \times 10^{-221}$), particularly for time horizons close to zero or infinity. Likewise, modal controllability demonstrated a high consistency across time systems (individual level: $r = 0.99$, $p = 3 \times 10^{-11}$; regional level $r = 1.0$, $p = 2 \times 10^{-16}$). Minimum control energy, however, was less consistent across time systems (individual level: $r_{min} = 0.77$, $p = 0.01$; regional level $r_{min} = 0.33$, $p = 2 \times 10^{-7}$), particularly for short time horizons. The observed results are in line with theoretical considerations that suggest a convergence of discrete- and continuous-time systems for infinite time. Overall, the consistency across discrete- and continuous-time systems was high but depended on the metric, the observation level, and the chosen time horizon.

4.2. Consistency across time horizons

Network control theory might lend itself particularly well to evaluate how local perturbations of the brain, for instance elicited by deep brain stimulation or transcranial magnetic stimulation, affect whole brain dynamics. In such a setting we might be interested in assessing different temporal scales of brain stimulation such as the effect of stimulation in the short term or in the long term. This question prompts the examination of the time horizon of the injected signal as another early modeling decision. We addressed this question by quantifying the Pearson correlation between values estimated for one time horizon T and for another time horizon T' , separately averaged across brain regions or

across individuals (figure 2(B)). We first noted that the time horizon affected the scaling of the metrics (Sfigure 2(A)). More specifically, average controllability monotonically increased in magnitude with larger time horizons. This observation is consistent with the theory—because we observed the impulse response of the system for a longer time interval, the magnitude of the impulse response was larger. Minimum control energy monotonically decreased with larger time horizons; this theoretically derivable relation is intuitive when we consider the fact that longer time horizons allow the system to capitalize on its own natural dynamics, thereby demanding less exogenous control input. In contrast, optimal control energy first rapidly decreased and then slightly increased with larger time horizons (Sfigure 2(A)). The increasing amount of optimal control energy might be required to additionally constrain the distance of traversed brain states over longer time horizons. In general, we found a high consistency between the metrics across a wide range of examined time horizons. However, smaller time horizons demonstrated a different control regime in which average controllability (individual level: $r_{min} = -0.56$, $p = 0.09$; regional level ($r_{min} = 0.88$, $p = 7 \times 10^{-78}$) and minimum control energy (individual level: $r_{min} = 0.28$, $p = 0.44$; regional level ($r_{min} = -0.71$, $p = 6 \times 10^{-37}$) were partly anti-correlated with the corresponding metrics in larger time horizons. In sum, short time horizons induced an alternative control regime in average controllability and minimum control energy compared to longer time horizons.

4.3. Impact of normalization

The normalization step represents another modeling decision that is related to time. For mathematical reasons, we often assume the neural dynamics to diminish and stabilize over time. Neurobiological considerations determine the degree of normalization; that is, how fast or slow we assume the neural system to stabilize. To investigate the effect of normalization on controllability metrics and control energies, we calculated average controllability, modal controllability, and minimum control energy for different choices of the normalization parameter c . At both individual and regional levels, we first observed that with increasing c , average controllability decreased whereas modal controllability and minimum control energy increased (Sfigure 3). Next, we investigated the consistency of the metrics across different manners of normalization by quantifying the Pearson correlation between metrics for two choices of the normalization parameter c , separately summarized across brain regions (figure 3(A)) and individuals (Sfigure 4(A)). In both cases, we observed two different control regimes depending on small ($c = 0.1$ to $c = 10^2$; figure 3(B)) and large ($c = 10^4$ to $c = 10^6$; figure 3(C)) normalization parameters. Within each regime, the results were highly consistent independent of the normalization

parameter c . Between both regimes, however, the consistency in average controllability (individual level: $r_{min} = -0.19$, $p = 0.61$; regional level: $r_{min} = 0.86$, $p = 1 \times 10^{-69}$), modal controllability (individual level: $r_{min} = 0.29$, $p = 0.41$; regional level: $r_{min} = 0.99$, $p = 7 \times 10^{-320}$), and minimum control energy (individual level: $r_{min} = 0.87$, $p = 2 \times 10^{-3}$; regional level: $r_{min} = 0.81$, $p = 6 \times 10^{-56}$) was reduced. This alternative control regime is due to a faster stabilization of the system. As directly follows from equation (3), the increase of the normalization parameter c leads to an increased decay rate of the slowest mode which, in turn, means a faster stabilization of the system. Taken together, a faster stabilization of the system introduced an alternative control regime that particularly affected controllability metrics.

4.4. Impact of control set size

In the study of the effects of brain stimulation on brain activity, we can also ask how many and which brain regions we should control in order to drive the system to a, for instance healthy, state. More concretely, we could compare the effects of targeting a specific neural circuit to the effects of whole-brain stimulation. This motivates the examination of a final modeling choice: the number of controlled brain regions. That is, we studied the number of brain regions into which we wish to inject signals. To probe the effect of control set size on minimum control energy, we generated random control sets of a varying number of brain regions that allow for a control input ranging from single-node to full-brain control. We then proceeded by testing the impact on minimum control energy and the numerical error in six brain state transitions: from activation of the default mode to activation of six canonical cognitive systems as defined by Yeo *et al* [69]. Importantly, the numerical error was reasonably small ($< 1 \times 10^{-6}$) when we controlled at least 28.3%–29.6% of brain regions ($N_{VIS} = 66$, $N_{SOM} = 67$, $N_{DOR} = 67$, $N_{VEN} = 68$, $N_{LIM} = 69$, $N_{FPC} = 68$), increasing our confidence in the results. We observed that minimum control energy and the numerical error decreased exponentially with increasing control set size, with differences depending on specific state transitions and control sets (Sfigure 5(A)). Intuitively, the control of a larger number of brain regions required less control energy. The exponential relationship between control energy and control node set can also be mathematically derived [10].

Next, we were interested in how control and state trajectories differ in partial- compared to full-brain control sets. We calculated the minimum control energy trajectory and the distance between the state trajectory and the target state for the same six state transitions controlling all versus randomly drawn sets of 150 brain regions. In full-brain control, we observed an exponential increase in energy (figure 4(A)) and an approximately linear decrease in the distance between

