

**Applications of optimal nonlinear control to a whole-brain network of FitzHugh-Nagumo oscillators**Teresa Chouzouris,<sup>1,\*</sup> Nicolas Roth<sup>1,\*</sup>,<sup>†</sup> Caglar Cakan,<sup>1,2</sup> and Klaus Obermayer<sup>1,2</sup><sup>1</sup>*Institut für Softwaretechnik und Theoretische Informatik, Technische Universität Berlin, Marchstraße 23, 10587 Berlin, Germany*<sup>2</sup>*Bernstein Center for Computational Neuroscience Berlin, Philippstraße 13, 10115 Berlin, Germany*

(Received 5 March 2021; accepted 26 July 2021; published 19 August 2021)

We apply the framework of optimal nonlinear control to steer the dynamics of a whole-brain network of FitzHugh-Nagumo oscillators. Its nodes correspond to the cortical areas of an atlas-based segmentation of the human cerebral cortex, and the internode coupling strengths are derived from diffusion tensor imaging data of the connectome of the human brain. Nodes are coupled using an additive scheme without delays and are driven by background inputs with fixed mean and additive Gaussian noise. Optimal control inputs to nodes are determined by minimizing a cost functional that penalizes the deviations from a desired network dynamic, the control energy, and spatially nonsparsely control inputs. Using the strength of the background input and the overall coupling strength as order parameters, the network's state-space decomposes into regions of low- and high-activity fixed points separated by a high-amplitude limit cycle, all of which qualitatively correspond to the states of an isolated network node. Along the borders, however, additional limit cycles, asynchronous states, and multistability can be observed. Optimal control is applied to several state-switching and network synchronization tasks, and the results are compared to controllability measures from linear control theory for the same connectome. We find that intuitions from the latter about the roles of nodes in steering the network dynamics, which are solely based on connectome features, do not generally carry over to nonlinear systems, as had been previously implied. Instead, the role of nodes under optimal nonlinear control critically depends on the specified task and the system's location in state space. Our results shed new light on the controllability of brain network states and may serve as an inspiration for the design of new paradigms for noninvasive brain stimulation.

DOI: [10.1103/PhysRevE.104.024213](https://doi.org/10.1103/PhysRevE.104.024213)**I. INTRODUCTION**

The widespread use of noninvasive electrical brain stimulation in clinical applications has sparked ongoing interest in studying the effects of external inputs on brain activity. Stimulation with electric fields in the range of 1–2 V/m can already modulate oscillatory brain activity [1,2], affect cross-regional synchronization [3,4], and modulate cognitive performance [5]. Clinical studies have successfully demonstrated the efficacy of targeted transcranial stimulation in the treatment of neurological and psychiatric disorders and diseases such as epilepsy [6,7], schizophrenia [8,9], Alzheimer's disease [10], and depression [11].

Brain network models offer a way to simulate and understand the human brain as a nonlinear dynamical system, in which each brain region is represented by a node, and the node dynamics is defined by a model of the average neural activity in that region [12]. Nodes interact with each other according to empirically measured human structural neural connectivity, which quantifies how neural activity in one brain region is coupled to the activity of connected regions.

Parcellation of human brains has yielded various brain atlases [13], which provide information on spatial and functional segregation, dividing the brain into distinct areas

[14,15]. The relative connectivity strength between these areas is defined by the number of axonal fibers, which can be estimated using structural neuroimaging scans of individual subjects. This results in a network model of the human brain where the edges reflect the relative strengths of axonal fiber bundles and the nodes represent individual brain regions. These nodes are then equipped with a dynamical system modeling the average neuronal activity in that region [16]. It has been repeatedly shown that when the parameters of brain network models are fitted to functional brain data, optimal operating points were close to the bifurcation lines of these models. This ensures that the model is in a state in which noise fluctuations can be amplified and produce realistic spatial correlation structures which are similar to empirical functional connectivity measurements [17–20].

Previous theoretical investigations into the impact of external perturbations mostly relied on the assumption of a linear node dynamics, allowing the application of methods from linear control theory [21]. Thus, one can draw conclusions on the effects of external inputs to the system, based on the network topology and independent of its dynamical state [22,23]. However, linear node dynamics cannot reproduce the dynamics of neural processes close to bifurcations [17,24]. To describe the transitions from one dynamical regime to another, Muldoon *et al.* [25] consider nonlinear node dynamics. They conclude that the effects of stimulation-based control can be predicted by diagnostics from linear control theory based only on the structural connectivity of the network.

\*The first two authors contributed equally to this work.

<sup>†</sup>roth@tu-berlin.de

In this work we go beyond linear control theory and explore the framework of optimal nonlinear control [26] for the assessing the impact of perturbations on networks of coupled nonlinear systems. Optimal control is an optimization method that derives control policies based on the minimization of a cost functional which depends on the state and control variables. Here, we define a cost functional that penalizes the deviation to the desired network dynamics, the control energy, and the spatial sparsity. The latter allows us to find optimal control signals that apply to only a few control sites, as introduced by Ref. [27] and further studied by Ref. [28]. We used methods presented by Tröltzsch in Ref. [29] for their calculation, while the handling of the stochastic term is based on the work of Stannat *et al.* [30]. The resulting mathematical formulation is analogous to the formulation presented by Casas *et al.* [31] for partial differential equations.

We evaluate the framework of optimal control for a brain network of FitzHugh-Nagumo (FHN) [32,33] oscillators with additive coupling and white Gaussian noise, where each oscillator represents a brain region and where the connections between them were chosen according to the human connectome derived from diffusion tensor imaging (DTI) measurements. FHN oscillators are well studied models for neural dynamics and detailed analyses of the dynamical states exist for single oscillators [34] and various network configurations, like two coupled units [35,36], lattices [37], rings and hierarchical architectures [38], as well as random and small-world topologies [39,40], and multiplex networks [41–43]. Similar empirical DTI-measured brain connectivities were used with FHN dynamics to model highly synchronized epileptic-seizure-like states [44,45], unihemispheric sleep [46], and the functional organization of the resting brain [47–49].

We consider two different classes of control problems, targeted attractor switching between multistable network states and increasing network synchronization. The solutions obtained for given energy, precision, and sparseness constraints are well interpretable and result in intuitively sensible optimal control inputs for all network nodes over time.

We then correlate the control energies to controllability measures from linear control theory. We confirm the findings of Muldoon *et al.* [25] for the investigated state transition (from the low fixed-point state to the oscillatory regime). For other control tasks, however, we show, that diagnostics from linear control theory do not correlate with results from optimal nonlinear control. Conclusions drawn from the structural connectivity alone lead to contradictions, which can only be resolved if the dynamical state of the network and the nonlinear interactions between nodes are taken into account. Applications of nonlinear control theory to whole-brain models enables us to investigate control mechanisms also close to bifurcations. It thus may help in the search for more effective paradigms for realistic transcranial brain stimulation protocols.

## II. NONLINEAR OPTIMAL CONTROL

### A. Network model and control inputs

We consider networks of  $N$  equivalent  $d$ -dimensional and “noisy” nodes with additive and zero-delay internode cou-

pling. Internode coupling strength is described by a  $N \times N$ -dimensional adjacency matrix  $\mathbf{A}$  which is scaled by a global coupling strength  $\sigma$ . We allow for an instantaneous and additive control input to the network, which is described by  $N \times d$  independent control variables  $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_N)$  with  $\mathbf{u}_i = (u_{i1}, \dots, u_{id})$ . The equations describing the controlled network dynamics thus read

$$\begin{aligned} \frac{d}{dt}\mathbf{x}(t) = & \mathbf{h}[\mathbf{x}(t)] + \sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x}(t) \\ & + (\mathbf{B} \otimes \mathbf{K})\mathbf{u}(t) + (\mathbf{I}_N \otimes \mathbf{D})\boldsymbol{\xi}(t), \end{aligned} \quad (1)$$

where  $\otimes$  denotes the Kronecker product. The state of the network is described by a vector  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ , where  $\mathbf{x}_i = (x_{i1}, \dots, x_{id})$  characterize the individual states of the nodes. The local node dynamics is given by  $\mathbf{h}(\mathbf{x}) = (\mathbf{h}(\mathbf{x}_1), \dots, \mathbf{h}(\mathbf{x}_N))$  with  $\mathbf{h}(\mathbf{x}_i) = [h_1(\mathbf{x}_i), \dots, h_d(\mathbf{x}_i)]$ . All nodes additionally receive Gaussian white noise of similar intensity  $\eta$ , which may be correlated within but is uncorrelated across the nodes of the network. The stochastic variables  $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N)$  with  $\boldsymbol{\xi}_i = (\xi_{i1}, \dots, \xi_{id})$  are independently drawn from a Gaussian distribution,  $\xi_{ij} \in \mathcal{N}(0, 1)$ , with zero mean and unit variance. Within node correlations are quantified by the  $d \times d$  dimensional local noise scheme  $\mathbf{D}$ , while the across-node statistical independence is assured by a  $N$ -dimensional identity matrix  $\mathbf{I}_N$ . The internode coupling term consists of the Kronecker product of the adjacency matrix  $\mathbf{A}$  and the  $d \times d$  dimensional local coupling scheme  $\mathbf{G}$ , where the purpose of the former is to describe the relative interaction strength between nodes while the purpose of the latter is to distribute the between-node interactions across the  $d$  variables which describe the local node dynamics. The  $N \times N$  matrix  $\mathbf{B}$  in the control term allows for the control of multiple nodes with different strength. The  $d \times d$  dimensional matrix  $\mathbf{K}$  implements the local control scheme, which is similar for every node in the network. Initial conditions are denoted by  $\mathbf{x}(t = 0) = \mathbf{x}_0$ .

### B. The cost functional and the optimality condition

The control  $\mathbf{u}$  is considered to be optimal ( $\mathbf{u} = \bar{\mathbf{u}}$ ), if it minimizes an appropriate cost functional. To this end, we construct a cost functional  $F[\mathbf{x}(\mathbf{u}), \mathbf{u}]$  for a state switching task, where the control input drives the network model from one stable state to another (see Sec. IV A), and for a node synchronization task, where the control input increases the degree of synchronization among its nodes (see Sec. IV B). For finite noise, i.e., for finite values of  $\eta$  in Eq. (1), the overall cost functional  $F$  is defined as a mean over noise realizations [30],

$$F[\mathbf{x}(\mathbf{u}), \mathbf{u}] = \langle F_n[\mathbf{x}(\mathbf{u}), \mathbf{u}] \rangle = \langle F_n^x(\mathbf{x}) \rangle + F^u(\mathbf{u}), \quad (2)$$

where  $F_n$  denotes the cost functional for one noise realization  $n$  and the angle brackets  $\langle \cdot \rangle$  denote the mean. In the noise-free case no averaging is performed. The state dependent term  $F_n^x(\mathbf{x})$  only implicitly depends on the control and penalizes the deviation from the desired output. It will be different for the switching,  $F_{n,sw}^x$ , and synchronization,  $F_{n,syn}^x$ , tasks (see below). The control dependent term  $F^u(\mathbf{u})$  accounts for the cost of the control itself.

For the state switching task, we consider a noise-free system ( $\eta = 0$ ), and the state dependent cost functional  $F_{n,sw}^x$  is

defined in terms of the deviation of the controlled state to a predefined target state  $\mathbf{x}_T(t)$ :

$$\begin{aligned} F_{n,\text{sw}}^x(\mathbf{x}, t) &= F_{\text{sw}}^x(\mathbf{x}, t) \\ &= \int_0^T \frac{I_p(t)}{2} [\mathbf{x}(t) - \mathbf{x}_T(t)]^2 dt, \end{aligned} \quad (3)$$

where we consider the control being active within the time period  $0 \leq t \leq T$ . To penalize the precision only toward the end of the controlled time interval rather than during the transition between the initial and target states, the precision-penalizing variable  $I_p$  can be time-dependent (see Sec. IV A).

The state dependent cost functional for the synchronization task  $F_{n,\text{syn}}^x$  is defined in terms of the deviations of the normalized pairwise cross-correlations  $R_{ij}$  for all nodes  $i$  and  $j$  from  $R_T$ , the desired mean cross-correlation in the synchronized target state:

$$F_{n,\text{syn}}^x(\mathbf{x}, t) = \frac{I_p}{4N^2} \sum_{i,j=1}^N (R_{ij} - R_T)^2. \quad (4)$$

The cross-correlation in component  $m$  is defined as

$$R_{ij}^m = \int_0^T \frac{[x_{im}(t) - \langle x_{im} \rangle_t][x_{jm}(t) - \langle x_{jm} \rangle_t]}{\sigma_{\langle x_{im} \rangle_t} \sigma_{\langle x_{jm} \rangle_t}} dt, \quad (5)$$

with  $i, j \in [0, \dots, N]$  and where  $\langle \cdot \rangle_t$  and  $\sigma_{\langle \cdot \rangle_t}$  denotes the temporal mean and the standard deviation. The mean  $R^m$  over all values  $R_{ij}^m$ ,

$$R^m = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N R_{ij}^m, \quad (6)$$

is the componentwise network cross-correlation. Later we specialize to  $m = 1$  and suppress the index  $m$ .