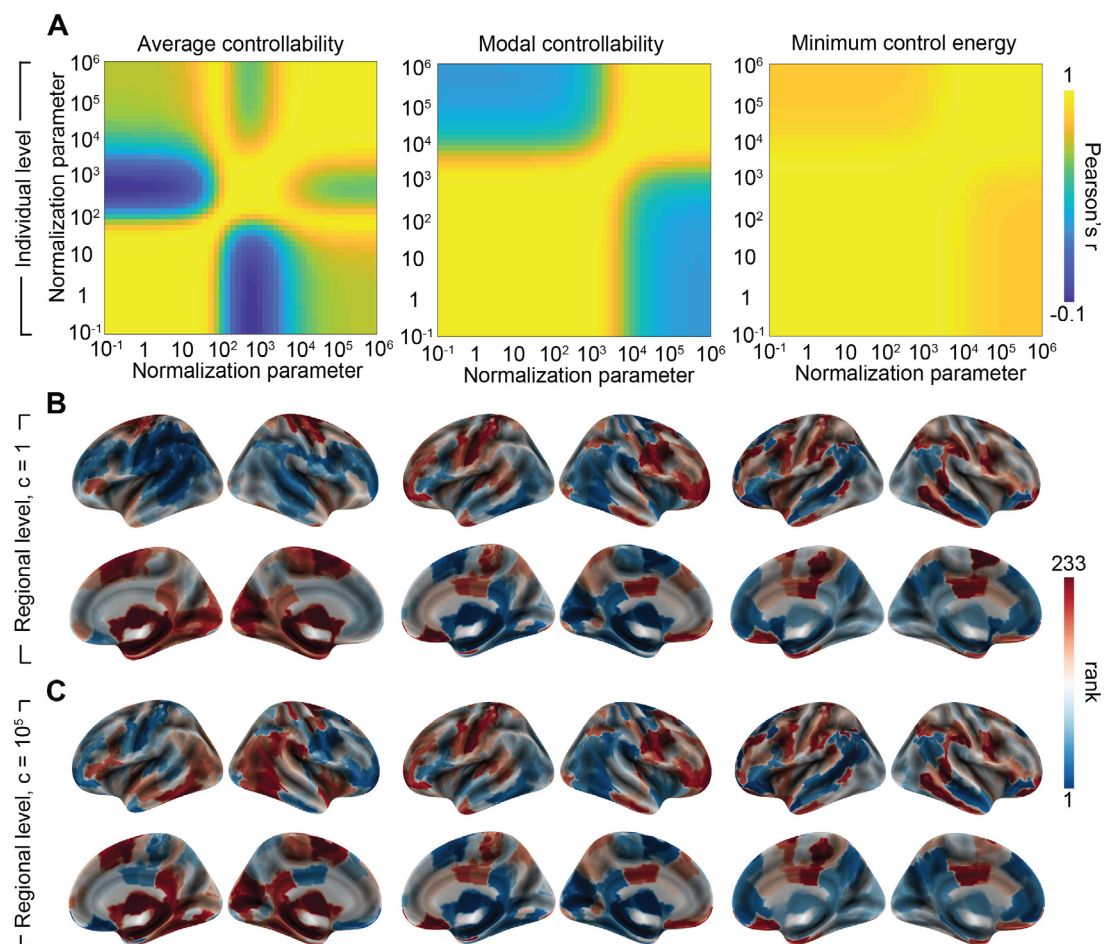
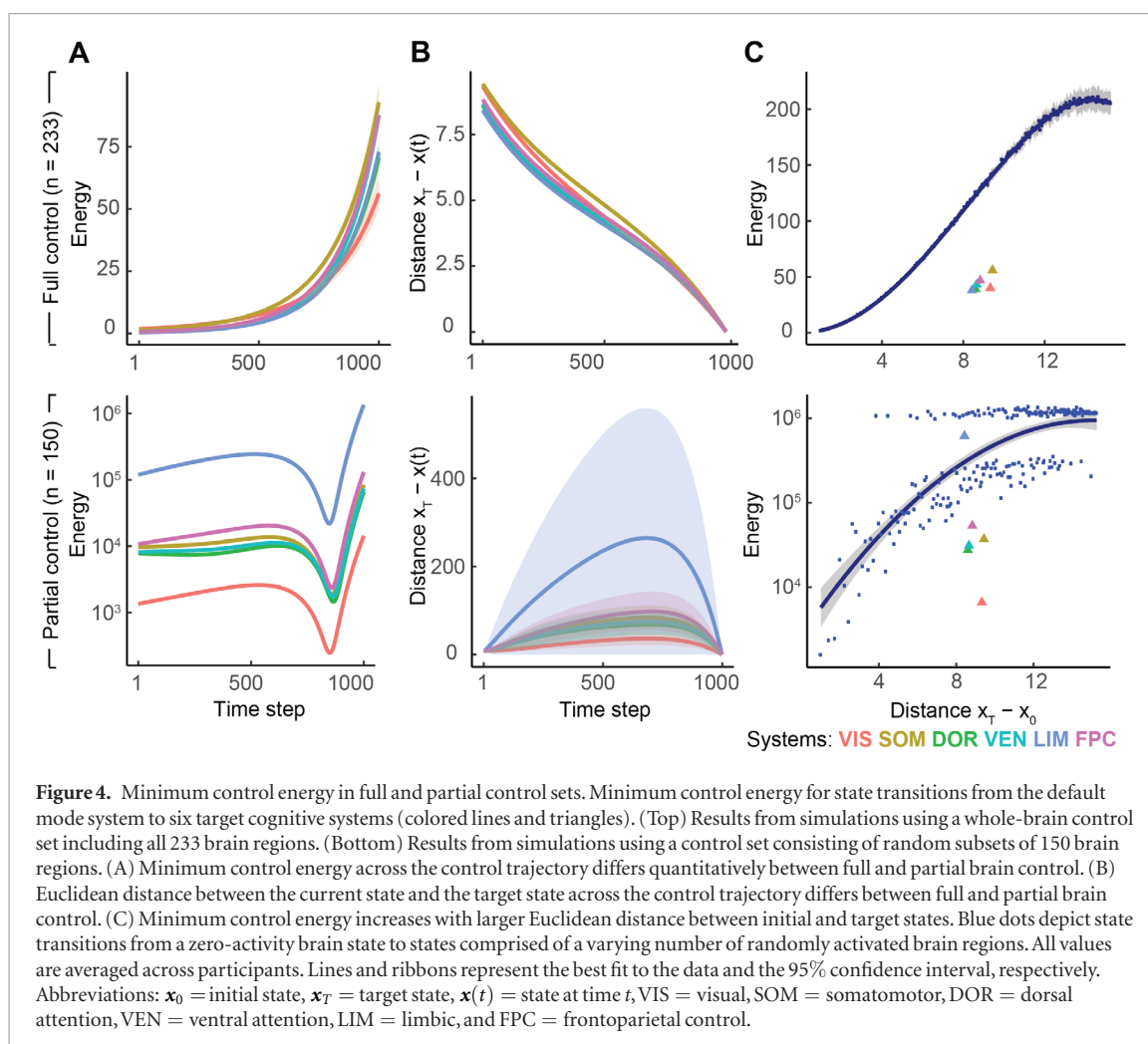


Figure 3. Different control regimes depending on normalization. Consistency of (left) average controllability, (middle) modal controllability, and (right) minimum control energy for different choices of the normalization parameter c . (A) Heat maps depict correlation matrices of different normalization parameters. Each heat map entry corresponds to the Pearson correlation between a metric calculated using one normalization parameter and the same metric calculated using a second normalization parameter. Pearson correlation of individual metrics summarized across brain regions. Normalizing the system such that it stabilizes faster introduces a different control regime. (B) Regional controllability metrics and control energy projected onto the brain surface using a normalization parameter $c = 1$. (C) Analogously, the same metrics plotted onto the brain surface using a normalization parameter $c = 10^5$. Together, panels (B) and (C) illustrate the fact that small and large choices of the normalization parameter can induce two different control regimes. Within each regime, the resulting controllability metrics were highly consistent whereas the consistency between both regimes was reduced. Note: metric values are ranked for visualization purposes only.

current and target state (figure 4(B)) across the control horizon. When we controlled only a part of the brain, control and state trajectories differed considerably. For instance, instead of taking the direct route through the state space, the system traversed more distant states before it reached the target state. Theoretical work has indeed shown that such non-local trajectories generally emerge if only a subset of nodes is controlled [74].