The input dependent cost functional  $F^u(\mathbf{u})$  penalizes the energy of the control signal and enforces its directional sparsity [27,28]. It is given by

$$F^u(\mathbf{u}) = \frac{I_e}{2} \int_0^T \mathbf{u}^2(t) dt + I_s \sum_{k=1}^N \left[ \int_0^T u_k^2(t) dt \right]^{\frac{1}{2}}, \quad (7)$$

where  $I_e$  and  $I_s$  are the weights for the energy and sparsity terms. The first term corresponds to the L2-regulation and the second term to the L1-regulation of the cost-functional. Typically, increasing the sparsity  $I_s$  reduces the number of controlled nodes, i.e., the nodes for which the control input is nonzero for at least part of the time period  $0 \leq t \leq T$ , while a higher penalty on the energy term typically leads to an overall reduction of control strengths.

Our goal is to find the optimal control  $\bar{\mathbf{u}}$  that minimizes the cost functional for chosen  $I_p$ ,  $I_e$ , and  $I_s$ , leading to the minimization problem

$$\bar{\mathbf{u}} = \arg \min_{\mathbf{u}} (F_n[\mathbf{x}(\mathbf{u}), \mathbf{u}]). \quad (8)$$

Similar to Troeltzsch *et al.* and Casas *et al.* [29,31,50], we analyze the gradient  $g = \nabla_{\mathbf{u}} F$  of the cost functional, which has to vanish when evaluated at the optimal control  $\bar{\mathbf{u}}$ . By applying the method of Lagrange multipliers, we obtain an expression for the optimality condition that depends on the

adjoint states  $\boldsymbol{\phi}(\mathbf{x}, \mathbf{u}, t)$  corresponding to the Lagrange multipliers. The control  $\bar{\mathbf{u}}$  is optimal, if

$$\mathbf{g}(t) = (\mathbf{B} \otimes \mathbf{K})^T \langle \boldsymbol{\phi}(\mathbf{x}, \bar{\mathbf{u}}, t) \rangle + I_e \bar{\mathbf{u}}(t) + I_s \bar{\lambda}(t) \stackrel{!}{=} \mathbf{0} \quad (9)$$

for  $0 \leq t \leq T$ , where  $\bar{\lambda}(t)$  is the derivative of the sparsity term, Eq. (7), with respect to the control inputs  $\mathbf{u}$ . The adjoint states are governed by a set of linear differential equations:

$$-\frac{d}{dt} \boldsymbol{\phi}(t) = [D_x(\mathbf{h}) + \sigma(\mathbf{A} \otimes \mathbf{G})]^T \boldsymbol{\phi}(t) + \nabla_{\mathbf{x}} f_n^x(\mathbf{x}, t), \quad (10)$$

where  $D_x(\mathbf{h})$  is the Jacobian matrix of the state equations of the dynamical system, and  $f_n^x(\mathbf{x})$  is the integrand of the cost functional,  $F_n^x(\mathbf{x}) = \int_0^T f_n^x(\mathbf{x}) dt$ . The adjoint state satisfies the boundary condition  $\boldsymbol{\phi}(T) = \mathbf{0}$ , and the differential equation is solved backwards in time.

Following Ref. [28],  $\bar{\lambda}(t)$  is given by

$$\bar{\lambda}_k(t) = \begin{cases} \frac{\bar{u}_k(t)}{\sqrt{E_k}} & \text{if } E_k \neq 0, \\ -\frac{1}{I_s} [(\mathbf{B} \otimes \mathbf{K})^T \langle \boldsymbol{\phi}(\mathbf{x}, \bar{\mathbf{u}}, t) \rangle]_k & \text{otherwise.} \end{cases} \quad (11)$$

Here,  $k \in [1, N]$ , and  $E_k$  is the nodewise control energy of the resulting optimal control  $\bar{\mathbf{u}}$ , which is defined as

$$E_k = \int_0^T \bar{u}_k^2(t) dt, \quad (12)$$

resulting in the total control energy  $E = \sum_k E_k$ . A detailed derivation of Eqs. (9) and (10) is provided in Appendix A.

### C. Numerical solution of the minimization problem

The optimization problem is numerically solved using the conjugate gradient method. We integrate the equations for the network state and the adjoint given in Eqs. (1) and (10) (see Appendix B for details). The direction  $\mathbf{d}_l$  for each step of the conjugate gradient algorithm is defined by the Polak-Ribiere method [51], while its step size  $s_l$  is derived using simple bisection [52]. We apply the Fletcher and Reeves algorithm [53] as presented in the following.

We initialize at iteration  $l = 0$  by choosing an initial control  $\mathbf{u}_0$  and drawing 20 noise realizations  $\xi_n(t)$ . The corresponding states  $\mathbf{x}_0(t)$ , Eq. (1), and adjoint states  $\boldsymbol{\phi}_0(t)$ , Eq. (10), are calculated for every noise realization. Then  $\mathbf{g}(t)$  as given in Eq. (9) is evaluated. The descent direction is initialized with  $\mathbf{d}_0(t) = -\mathbf{g}(t)$ . We then iterate until convergence:

While  $\|\mathbf{g}(t)\|_{\infty} > \epsilon$ :

- (1) compute the step size  $s_l$  using bisection
- (2) set  $\mathbf{u}_{l+1}(t) = \mathbf{u}_l(t) + s_l \mathbf{d}_l(t)$
- (3) calculate the states  $\mathbf{x}_{l+1}(t)$  and the adjoint states  $\boldsymbol{\phi}_{l+1}(t)$  for every noise realization and compute the mean  $\langle \boldsymbol{\phi}_{l+1}(t) \rangle$
- (4) evaluate the gradient  $\mathbf{g}_{l+1}(t)$
- (5) compute  $\beta_l$  using the Polak-Ribiere method:  $\beta_l = \frac{\mathbf{g}_{l+1}(\mathbf{g}_{l+1} - \mathbf{g}_l)}{\|\mathbf{g}_{l+1}\|^2}$
- (6) set the direction  $\mathbf{d}_{l+1}(t) = -\mathbf{g}_{l+1}(t) + \beta_l \mathbf{d}_l(t)$
- (7) if  $\mathbf{d}_{l+1}(t)$  is not a descent direction, set  $\mathbf{d}_{l+1}(t) = -\mathbf{g}_{l+1}(t)$
- (8) set  $l = l + 1$

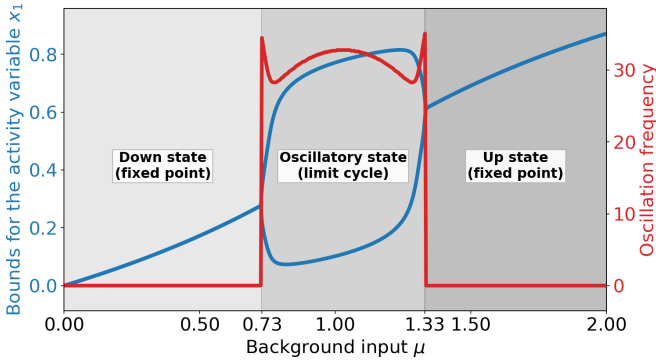


FIG. 1. Bifurcation diagram of an uncoupled FitzHugh-Nagumo oscillator, as given in Eq. (13), with the parameters  $\alpha = 3$ ,  $\beta = 4$ ,  $\gamma = 1.5$ ,  $\delta = 0.5$ , and  $\tau = 20$ . The blue line shows the minimal and maximal values of the activity variable  $x_1$  (identical in fixed points) for the respective background input  $\mu$  to the node. The red line shows the frequency of the oscillation (time and therefore also frequency are measured in arbitrary units).

Gradient-based optimization is only guaranteed to converge to local optima of the cost function. In general, however, multiple initial conditions  $\mathbf{u}_0$  converged to the same optimum. The only exception occurred for the state-switching task for high values of the sparseness parameter  $I_s$ , for which a solution with finite values of  $\mathbf{u}$  coexists with the solution  $\mathbf{u}(t) = 0$  (see Appendix B for details).

### III. THE WHOLE-BRAIN NETWORK

#### A. The local node dynamics

We consider a single FitzHugh-Nagumo (FHN) oscillator, with the activity variable  $x_1$  and the linear recovery variable  $x_2$ :

$$\mathbf{h}(\mathbf{x}) = \begin{pmatrix} h_1(\mathbf{x}) \\ h_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \frac{d}{dt}x_1 \\ \frac{d}{dt}x_2 \end{pmatrix} = \begin{pmatrix} R(x_1) - x_2 + \mu \\ \frac{1}{\tau}(x_1 - \delta x_2) \end{pmatrix}, \quad (13)$$

where  $\mu$  is a node-independent, constant background input and  $R(x) = -\alpha x^3 + \beta x^2 - \gamma x$ . The parameters in  $R$  are chosen to obtain the bifurcation diagram shown in Fig. 1. This bifurcation diagram, depending on the background input  $\mu$ , shows three distinct states. In the down state, the node is in a stable fixed point and has a low constant value of the activity variable  $x_1$  (in blue). In the oscillatory state, the activity variable  $x_1$  oscillates at an input-dependent frequency (in red). In the up state, the node is again in a stable fixed point with a high constant value of the activity variable  $x_1$ .

The succession of these states in the single node dynamics—a supercritical Andronov-Hopf bifurcation from the down state to the oscillatory state and another supercritical Andronov-Hopf bifurcation from the oscillatory state to the up state [34,54]—closely resembles the states found in large random networks of excitatory and inhibitory spiking neuron models and their corresponding mean-field description [55]. The mean firing rate of these models changes as a function of background input in a way which is qualitatively similar to the value of the activity variable shown in Fig. 1. For this reason one can interpret the value of the activity variable  $x_1$  as the

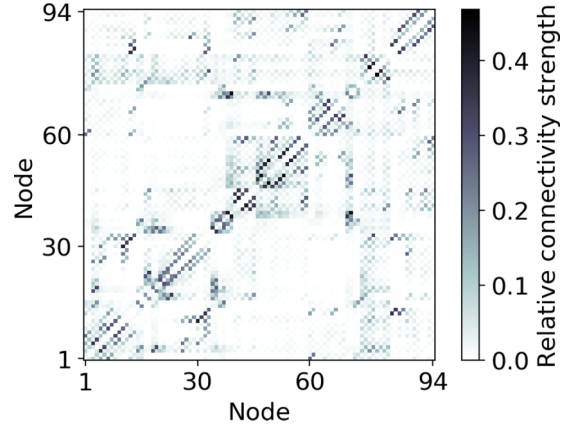


FIG. 2. Weighted adjacency matrix  $\mathbf{A}$  describing the topology of the whole-brain network. Color denotes the relative connection strength of the structural connectivity for every pair of the  $N = 94$  nodes.

difference between the output firing rate of a cortical node to a baseline value.

#### B. The brain network model

The topology of the whole-brain network model is derived from diffusion tensor imaging (DTI) data of 12 human subjects (all male and 26–30 years old) from the Human Connectome Project [56] (see Appendix D for subject IDs). The  $N = 94$  nodes in the network correspond to the cortical and subcortical regions defined by the AAL2 atlas-based segmentation [14] (cerebellum excluded). We performed probabilistic fiber tracking (using the Functional Magnetic Resonance Imaging of the Brain Software Library (FSL) [57], details on the processing pipeline in Appendix D) to determine the relative connection strength (edges) between these brain regions (nodes). The pairwise structural connectivity is summarized in the weighted adjacency matrix  $\mathbf{A}$  shown in Fig. 2. Since DTI data carries no directional information,  $\mathbf{A}$  must be symmetric. For the following, we define the weighted degree  $d_k$  of a node  $k$  as the sum over all afferent connection strengths, i.e.,  $d_k = \sum_{i=1}^N A_{ik}$ .

As motivated in Sec. III A, the dynamics of each node in the network is described by a FHN oscillator [Eq. (13)]. The general network dynamics, Eq. (1), then simplifies to

$$\begin{aligned} \frac{d}{dt}x_{k1}(t) &= h_1[\mathbf{x}_k(t)] + \sigma \sum_{i=1}^N A_{ki}x_{i1}(t) + \xi_k(t) + u_k(t), \\ \frac{d}{dt}x_{k2}(t) &= h_2[\mathbf{x}_k(t)], \end{aligned} \quad (14)$$

for all nodes  $k \in [1, N]$ . This implies setting the control matrix  $\mathbf{B} = \mathbf{I}_N$ , which means that all individual nodes can potentially receive independent control inputs. The local coupling, control, and noise schemes are set to  $\mathbf{G} = \mathbf{K} = \mathbf{D} = [[1, 0], [0, 0]]$ , i.e., the individual FHN nodes receive node-external inputs only through their activity variables  $x_{k1}$ .

We do not consider finite delays for reasons of simplicity. Finite delays induce additional dynamical states. Although of high interest in terms of whole-brain modeling, these would