Finally, we wished to study the effect of distance between initial and target state on minimum control energy. Because the set of state transitions that we studied lacked sufficient variability in these distances, we additionally simulated trajectories from a zero-activity initial state to random target states with a varying size of brain regions activated. We found a monotonic increase of minimum control energy with increasing distance between initial and target states (figure 4(C)), which is in line with theoretical considerations. When employing a partial control set, a subset of the random state transitions required massive amounts

of control energy. A further exploration revealed that these hardly controllable state transitions involved an activation of two weakly connected limbic regions that were not part of the random control set. Similarly, the six state transitions likely required less control on average because the activation of densely connected cognitive systems is an easier control task than the activation of randomly chosen regions in target states of equal distance. This could be explained by the fact that densely connected brain regions influence each other more strongly in their neurophysiological activity than loosely connected brain regions. Thus, it is a very difficult control task to activate brain regions that are only loosely connected to the other, active brain regions, particularly if we cannot directly control the target brain regions. Similarly, due to their stronger mutual influence, it is easier to reach a similar activation level in densely connected brain regions than in a randomly chosen set of regions. The findings in optimal control energy were highly similar, even if the exact control



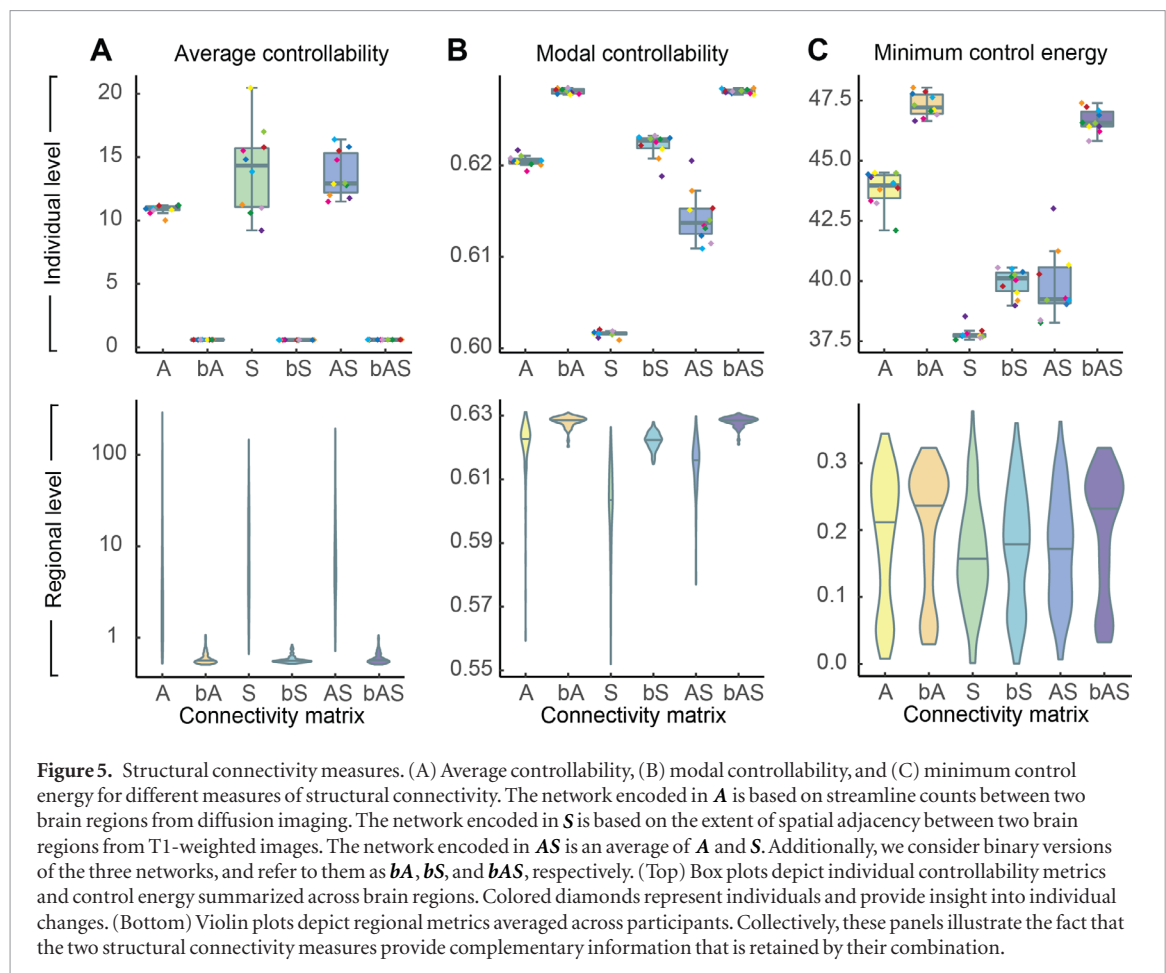
and state trajectories were different (Sfigure 6). Overall, state and control trajectories differed substantially depending on which brain regions were allowed to receive energy input.

4.5. Structural connectivity measures

After systematically examining the impact of diverse modeling choices, we wished to provide several, potentially useful extensions of the theoretical framework. We begin with a consideration of the architecture of the brain which represents the core of network control theory. Thus, it is particularly relevant how we define the inter-connections between brain regions. Typically used DTI data do not take into account the fact that the signal can theoretically diffuse via physical contact between two brain regions. To evaluate the consequences of different forms of the adjacency matrix reflecting different modes of signal propagation in the brain, we additionally built structural connectivity networks based on the amount of shared neighborhood between two brain regions. Then, we calculated controllability metrics and control energies for the two alternative measures of structural connectivity, their combination, and their binarized versions (figure 5). We first examined the similarity in controllability of structural networks based on diffusion imaging (**A**) and based on spatial adjacency

(**S**). Between **A** and **S**, we found small- to medium-sized Pearson correlations in average controllability (individual level: $r = 0.02$, $p = 0.95$; regional level: $r = -0.01$, $p = 0.92$), modal controllability (individual level: $r = -0.15$, $p = 0.67$; regional level: $r = 0.41$, $p = 10 \times 10^{-11}$), and minimum control energy (individual level: $r = 0.36$, $p = 0.31$; regional level: $r = 0.64$, $p = 2 \times 10^{-16}$). Thus, the two measures of structural connectivity provide complementary information. Next, we quantified the effect of binarization, matrix type (**A** versus **S**), and their combination **AS**, on controllability metrics and control energy. Repeated measures ANOVAs revealed significant main effects of matrix type and binarization on a Bonferroni-corrected level of $\alpha = 0.01$ except for the effect of matrix type on average controllability (individual level: $F = 5.98$, $p = 5 \times 10^{-3}$; regional level $F = 1.30$, $p = 0.27$).

To ensure that these results were not exclusively due to different edge weight distributions, we verified these results using **S** and **AS** based on the same edge weight distribution as **A**. When we examined the effects in more detail, we observed that the binarization reduced the absolute values and variance of average controllability on both the regional and individual level, whereas modal controllability displayed a reverse effect. This pattern of results is in line with findings