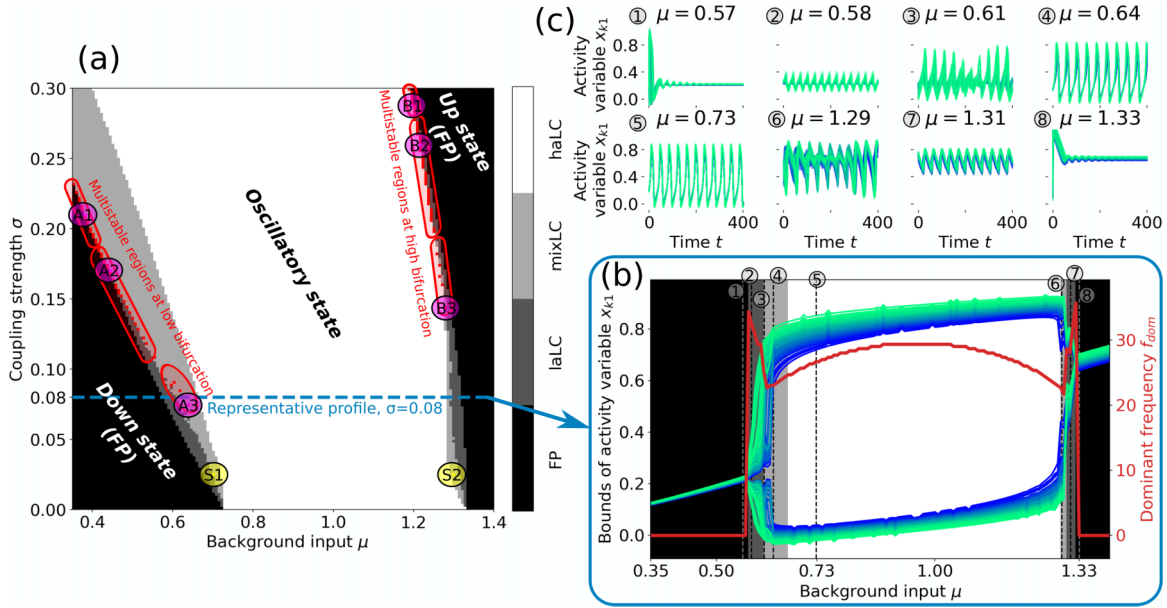


FIG. 3. Overview of the state space of the brain network model, Eqs. (13) and (14), for the noise free case ( $\eta = 0$ ). Time  $t$  measured in arb. units and simulations of each parameter configuration are started with initial conditions  $\mathbf{x}_0$  based on continuation and evaluated [except for (c1) and (c8)] after  $\bar{t} = 5000$  (see text for details). (a) Each pixel corresponds to a network with a particular parameter configuration for  $\mu \in [0.35, 1.4]$  and  $\sigma \in [0, 0.3]$ . States are classified based on the network oscillations, and oscillatory states are distinguished depending on whether no (laLC, light gray), some (mixLC, dark gray), or all (haLC, white) nodes fulfill the criterion of Eq. (16). Parameter configurations for which multiple stable solutions were detected (see text for details) and visually confirmed are marked with red pixels. Red lines enclose regions where we observe bistability between FP and mixLC (around point A1), laLC and mixLC (around points A2, B2), and mixLC and haLC (around point B1), as well as multistability in mixLC (around points A3, B3). Examples for bi- and multistable states are shown in Fig. 4. Points S1 and S2 indicate states with a low mean network cross-correlation  $R$  that are explored in Sec. IV B. (b) One dimensional bifurcation diagram for a network as a function of  $\mu$  for a fixed coupling strength of  $\sigma = 0.08$ . The dominant network frequency (red line) is calculated according to Eq. (15). Minimal and maximal values of the activity variable  $x_{k1}$  are plotted for each node individually, where line color indicates the weighted degree (green to blue indicates high to low values). (c) Example traces of the activity of all nodes in the network for different values of  $\mu$  [see (b)] and for  $\sigma = 0.08$ , showing different dynamical states. Line color indicates the weighted degree of the nodes [see (b)]. Traces at the FPs (1, 8) are shown after  $\bar{t} = 0$  (see text for details) to show the stability of the respective fixed point.

not add to the conclusions drawn in Secs. IV and IV B about the impact of nonlinear optimal control and its comparison with diagnostics [23] derived from linear control and based on connectome properties only. Please refer to Sec. VI for a discussion of the biological plausibility of this assumption.

### C. State space exploration

In this section, we describe and characterize the different dynamical states that can emerge in a whole-brain network of coupled FHN oscillators for a large range of parameter configurations. Having such an overview of the dynamical landscape of the system will allow us to formulate well defined control tasks in Sec. IV. Throughout the state space exploration we do not apply any control input [ $u_k(t) = 0 \forall k, t$ ].

Initially we consider the noise free case ( $\eta = 0$ ). Here, only two free model parameters remain: the global coupling strength  $\sigma$  and the time independent background input  $\mu$  which is the same for each FHN oscillator. The state space of the FHN network is explored by simulating the network dynamics for wide ranges of these parameters (see Appendix B for details). The initial conditions  $\mathbf{x}_0$  are drawn randomly from a uniform distribution in the interval  $[0, 1]$  or taken as the state vector  $\mathbf{x}'(t_{\text{end}})$  of the last time step of the previous simulation with same  $\sigma$  and slightly smaller  $\mu$  (continuation).

To avoid analyzing transient effects, we show and evaluate the network states only after a sufficiently long transient time  $\bar{t}$ .

If oscillations are present, then the dynamical state is characterized by the strength of synchronization quantified by the total cross-correlation, Eq. (6), and by the dominant frequency  $f_{\text{dom}}$ . The latter is given by the frequency of the highest peak of the combined power spectrum of all nodes,

$$f_{\text{dom}} = \arg \max_f \sum_{i=1}^N S_{xx,il}(f), \quad (15)$$

where  $S_{xx,il} = |(\mathcal{F}x_{il})(f)|^2$ , with Fourier transform  $\mathcal{F}$ , denotes the power spectral density of an individual node.

Figure 3 provides an overview of the state space of the FHN network. Similar to the case of a single FHN oscillator (cf. Fig. 1), the network shows two regions in parameter space, for which a stable fixed point (FP) exists. For small values of  $\mu$  and  $\sigma$  the activity variable  $x_{k1}$  has a low value for all nodes [down state, cf. Fig. 3(c)(1)]; for large values of  $\mu$  and  $\sigma$  the activity value is high [up state, cf. Fig. 3(c)(8)]. Figure 3(a) shows that the dynamics can transition from an down state to an oscillatory state, as well as from an oscillatory state to an up state by either increasing  $\mu$  or  $\sigma$ . For convenience, we call these transitions *low* and *high bifurcation*, respectively.

Within the oscillatory regime, we observe network states that are qualitatively different in their appearance, which we explain in the following by their different underlying mechanisms. A single, uncoupled FHN node oscillates in its limit cycle (LC) if it receives a background input in the range of  $\mu \in [0.73, 1.33]$  (cf. Fig. 1). We therefore expect a node in the network to show sustained oscillation (indiv. LC)—and consequently affecting the dynamics of the network—if its combined input is in this interval. We thus call a node “to be in its individual LC regime” if

$$0.73 \lesssim \mu + \sigma \sum_{i=1}^N A_{ki}x_{i1}(t) \lesssim 1.33, \quad (16)$$

for at least one point in time within the considered time interval. This criterion is further motivated by observations of the dynamics (cf. Fig. S1 in Ref. [78]) which show qualitative changes in behavior depending on whether it is met by all, some, or none of the nodes in the network.

If all nodes fulfill Eq. (16), then we observe the network to be in a synchronous, high-amplitude limit cycle [haLC in Fig. 3(a), cf. Fig. 3(c)(5)]. If, however, no node fulfills Eq. (16), then the network either inherits the fixed point solution from the individual nodes, or a new state emerges due to network effects. We observe both cases: The first one is true in both down and up state of the network in Fig. 3(a). The second case appears for low and intermediate coupling strength  $\sigma$  at the transition from down or up state to oscillatory state. In these states the FP of the network dynamics is unstable due to coupling effects and a self-sustained low amplitude limit cycle [laLC in Fig. 3(a), cf. Fig. 3(c)(2, 7)] emerges. The geometric interpretation of this laLC is a rotation around the FP that would be transient for an isolated node, but is prevented from converging to a fixed-point by the network couplings.

It is also possible that only a fraction of the network nodes fulfill the criterion in Eq. (16), which can lead—even in the absence of noise and without network delays—to asynchronous and apparent aperiodic behavior (mixLC in Fig. 3(a), cf. Fig. 3(c)(3, 6), further indications for aperiodicity are provided in Ref. [78]). On the network level, this can be seen as the result of the different frequencies of the two coexisting network limit cycles, laLC and haLC, interacting. Close to the bifurcation the frequency of the individual nodes (cf. Fig. 1), and their trajectories (cf. Fig. S3 in Ref. [78]), are particularly sensitive to their respective input. If the additive coupling input between the nodes is not sufficient to entrain a common frequency, this may result in asynchronous and potentially aperiodic network oscillations. States that are classified as mixLC are, however, not necessarily asynchronous. If the driving force of nodes that fulfill the criterion in Eq. (16) is high enough, it will lead to frequency entrainment. Therefore, the oscillations of the driven nodes [where Eq. (16) is not fulfilled] are similar to the ones of the driver nodes, resulting in dynamics that are similar to the haLC state [cf. Fig. 3(c)(4, 5)].

The distinction between different dynamical states and types of network oscillations (laLC, mixLC, haLC) is particularly interesting since the transitions between these are the regions in state space, where we find multistable network states. Multistability is detected by simulating the network

dynamics for 21 different initial conditions  $\mathbf{x}_0$  and comparing the resulting time series by calculating their nodewise correlation in the activity variable. We ignore the first  $\bar{t} = 5000$  arb. units of each time series to avoid analyzing transient effects and then compare the interval  $t \in [5000, 6000]$  (equiv. to 10–35 periods) of one initial condition with  $t \in [5000, 6200]$  of all other traces (second interval is sufficiently longer to account for all phase shifts). If the autocorrelation of at least one initial condition is close to one, this indicates the existence of a stable (periodic) state, and if the cross-correlation between two (or more) initial conditions is different from one, this indicates bistability (or multistability). Red pixels in Fig. 3(a) show states with more than one stable solution, which are observed along both, the low and high bifurcation.

Numerically we found regions of bistability between FP and mixLC (A1), laLC and mixLC (A2, B2), and mixLC and haLC (B1), as well as multistability in mixLC with different numbers of nodes fulfilling the criterion in Eq. (16) (A3, B3) [cf. Fig. 3(a)]. The corresponding time series of the activity variables  $x_{k1}$  are shown in Fig. 4. The automatic detection of multistability might not capture all multistable states, especially because some of the states may be very similar [cf. Fig. 4(b), bottom panel]. More detailed analyses could thus also uncover bistabilities between up state and mixLC, as well as between mixLC and haLC at the low bifurcation. We inspected the detected multistable states visually to assure that their difference are not due to transient effects and confirm their assignment to the specified multistable regions in Fig. 3(a). We do not observe state switches in this noise-free case even for very long simulation times [200 000 arb. units simulated for multiple initial conditions in each multistable region, A1–3, B1–3, of Fig. 3(a)], meaning that the states are stable and no temporal intermittency is found.

Adding noise to the network most strongly affects the dynamics of the states close to the bifurcations. The influence of noise on synchronous, mono-stable oscillatory states is, as expected, comparably small (cf. Figs. S4(d) and S4(e) in Ref. [78]). If the network dynamics is in a FP far away from the bifurcation line, the additional Gaussian white noise results in uncorrelated fluctuations in the activity variable (cf. Fig. S4(i) in Ref. [78]). Parametrizations that without noise would lead to a stable FP close to the bifurcation, show oscillations when sufficient noise is introduced (cf. Figs. S4(a) and S4(h) in Ref. [78]). The clear distinction between the FP and laLC network states in the noise-free case [cf. Fig. 3(a)] is therefore blurred for noisy dynamics. For asynchronous mixLC states, the additional noise can have a synchronizing effect (cf. Fig. S4(c) in Ref. [78]). This can be explained by more nodes fulfilling the criterion in Eq. (16), resulting in a stronger drive for the remaining nodes and therefore more synchronous oscillation. These effects lead to a shift of the bifurcation lines compared to the noise-free case and consequently also affects the location of multistable network states.

In the case of multistability, we observe that this stochastic additional input can lead to sufficient perturbations that drive the dynamics from one attractor to the other. This results in noise-induced state switching, as shown in Fig. 4(c), at the bifurcation lines of the noisy state space. Such noise-induced transitions are a known phenomenon in systems of coupled