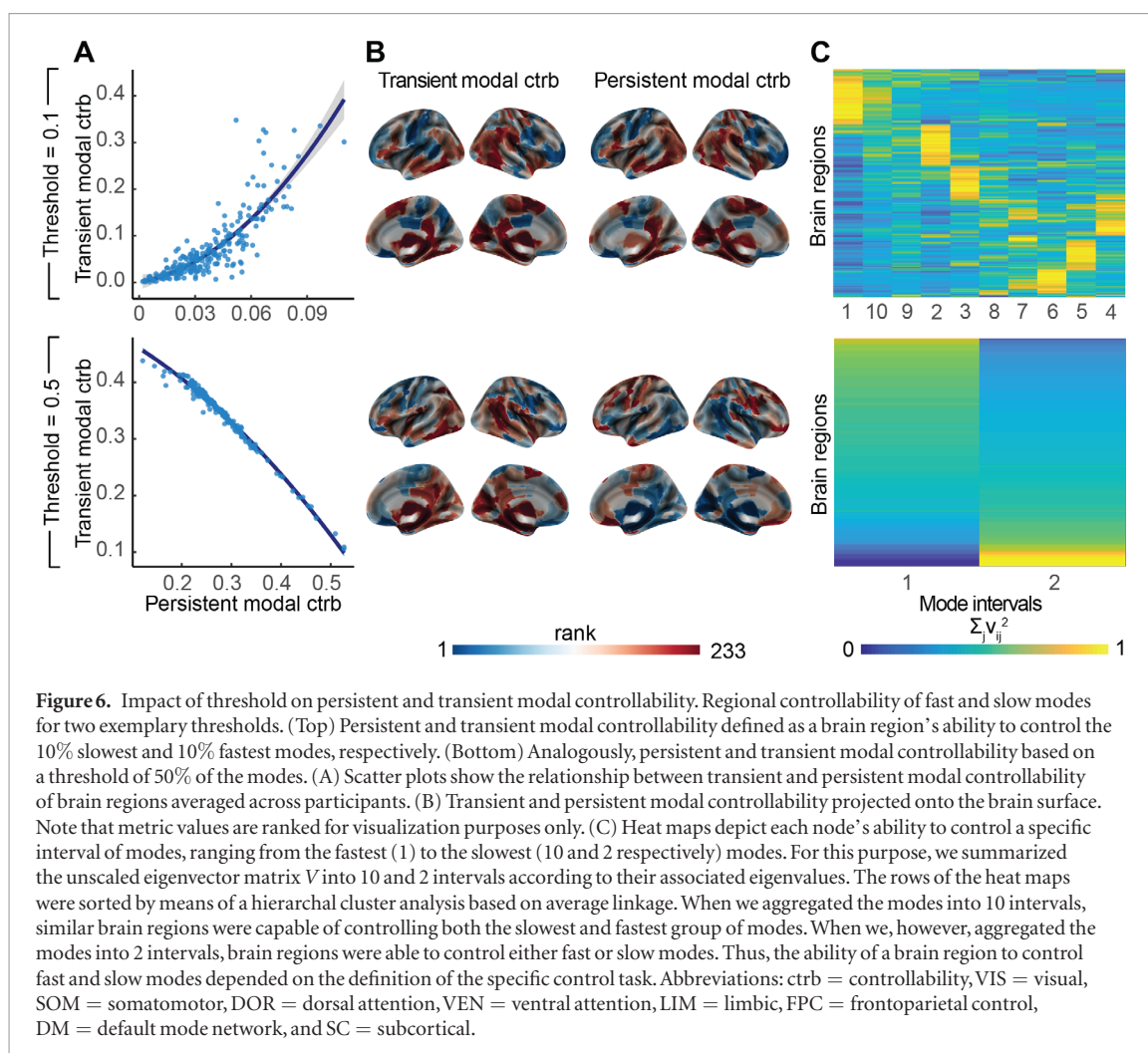
that less connected brain regions exhibit lower average controllability but higher modal controllability [35]. Similarly, individual minimum control energy was increased for binary matrices compared to fully weighted matrices; this result is consistent with previous evidence demonstrating that control nodes with more homogeneous edge weights require larger control energy [29]. Overall, the binarization of the structural connectivity matrix substantially reduced the variance of controllability metrics but not minimum control energy, suggesting that the edge weights carry valuable information especially for controllability metrics.

4.6. Persistent and transient modal controllability

Many neuroscientific endeavors focus on the speed of neural dynamics. Network control theory allows us to explicitly study whether a brain region is capable of controlling fast and slowly changing activity modes by means of transient and persistent modal controllability. However, there is no clear definition of which activity modes are considered as fast or slow. Thus, we wished to further inspect how the definition of fast and slow temporal dynamics affects transient and persistent modal controllability. We began with the calculation of both metrics across various thresholds for determining which modes were considered to be transient versus persistent. First, we observed that with increasing threshold the magnitude

of both transient and persistent modal controllability increased because the number of summed modes was expanded. As expected, we further noted that transient and persistent modal controllability based on a threshold of 0.5 summed up to modal controllability. The initially positive Pearson correlation between transient and persistent modal controllability of brain regions reduced and turned into a negative association with increasing thresholds ($r_{0.1} = 0.82$, $p_{0.1} = 7 \times 10^{-59}$; $r_{0.2} = 0.78$, $p_{0.2} = 7 \times 10^{-49}$; $r_{0.3} = 0.65$, $p_{0.3} = 1 \times 10^{-29}$; $r_{0.4} = -0.20$, $p_{0.4} = 3 \times 10^{-3}$; $r_{0.5} = -0.99$, $p_{0.5} = 4 \times 10^{-187}$) (figure 6(A)). Notably, for small thresholds such as 0.1, a subset of brain regions was found to be capable of controlling both fast and slow temporal dynamics (figure 6(B)). While controlling for the size of each cognitive system, we found that these brain regions belonged primarily to the subcortex (36%) and VIS (22%) systems, but also VEN (12%), DOR (9%), SOM (8%), DM (8%), and FPC (5%) systems. For large thresholds such as 0.5, brain regions seem to be either able to control fast dynamics (39% SC, 14% DOR, 12% VIS, 11% DM, 8% FPC, 6% SOM, 5% VEN, and 5% LIM systems) or slow dynamics (31% FPC, 26% subcortex, 22% SOM, 10% DOR, 7% DM, and 3% VIS systems), but not both.

To explore this ambiguous relationship in more detail, we disentangled the overlapping thresholds by considering the unscaled eigenvector matrix **V**, and then by summarizing the modes into 10 intervals



versus 2 intervals (figure 6(C), top versus bottom). Interestingly, this investigation into the controllability of separate mode intervals also supported the notion that similar brain regions were capable of controlling fast and slow dynamics in the strict definition of these control tasks (10 intervals) but not in the broader definition of these control tasks (2 intervals). Importantly, we note that these results do not extend to discrete-time systems because the definition of modes that are considered as fast versus slow differs substantially between time systems. Overall, the ability of a brain region to control fast and slow modes largely depended on the definition of the control tasks.

4.7. Complexity of energy landscape

Finally, we sought to extend the types of research question we can address with the set of currently available controllability and energy metrics. For this purpose, we developed and validated a complementary metric that measures the heterogeneity of all possible minimum control energy trajectories. The complexity of the energy landscape allows us to quantify the similarity or dissimilarity of all possible state transitions in respect to their required amount of control energy. Based on the variability of the eigenvalues of the controllability Gramian, we

quantified the complexity of the minimum control energy landscape in each individual. Probing the consistency of the complexity of the energy landscape across time systems, we observed a large positive Pearson correlation between discrete- and continuous-time systems ($r = 0.87$, $p = 1 \times 10^{-3}$). We further examined the complementarity of the complexity of the energy landscape by calculating the Pearson correlation between the complexity measure and the other established control metrics defined earlier. We found a small negative association between complexity and average controllability ($r = -0.15$, $p = 0.68$), a large negative association with modal controllability ($r = -0.67$, $p = 0.04$), and a medium negative association with minimum control energy ($r = -0.40$, $p = 0.26$). Next, we validated the complexity of the energy landscape of the brain against three null models, preserving either the strength distribution or the spatial embedding, or both. Brain networks showed a significantly lower complexity of the energy landscape than the topological null model ($W = 65$, $p = 8 \times 10^{-8}$), the spatial null model ($W = 0$, $p = 5 \times 10^{-8}$), and the combined null model ($W = 2498$, $p = 6 \times 10^{-3}$), as quantified by a Wilcoxon test (figure 7). Interestingly, the combination of topological and spatial characteristics seemed to

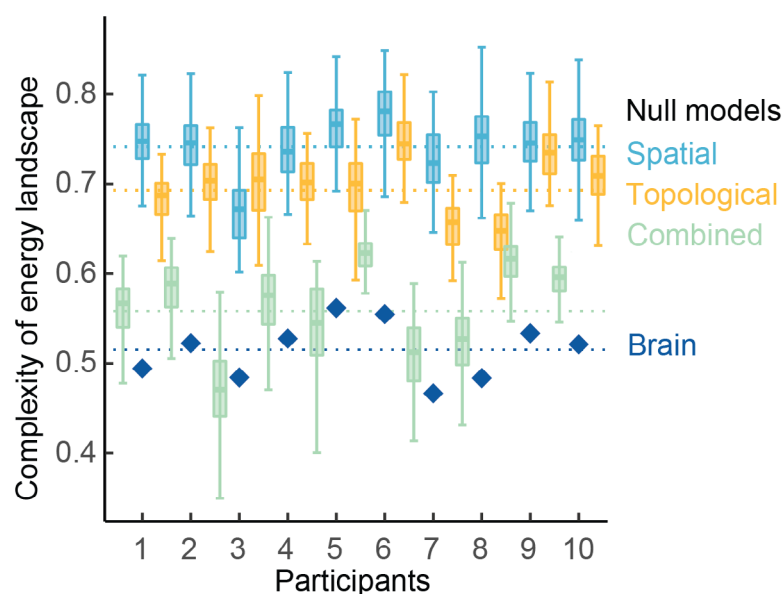


Figure 7. Complexity of the energy landscape of the human brain. Heterogeneity of the minimum control energy landscape of individual participants (dark blue diamonds) as compared to three null models preserving different characteristics of brain networks. The complexity of the energy landscape was quantified by the variability of the eigenvalue distribution of the controllability Gramian. Null model distributions (box plots) were estimated by randomly rewiring each brain network 100 times. Spatial null models (blue box plots) preserved the relationship between edge weight and Euclidean distance. Topological null models (yellow box plots) preserved degree and strength distributions. Combined null models (green box plots) preserved both strength distribution and spatial embedding. Dashed lines indicate complexity of the energy landscape of brain networks and null models averaged across individuals. The combination of topological and spatial characteristics partially explains the homogeneous energy landscape of the brain.

partially explain the brain's higher homogeneity of the energy landscape. We found consistent evidence in discrete-time systems (Sfigure 9; further details are provided in the supplementary results). Overall, the complexity of the energy landscape of the brain was complementary to other controllability metrics and low compared to several null models.

5. Discussion

Network control theory is an emerging field in neuroscience that has the potential to yield promising insights into structure-function relationships in health and disease. Here, we provided an overview of the theoretical framework by illustrating the underlying model of neural dynamics and commonly studied controllability concepts. Based on the structural brain networks from ultra high-resolution diffusion imaging data (730 diffusion directions) of 10 healthy adults, we calculated average and modal controllability as well as minimum and optimal control energy. We then systematically probed the impact of different modeling choices, specifically the choice of time system, time horizon, normalization, and size of the control set, on these metrics. We further suggested potentially useful model extensions such as an alternative measure of structural connectivity accounting for propagation of signals through gray matter to abutting regions, and a complementary metric quantifying the complexity of the energy landscape of brain networks.

5.1. Specific modeling recommendations

Based on theoretic considerations and on our systematic examination of different modeling choices, we derived several specific recommendations. First, we observed a generally high consistency between the behavior of discrete- and continuous-time systems, which depended on the metric, observation level, and time horizon. Classifying the neural dynamics under study as clearly discrete- or continuous-time is often challenging. Unless an investigator has a clear justification for choosing one time system over another, we recommend to verify the obtained results in the alternative time-system to allow for a better generality of the findings and inferences drawn therefrom. Second, we demonstrated that short time horizons led to an alternative time system compared to longer time horizons. Note that too short time horizons could be neurobiologically implausible considering that the transition between brain states is a dynamic process. The arbitrary units of the time scale further challenge the decision of which time horizon to choose. If there exists no concrete justification for the choice of time horizon, we recommend to validate the obtained findings using several different time horizons. Third, we found that a fast stabilization of the system induced a substantially different control scenario, which is in line with theoretic considerations. Again, if there are no concrete neurobiological variables that can be used to constrain one's choice, we suggest that a slow stabilization could be a plausible representation of

most neural dynamics, allowing for a broader range of dynamics. Since the influence of the normalization parameter c depends on the largest eigenvalue, the same c can have different stabilization effects in different brain networks. To ensure consistency across studies, we suggest to make c dependent on the largest positive eigenvalue of the structural connectivity matrix, for instance by $c = 0.01 \cdot \lambda(\mathbf{A})_{\max}$. Finally, the observation that controlling brain dynamics becomes increasingly difficult with reduced control sets is coherent with the complexity of cognitive functions and brain diseases. The amount of required control energy for a specific state transition depended on the size and composition of the control region set. The decision critically depends on the individual research question and hence, should be well informed by theoretical or practical considerations. From a methodological perspective, it is important to control a sufficiently large number of brain regions to robustly estimate control energies. An important next step is the development of tools to determine the most efficient control set for a specific state transition [29]. Lastly, we recommend that future studies assess and report the relation between controllability metrics and weighted degree (Sfigure 10 and 11). Because different network topologies show different relationships between these two metrics, a systematic examination of such relationships in diverse graph ensembles is an important direction for future research. In sum, these recommendations could guide more informed modeling choices in future applications of network control theory to pressing questions in cognitive, developmental, and clinical neuroscience.

5.2. The role of time in network controllability

In our examination of different modeling choices, we found that both a short time horizon and a fast stabilization of the system induced an alternative control regime. We suggest a common mechanism underlying both time-related observations. Whereas the injected control input has time to diffuse along inter-connections between brain regions over longer time horizons, it might be possible that this diffusion process is constrained over short time horizons. Instead, a different control regime could come into effect in which the injected input primarily controls each brain region independently rather than capitalizing on their interconnections. This finding suggests that time might play a more important role in the controllability of structural brain networks than is commonly assumed. Thus, it could be interesting to further investigate the factor of time, for instance by linking control to real-time measures of brain function [17, 57]. Another potentially fruitful venture could be to determine optimal control horizons by capitalizing on the natural dynamics of the system or by changing inter-connections in more advanced dynamic models [75]. Such methods emphasizing the role of time could help to develop minimal clinical interventions such as

neuromodulation [76], which is immediately relevant for the control of seizures in epilepsy [77–81]. The temporal nature of control is also potentially relevant for further refining brain-machine interfaces [82, 83].

5.3. Future directions for proposed model extensions

Moreover, the present work provides several potentially useful extensions of network control theory. We first developed and validated a complementary measure of structural connectivity motivated by the fact that brain networks based on diffusion imaging data disregard the potential for neural signals to diffuse between spatially adjacent brain regions. We demonstrated that this alternative structural connectivity measure based on the amount of shared neighborhood between two brain regions was complementary to the tractography version. We further showed that their combination introduced more inter-individual variability in controllability metrics, motivating future efforts to employ this approach in studies of individual differences. An important next step is to test whether structural brain networks based on both diffusion imaging and spatial adjacency outperform networks purely based on diffusion imaging data by better accounting for the observed neural dynamics [44, 45]. Additionally, we examined the ability of the brain to control slow and fast dynamics. We found that the capability of a brain region to control different fast modes depended on the specific definition of the control task and was not consistent between time-systems. Neuroscientists interested in the speed of neural changes such as different frequency bands [84, 85] should be careful in justifying their choice of time system and the threshold which defines slow versus fast modes.