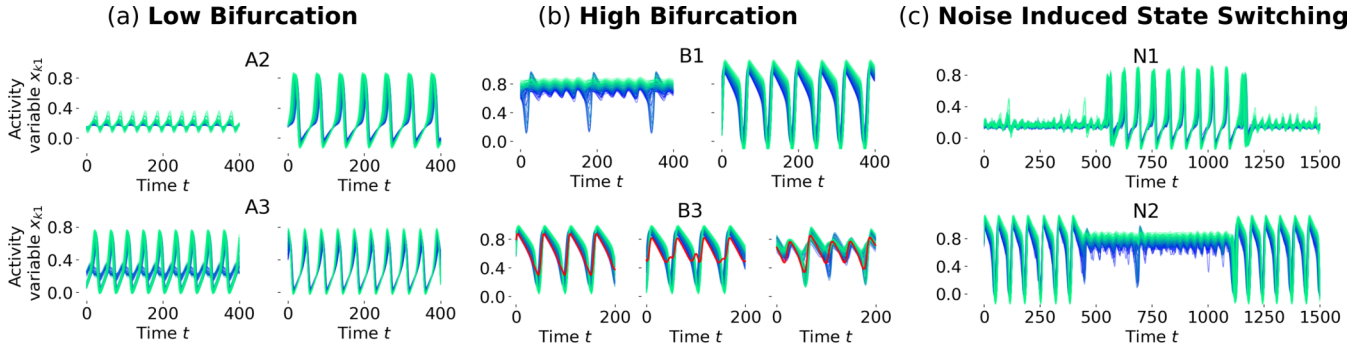


FIG. 4. Example traces of the activity variable  $x_{k1}$  of each node from the regions of bi- and multistability indicated by the labeled points in Fig. 3(a). Plots show the dynamics for random initial conditions after a sufficiently long transient time ( $\bar{t} = 5000$  arb. units). Line color indicates the weighted degree (green to blue indicates high to low values). (a) Traces at low bifurcation, top: bistability between laLC (left) and mixLC (right) in point A2 ( $\mu = 0.44$ ,  $\sigma = 0.17$ ,  $\eta = 0$ ); bottom: bistability between states in mixLC in point A3 ( $\mu = 0.64$ ,  $\sigma = 0.07$ ,  $\eta = 0$ ). (b) Traces at high bifurcation, top: bistability between mixLC (left) and haLC (right) in point B1 ( $\mu = 1.2$ ,  $\sigma = 0.29$ ,  $\eta = 0$ ); bottom: multistability between states in mixLC in point B3 ( $\mu = 1.28$ ,  $\sigma = 0.17$ ,  $\eta = 0$ ). The main difference between left and middle panels is the trajectory of the node with the smallest in-degree (highlighted in red). (c) Traces showing noise induced state switching. Locations N1 and N2 are also close to the bifurcations but are not shown in Fig. 3(a) since the state space changes under the influence of noise. Top: Example at the low bifurcation in point N1 ( $\mu = 0.39$ ,  $\sigma = 0.2$ ,  $\eta = 0.024$ ). Bottom: Example at the high bifurcation in point N2 ( $\mu = 1.2$ ,  $\sigma = 0.3$ ,  $\eta = 0.024$ ).

oscillators [58] and may be exploited for the control of the neural dynamics in practice [59].

#### IV. OPTIMAL CONTROL OF THE BRAIN NETWORK DYNAMICS

##### A. Switching between bistable network states

In this section, we present optimal control inputs that induce a switch between previously identified multistable states. All control inputs optimize the cost functional given in Eq. (2) with the state dependent terms given in Eq. (3) and the energy and sparsity terms given in Eq. (7). Energy and sparsity terms are evaluated over the whole time interval during which the control is active. The cost functional itself, however, considers the deviation from the target state only the end of the control period. For this we set  $I_p(t) = I_p^*$  for  $(T - \tau) \leq t \leq T$  and  $I_p = 0$  else. The convergence criterion of the numerical solution of the minimization problem, as described in Sec. IIC, is set to  $\epsilon = 10^{-5}$  for all our applications. We ensured for all presented examples that the method actually converges and that results do not change for lower values for  $\epsilon$ . We computed the optimal control for different phase shifts of the initial and target states with respect to the control onset and present the results for the phase shift which leads to the smallest total control energy  $E$ .

Figure 5 shows the result of applying optimal control at point A1 in Fig. 3(a), where a low-activity fixed point coexists with an oscillatory mixed state. The weight  $I_s$  for the sparseness term in Eq. (7) was set to zero. Since the target state in this task is always stable, the network remains in this state once it is reached also after the control is turned off. If the initial state is the down state FP, then the obtained optimal control input oscillates synchronously for all nodes with increasing amplitude. The frequency of the control input [ $f_{\text{dom}}(\bar{u}_1) = 35.0$ ] corresponds to the frequency of the fixed point's focus [ $f_{\text{dom}}(x_1^{\text{FP}}) = 35.0$ , measured by perturbing the FP], inducing resonance effects in the node activity. This strat-

egy can be well observed in Video 1 (in Ref. [78]), showing the node oscillations and the respective optimal control inputs in state space.

When switching from the oscillatory to the down-state, as in Fig. 5(d), the optimal control input consists of one short biphasic pulse applied to all of the nodes followed by minor corrections. It is a known result from control theory, that applying a biphasic control pulse around an extreme point is an efficient way to drive an oscillating system to a stop. Interestingly, the optimal control strategy in this example is not to apply the pulse at an extreme point of the activity variables  $x_{k1}$ , but rather in a way that the gradient of the control is highest when the phase velocity of the limit cycle oscillation

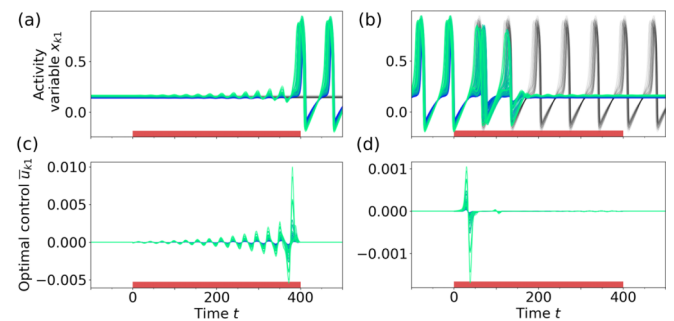


FIG. 5. State switching with optimal control between bistable states at the low bifurcation [point A1 in Fig. 3(a)]. (a) Switch from down-state to oscillatory state and (b) vice versa. Activities  $x_{k1}$  of each node  $k$  are shown over time. The colored lines show the controlled activities with green to blue indicating nodes with high to low weighted degree. The black lines correspond to the uncontrolled activities. (c, d) Corresponding optimal control inputs  $\bar{u}_{k1}$  to each node. The red bar indicates the time interval during which the control is active (from  $t = 0$  to  $t = 400$ ). The deviation from the target state is penalized in the last  $\tau = 25$  time units of the control. Parameters were:  $\mu = 0.378$ ,  $\sigma = 0.21$ ,  $\eta = 0.0$ ,  $I_p^* = 0.0005$ ,  $I_e = 1.0$ ,  $I_s = 0$ .

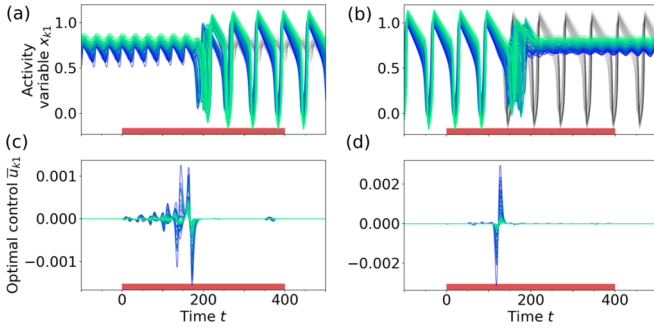


FIG. 6. State switching with optimal control between bistable states at the high bifurcation [point B2 in Fig. 3(a)]. (a) Switch from low-amplitude oscillation to high-amplitude oscillation and (b) vice versa. Activities  $x_{k1}$  of each node  $k$  are shown over time. The colored lines show the controlled activities with green to blue indicating nodes with high to low weighted degree. The black lines correspond to the uncontrolled activities. (c, d) Corresponding optimal control inputs  $\bar{u}_{k1}$  to each node. The red bar indicates the time interval during which the control is active (from  $t = 0$  to  $t = 400$ ). The deviation from the target state is penalized in the last  $\tau = 25$  time units of the control. Parameters were  $\mu = 1.22$ ,  $\sigma = 0.26$ ,  $\eta = 0.0$ ,  $I_p^* = 7 \times 10^{-5}$ ,  $I_e = 1.0$ ,  $I_s = 0$ .

is the lowest (best observed in Video 2 in Ref. [78]). Although the deviation from the target state is only penalized at the end of the control interval, the pulse is applied early to give the system enough time to come to rest.