Lastly, we wished to extend the existing set of controllability metrics. For this purpose, we developed and validated a new metric that quantifies the complexity of the energy landscape of a given brain network. In other words, the metric measures how heterogeneous all possible state transitions are in the control energy that they require. We showed that the brain exhibited a more homogeneous energy landscape compared to two different null models. We found that both the brain networks' strength distribution and spatial embedding partially explained this observation, which is in line with previous findings connecting local and global network characteristics to network controllability [29, 30, 35]. The requirement of a similar amount of energy to enable diverse state transitions implies that brain architecture supports diverse transitions, which in turn could explain the complex functional dynamics consistently observed in neural systems. A crucial next step is to test the practical utility of this new metric by linking it to development, cognition, and psychiatric disorders. Taken together, the proposed model extensions hopefully stimulate and enrich future research. A detailed outlook on further developments in network

control theory is provided in the Supplementary Discussion.

5.4. Methodological considerations

Several methodological aspects could potentially constrain the interpretability of our results. First, we capitalized on high-resolution diffusion weighted imaging data for the construction of structural connectivity networks. Associated tractography algorithms are still limited in their capacity to reliably track fiber bundles, particularly long-range connections [86, 87], in terms of their origin, exact direction, and intersection [88]. Nevertheless, diffusion weighted imaging serves as the state-of-the-art method to study white matter architecture in humans and therefore, tractography algorithms are continuously being refined [89]. In the future, the potential incorporation of directed structural networks enabling more complex brain dynamics may provide additional insight into control strategies utilized by the human brain [90, 91]. Second, our dynamic model of neural processes relied on several simplifying assumptions including linearity and time-invariance. However, such basic models often provide a good starting-point to approximate higher-order dynamics [44, 45] and can subsequently be adapted to contain more complex features such as non-linearity [92, 93] and time-dependence [75]. Third, it was beyond the scope of this paper to examine the impact of modeling choices in all of their theoretically possible combinations. Instead, we systematically varied one modeling choice at a time while keeping all other choices constant. Thus, the obtained results might not automatically generalize to left-out choices, for example in the presence of higher order interactions.

In a similar vein, our findings might not be generalizable to different edge weights due to the substantial impact they can have on controllability metrics [65]. For a more detailed elaboration, we refer the interested reader to a comprehensive study of the behavior of controllability metrics on networks with different edge weighting schemes such as Gaussian, power-law, and nonparametric distributions [65]. Because this practical guide is supposed to primarily target the neuroimaging community, choosing structural brain networks based on diffusion imaging data seemed most useful to us. Nevertheless, the generalizability to other edge weight distributions or structural connectivity measures remains an open question. A further limitation of our study is the investigation of control energies in a restricted set of six state transitions. For didactic purposes, the initial and target brain states were constructed in a controlled, yet unnatural way by capitalizing on the artificial activation of brain regions belonging to the same cognitive system. A thorough investigation of the impact of different choices of brain states would aid future application of network control theory to theoretical neuroscience, but is beyond

the scope of this work. For greater ecological validity, future studies should consider real brain states measured by functional neuroimaging [17, 57] as also provided by meta-analysis of large databases [60, 61]. The present work is furthermore limited to the examination of how the brain transitions between two states as opposed to how the brain maintains a specific state. The systematic study of more complex brain dynamics such as setpoint tracking is an exciting future direction in network control theory [57]. Lastly, we wish to point out that the high consistency between minimum and optimal control energy could also be due to different scales of distance and energy costs. To avoid such effects, future efforts could develop an optimal energy algorithm that balances both constraints equally independent of their scale.

5.5. Conclusions

Our systematic overview of network control theory and possible modeling choices aimed to facilitate a deeper understanding and better evaluation of network control theory applications in neuroscience. Future work can potentially benefit from our specific recommendations and the proposed model extensions. Overall, this work hopefully inspires the neuroscience community to fully exploit the potential of network control theory on multiple spatial scales ranging from single neurons to brain regions. Ultimately, such endeavors could advance our understanding of how the architecture of the brain gives rise to complex neural dynamics.

Acknowledgments

TMK was supported by the International Research Training Group (IRTG 2150) of the German Research Foundation and by a full doctoral scholarship of the German National Academic foundation. We gratefully acknowledge Eli J Cornblath, Xiaosong He, Urs Braun, and Lorenzo Caciagli for useful discussions. DSB acknowledges support from the John D and Catherine T MacArthur Foundation, the Alfred P Sloan Foundation, the ISI Foundation, the Paul Allen Foundation, the Army Research Laboratory (W911NF-10-2-0022), the Army Research Office (Bassett-W911NF-14-1-0679, DCIST-W911NF-17-2-0181, Grafton-W911NF-16-1-0474), the Office of Naval Research, the National Institute of Mental Health (2-R01-DC-009209-11, R01-MH112847, R01-MH107235, R21-MH-106799), the National Institute of Child Health and Human Development (1R01HD086888-01), National Institute of Neurological Disorders and Stroke (R01 NS099348), and the National Science Foundation (BCS-1441502, BCS-1430087, NSF PHY-1554488 and BCS-1631550). 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

DSB, JS and TMK designed the study. AEK preprocessed the data. JZK contributed analytic solutions. TMK and JZK wrote code. TMK analyzed the data and wrote the manuscript. All authors edited the manuscript and approved the final version.

Conflict of interest

The authors declare no competing interests.

ORCID iDs

Teresa M Karrer  <https://orcid.org/0000-0002-7444-7679>

Jennifer Stiso  <https://orcid.org/0000-0002-3295-586X>

Fabio Pasqualetti  <https://orcid.org/0000-0002-8457-8656>

Danielle S Bassett  <https://orcid.org/0000-0002-6183-4493>

References

- [1] Deco G, Jirsa V K and McIntosh A R 2011 Emerging concepts for the dynamical organization of resting-state activity in the brain *Nat. Rev. Neurosci.* **12** 43
- [2] Deco G and Kringelbach M L 2014 Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders *Neuron* **84** 892–905
- [3] Hermundstad A M, Brown K S, Bassett D S and Carlson J M 2011 Learning, memory, and the role of neural network architecture *PLoS Comput. Biol.* **7** e1002063
- [4] Hermundstad A M et al 2013 Structural foundations of resting-state and task-based functional connectivity in the human brain *Proc. Natl Acad. Sci.* **110** 6169–74
- [5] Hermundstad A M, Brown K S, Bassett D S, Aminoff E M, Frithsen A, Johnson A, Tipper C M, Miller M B, Grafton S T and Carlson J M 2014 Structurally-constrained relationships between cognitive states in the human brain *PLoS Comput. Biol.* **10** e1003591
- [6] Rajan K, Harvey C D and Tank D W 2016 Recurrent network models of sequence generation and memory *Neuron* **90** 128–42
- [7] Levy N, Horn D, Meilijson I and Ruppin E 1999 Distributed synchrony in a hebbian cell assembly of spiking neurons *Adv. Neural Inf. Process. Syst.* **14** 815–24
- [8] Fiete I R, Senn W, Wang C Z and Hahnloser R H 2010 Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity *Neuron* **65** 563–76
- [9] Liu Y-Y, Slotine J-J and Barabási A-L 2011 Controllability of complex networks *Nature* **473** 167
- [10] Pasqualetti F, Zampieri S and Bullo F 2014 Controllability metrics, limitations and algorithms for complex networks *IEEE Trans. Control Netw. Syst.* **1** 40–52
- [11] Bullmore E and Sporns O 2009 Complex brain networks: graph theoretical analysis of structural and functional systems *Nat. Rev. Neurosci.* **10** 186
- [12] Bassett D S, Zurn P and Gold J I 2018 On the nature and use of models in network neuroscience *Nat. Rev. Neurosci.* **19** 566–78
- [13] Shenoy K V, Kaufman M T, Sahani M and Churchland M M 2011 A dynamical systems view of motor preparation: implications for neural prosthetic system design *Progress in Brain Research* vol 192 (Amsterdam: Elsevier) pp 33–58
- [14] Freeman W J 1994 Characterization of state transitions in spatially distributed, chaotic, nonlinear, dynamical systems in cerebral cortex *Integr. Physiol. Behav. Sci.* **29** 294–306
- [15] Kalman R E, Ho Y-C and Narendra K S 1963 Controllability of linear dynamical systems *Contributions to Differential Equations* vol 1 pp 189–213
- [16] Muldoon S F, Pasqualetti F, Gu S, Cieslak M, Grafton S T, Vettel J M and Bassett D S 2016 Stimulation-based control of dynamic brain networks *PLoS Comput. Biol.* **12** e1005076
- [17] Stiso J et al 2019 White matter network architecture guides direct electrical stimulation through optimal state transitions *Cell Rep.* **28** 2554–66
- [18] Solomon E A et al 2018 Medial temporal lobe functional connectivity predicts stimulation-induced theta power *Nat. Commun.* **9** 4437
- [19] Medaglia J D, Harvey D Y, White N, Kelkar A, Zimmerman J, Bassett D S and Hamilton R H 2018 Network controllability in the inferior frontal gyrus relates to controlled language variability and susceptibility to TMS *J. Neurosci.* **38** 6399–410
- [20] Khambhati A N et al 2019 Functional control of electrophysiological network architecture using direct neurostimulation in humans *Netw. Neurosci.* **3** 848–77
- [21] Cui Z et al 2018 Optimization of energy state transition trajectory supports the development of executive function during youth *eLife* **9** e53060
- [22] Muldoon S F, Costantini J, Webber W R S, Lesser R and Bassett D S 2018 Locally stable brain states predict suppression of epileptic activity by enhanced cognitive effort *NeuroImage Clin.* **18** 599–607
- [23] Cornblath E J et al 2019 Sex differences in network controllability as a predictor of executive function in youth *NeuroImage* **188** 122–34
- [24] Dum R P, Levinthal D J and Strick P L 2016 Motor, cognitive, and affective areas of the cerebral cortex influence the adrenal medulla *Proc. Natl Acad. Sci. USA* **113** 9922–7
- [25] Goyal M S, Venkatesh S, Milbrandt J, Gordon J I and Raichle M E 2015 Feeding the brain and nurturing the mind: linking nutrition and the gut microbiota to brain development *Proc. Natl Acad. Sci. USA* **112** 14105–12
- [26] Yan G, Vértés P E, Towilson E K, Chew Y L, Walker D S, Schafer W R and Barabási A-L 2017 Network control principles predict neuron function in the caenorhabditis elegans connectome *Nature* **550** 519
- [27] Towilson E K, Vértés P E, Yan G, Chew Y L, Walker D S, Schafer W R and Barabási A 2018 Caenorhabditis elegans and the network control framework-faqs *Phil. Trans. R. Soc. B* **373** 20170372
- [28] Wiles L, Gu S, Pasqualetti F, Parvese B, Gabrieli D, Bassett D S and Meaney D F 2017 Autaptic connections shift network excitability and bursting *Sci. Rep.* **7** 44006
- [29] Kim J Z, Soffer J M, Kahn A E, Vettel J M, Pasqualetti F and Bassett D S 2018 Role of graph architecture in controlling dynamical networks with applications to neural systems *Nat. Phys.* **14** 91
- [30] Gu S et al 2015 Controllability of structural brain networks *Nat. Commun.* **6** 8414
- [31] Shine J M, Breakspear M, Bell P T, Ehgoetz Martens K A, Shine R, Koyejo O, Sporns O and Poldrack R A 2019 Human cognition involves the dynamic integration of neural activity and neuromodulatory systems *Nat. Neurosci.* **22** 289–96
- [32] Jegannathan J, Perry A, Bassett D S, Roberts G, Mitchell P B and Breakspear M 2018 Fronto-limbic dysconnectivity leads to impaired brain network controllability in young people with bipolar disorder and those at high genetic risk *NeuroImage Clin.* **19** 71–81
- [33] Bernhardt B C, Fadaie F, Liu M, Caldaïrou B, Gu S, Jefferies E, Smallwood J, Bassett D S, Bernasconi A and Bernasconi N 2019 Temporal lobe epilepsy: Hippocampal pathology modulates connectome topology and controllability *Neurology* **92** e2209–20
- [34] Gu S, Betzel R F, Mattar M G, Cieslak M, Delio P R, Grafton S T, Pasqualetti F and Bassett D S 2017 Optimal trajectories of brain state transitions *NeuroImage* **148** 305–17

- [35] Betzel R F, Gu S, Medaglia J D, Pasqualetti F and Bassett D S 2016 Optimally controlling the human connectome: the role of network topology *Sci. Rep.* **6** 30770
- [36] Bassett D S and Bullmore E T 2009 Human brain networks in health and disease *Curr. Opin. Neurol.* **22** 340
- [37] Medaglia J D, Lynall M-E and Bassett D S 2015 Cognitive network neuroscience *J. Cogn. Neurosci.* **27** 1471–91
- [38] Van den Heuvel M P, Bullmore E T and Sporns O 2016 Comparative connectomics *Trends Cogn. Sci.* **20** 345–61
- [39] Braun U, Schaefer A, Betzel R F, Tost H, Meyer-Lindenberg A and Bassett D S 2018 From maps to multi-dimensional network mechanisms of mental disorders *Neuron* **97** 14–31
- [40] Tang E et al 2017 Developmental increases in white matter network controllability support a growing diversity of brain dynamics *Nat. Commun.* **8** 1252
- [41] Tang E and Bassett D S 2018 Colloquium: control of dynamics in brain networks *Rev. Mod. Phys.* **90** 031003
- [42] Kim J Z and Bassett D S 2019 Linear dynamics & control of brain networks (arXiv:1902.03309)
- [43] Kailath T 1980 *Linear Systems* vol 156 (Englewood Cliffs, NJ: Prentice-Hall)
- [44] Galán R F 2008 On how network architecture determines the dominant patterns of spontaneous neural activity *PLoS One* **3** e2148
- [45] Honey C, Sporns O, Cammoun L, Gigandet X, Thiran J-P, Meuli R and Hagmann P 2009 Predicting human resting-state functional connectivity from structural connectivity *Proc. Natl Acad. Sci.* **106** 2035–40
- [46] May A 2011 Experience-dependent structural plasticity in the adult human brain *Trends Cogn. Sci.* **15** 475–82
- [47] Baum G L et al 2017 Modular segregation of structural brain networks supports the development of executive function in youth *Curr. Biol.* **27** 1561–72.e8
- [48] Brinkman B A, Weber A I, Rieke F and Shea-Brown E 2016 How do efficient coding strategies depend on origins of noise in neural circuits? *PLoS Comput. Biol.* **12** e1005150
- [49] Mlynarski W F and Hermundstad A M 2018 Adaptive coding for dynamic sensory inference *eLife* **7** e32055
- [50] Gai Y, Doiron B and Rinzel J 2010 Slope-based stochastic resonance: how noise enables phasic neurons to encode slow signals *PLoS Comput. Biol.* **6** e1000825
- [51] Garrett D D, McIntosh A R and Grady C L 2014 Brain signal variability is parametrically modifiable *Cereb. Cortex* **24** 2931–40
- [52] Breakspear M and McIntosh A R 2011 Networks, noise and models: reconceptualizing the brain as a complex, distributed system *NeuroImage* **58** 293–5
- [53] Gollo L L, Mirasso C, Sporns O and Breakspear M 2014 Mechanisms of zero-lag synchronization in cortical motifs *PLoS Comput. Biol.* **10** e1003548
- [54] Taylor P N, Wang Y, Goodfellow M, Dauwels J, Moeller F, Stephani U and Baier G 2014 A computational study of stimulus driven epileptic seizure abatement *PLoS One* **9** e114316
- [55] Boltyanskii V G, Gamkrelidze R V and Pontryagin L S 1960 The maximum principle in the theory of optimal processes of control *IFAC Proc. Vol.* **1** 464–9
- [56] Hespanha J P 2018 *Linear Systems Theory* (Princeton, NJ: Princeton University Press)
- [57] Cornblath E J et al 2018 Context-dependent architecture of brain state dynamics is explained by white matter connectivity and theories of network control (arXiv:1809.02849)
- [58] Zöller D, Sandini C, Schaer M, Eliez S, Bassett D S and Van De Ville D 2019 Structural control energy of resting-state functional brain states reveals inefficient brain dynamics in psychosis vulnerability Submitted (<https://doi.org/10.1101/703561>)
- [59] Braun U et al 2019 Brain state stability during working memory is explained by network control theory, modulated by dopamine d1/d2 receptor function, and diminished in schizophrenia Epub ahead of print
- [60] Yarkoni T, Poldrack R A, Nichols T E, Van Essen D C and Wager T D 2011 Large-scale automated synthesis of human functional neuroimaging data *Nat. Methods* **8** 665
- [61] Fox P T and Lancaster J L 2002 Mapping context and content: the brainmap model *Nat. Rev. Neurosci.* **3** 319
- [62] Menara T, Bassett D and Pasqualetti F 2018 Structural controllability of symmetric networks *IEEE Trans. Autom. Control* **64** 3740–7
- [63] Hamdan A and Nayfeh A 1989 Measures of modal controllability and observability for first- and second-order linear systems *J. Guid. Control Dyn.* **12** 421–8
- [64] Murphy A C, Bertolero M A, Papadopoulos L, Lydon-Staley D M and Bassett D S 2019 Multiscale and multimodal network dynamics underpinning working memory (arXiv:1901.06552)
- [65] Wu-Yan E, Betzel R F, Tang E, Gu S, Pasqualetti F and Bassett D S 2018 Benchmarking measures of network controllability on canonical graph models *J. Nonlinear Sci.* pp 1–39
- [66] Fortunato S and Hric D 2016 Community detection in networks: a user guide *Phys. Rep.* **659** 1–44
- [67] Cammoun L, Gigandet X, Meskaldji D, Thiran J P, Sporns O, Do K Q, Maeder P, Meuli R and Hagmann P 2012 Mapping the human connectome at multiple scales with diffusion spectrum mri *J. Neurosci. Methods* **203** 386–97
- [68] Daducci A, Gerhard S, Griffa A, Lemkaddem A, Cammoun L, Gigandet X, Meuli R, Hagmann P and Thiran J-P 2012 The connectome mapper: an open-source processing pipeline to map connectomes with mri *PLoS One* **7** e48121
- [69] Yeo B T et al 2011 The organization of the human cerebral cortex estimated by intrinsic functional connectivity *J. Neurophysiol.* **106** 1125–65
- [70] Eisenreich B R, Akaishi R and Hayden B Y 2017 Control without controllers: toward a distributed neuroscience of executive control *J. Cogn. Neurosci.* **29** 1684–98
- [71] Tang E, Baum G L, Roalf D R, Satterthwaite T D, Pasqualetti F and Bassett D S 2019 The control of brain network dynamics across diverse scales of space and time (arXiv:1901.07536)
- [72] Rubinov M and Sporns O 2010 Complex network measures of brain connectivity: uses and interpretations *NeuroImage* **52** 1059–69
- [73] Roberts J A, Perry A, Lord A R, Roberts G, Mitchell P B, Smith R E, Calamante F and Breakspear M 2016 The contribution of geometry to the human connectome *NeuroImage* **124** 379–93
- [74] Sun J and Motter A E 2013 Controllability transition and nonlocality in network control *Phys. Rev. Lett.* **110** 208701
- [75] Li A, Cornelius S P, Liu Y-Y, Wang L and Barabási A-L 2017 The fundamental advantages of temporal networks *Science* **358** 1042–6
- [76] Afshar P, Wei X, Lazarewicz M, Gupta R, Molnar G and Denison T 2011 Advancing neuromodulation using a dynamic control framework *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2011** 671–4
- [77] Ching S, Brown E N and Kramer M A 2012 Distributed control in a mean-field cortical network model: implications for seizure suppression *Phys. Rev. E* **86** 021920
- [78] Stanslaski S, Giftakis J, Stypulkowski P, Carlson D, Afshar P, Cong P and Denison T 2011 Emerging technology for advancing the treatment of epilepsy using a dynamic control framework *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2011** 753–6
- [79] Taylor P N, Thomas J, Sinha N, Dauwels J, Kaiser M, Thesen T and Ruths J 2015 Optimal control based seizure abatement using patient derived connectivity *Frontiers Neurosci.* **9** 202
- [80] Olmi S, Petkoski S, Guye M, Bartolomei F and Jirsa V 2019 Controlling seizure propagation in large-scale brain networks *PLoS Comput. Biol.* **15** e1006805
- [81] Ehrens D, Sritharan D and Sarma S V 2015 Closed-loop control of a fragile network: application to seizure-like dynamics of an epilepsy model *Frontiers Neurosci.* **9** 58
- [82] Gowda S, Orsborn A L and Carmenta J M 2012 Parameter estimation for maximizing controllability of linear brain-machine interfaces *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2012** 1314–7
- [83] Stiso J, Corsi M-C, Garcia J O, Vettel J M, De Vico Fallani F, Lucas T H and Bassett D S 2019 Learning in brain-computer interface control evidenced by joint decomposition of brain and behavior (psyArxiv:en69j)

- [84] Siegel M, Donner T H and Engel A K 2012 Spectral fingerprints of large-scale neuronal interactions *Nat. Rev. Neurosci.* **13** 121
- [85] Baillet S 2017 Magnetoencephalography for brain electrophysiology and imaging *Nat. Neurosci.* **20** 327
- [86] Thomas C, Frank Q Y, Irfanoglu M O, Modi P, Saleem K S, Leopold D A and Pierpaoli C 2014 Anatomical accuracy of brain connections derived from diffusion mri tractography is inherently limited *Proc. Natl Acad. Sci.* **111** 16574–9
- [87] Reveley C, Seth A K, Pierpaoli C, Silva A C, Yu D, Saunders R C, Leopold D A and Frank Q Y 2015 Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography *Proc. Natl Acad. Sci.* **112** E2820–8
- [88] Jbabdi S and Johansen-Berg H 2011 Tractography: Where do we go from here? *Brain Connectivity* **1** 169–83
- [89] Pestilli F, Yeatman J D, Rokem A, Kay K N and Wandell B A 2014 Evaluation and statistical inference for human connectomes *Nat. Methods* **11** 1058
- [90] Pasqualetti F and Zampieri S 2014 On the controllability of isotropic and anisotropic networks *53rd IEEE Conf. on Decision and Control (IEEE)* pp 607–12
- [91] Felleman D J and Van Essen D 1991 Distributed hierarchical processing in the primate cerebral cortex *Cerebral Cortex (New York, NY: 1991)* **1** 1–47
- [92] Fiedler B, Mochizuki A, Kurosawa G and Saito D 2013 Dynamics and control at feedback vertex sets. I: informative and determining nodes in regulatory networks *J. Dyn. Differ. Equ.* **25** 563–604
- [93] Zañudo J G T, Yang G and Albert R 2017 Structure-based control of complex networks with nonlinear dynamics *Proc. Natl Acad. Sci.* **114** 7234–9