Figure 6 shows the result of applying optimal control at point B1 in Fig. 3(a), where a low amplitude limit cycle (laLC) around the high-activity up-state coexists with an oscillatory mixed state (mixLC, cf. Sec. III). When switching from the laLC to the mixLC state [Figs. 6(a) and 6(c)], the frequency of the control input [ $f_{\text{dom}}(\bar{u}_1) = 30.0$ ] is again adapted to the frequency of the initial state [ $f_{\text{dom}}(x_1^{\text{laLC}}) = 30.0$ ], utilizing resonance effects. While the nodes in the laLC state all oscillate with the same frequency ( $f_k = f_{\text{dom}} = 30.0 \forall k$ ), their phase speed is not the same at all times (reflected by lower synchrony  $R(x_1^{\text{laLC}}) = 0.861$ , best observed in Video 3 in Ref. [78]). This results in a less synchronous oscillation of the control signals in Fig. 6(c) ( $R(\bar{u}_1) = 0.714$ ) compared to the control inputs at the low bifurcation [Fig. 5(c),  $R(\bar{u}_1) = 0.999$ ]. In the other direction [Figs. 6(b) and 6(d), Video 4 in Ref. [78]), the optimal control finds a short, off-phase biphasic pulse analogous to Fig. 5(d). As a result, the amplitude of the network oscillation decreases to almost zero [for  $t \in [200, 250]$  in Fig. 6(b)]. Since there is no stable FP, the oscillation amplitude increases again over time and converges to the laLC target state approx. 550 time units after the control pulse (see also Fig. S5 in Ref. [78]).

By increasing the sparsity parameter  $I_s$  of the cost functional, we can tune the number of controlled nodes. Figure 7 shows the control energy  $E_k$  [Eq. (12)] of the optimal control input  $\bar{u}_{k1}$  to each node  $k$  as a function of the sparsity parameter  $I_s$ . For low values of  $I_s$ , all nodes receive a finite control signal. When  $I_s$  is increased, less nodes are controlled, until  $I_s$  becomes so large, that the target state can no longer be reached. As expected, decreasing the number of controlled nodes needs to be compensated for with a higher control energy.

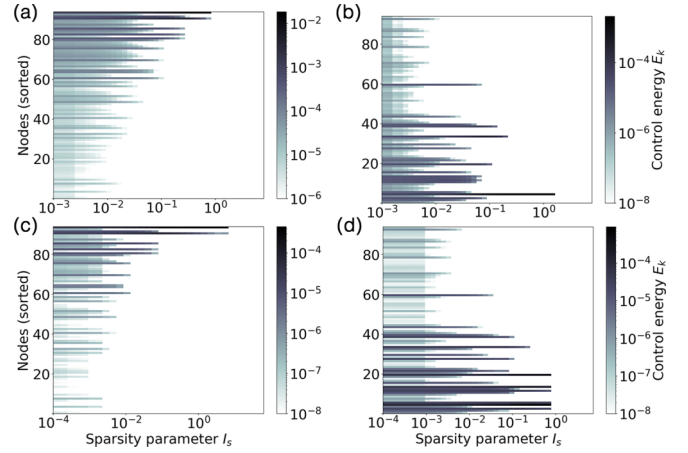


FIG. 7. State switching with sparse optimal control. The nodes are sorted from lowest (bottom) to highest (top) weighted degree. The length of the bars indicate up to which value of the spatial sparsity parameter  $I_s$  the node still receives finite control input. Their color denotes the corresponding optimal control energies  $E_k$  [Eq. (12)] for each node. (a) Switching from the down state (FP) to the oscillatory state (mixLC) with parameters close to the low bifurcation as in Fig. 5 and (b) Switching from low-amplitude (laLC) to high-amplitude oscillatory state (mixLC) with parameters close to the high bifurcation as in Fig. 6. (c) Same as panel (a) but switching from mixLC to FP. (d) Same as panel (b) but switching from mixLC to laLC.

The results also show that at the low bifurcation [Figs. 7(a) and 7(c)], nodes with a high weighted degree receive, independent of the switching direction, a stronger control signal and remain controlled even for high values of  $I_s$ . This shows that at the low bifurcation it is most important to control the network hubs since they act as driver nodes for attractor switching. At the high bifurcation [Figs. 7(b) and 7(d)], however, nodes with low weighted degree receive the strongest control inputs.

This different behavior can be explained based on the additive coupling scheme between the nodes (cf. Sec. III), which causes nodes with higher degree to typically receive stronger inputs. When choosing parameters close to the low bifurcation, these high degree nodes therefore have higher oscillation amplitudes (cf. Fig. 1) and are consequently driving the oscillation of the remaining nodes in the network. Close to the high bifurcation the nodes with lower degree, who receive less inputs, are more likely to remain in the limit cycle regime. Consequently – and in contrast to the dynamics close to the low bifurcation – the low degree nodes drive the oscillation of the network hubs. Videos 1 and 3 (in Ref. [78]) illustrate how the control inputs on the respective driver nodes force the network dynamics to the high-amplitude oscillation target states.

The optimal control inputs to the control sites have well interpretable shapes. When the objective is to transition from a low-amplitude state to one with a higher amplitude [mixLC, Figs. 5(a) and 6(a)], the optimal control strategy utilizes the characteristics of the flow field in state space and synchronously drives the network with its resonant frequency from one attractor toward the other. Switching in the opposite direction, and therefore leaving this basin of attraction,



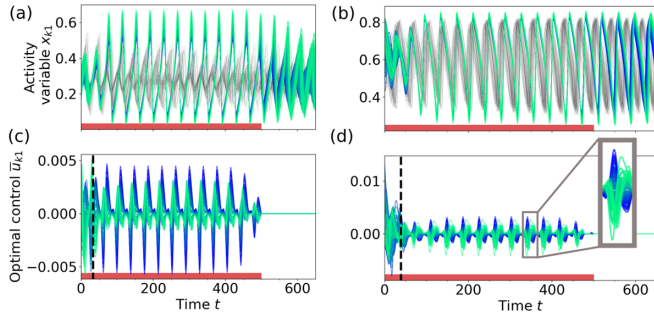


FIG. 8. Synchronizing the noise-free network with optimal control inputs. (a) Synchronization task at point S1 (low bifurcation) in Fig. 3(a). Activities  $x_{k1}$  of each node  $k$  are shown over time. The colored lines show the controlled activities (average cross-correlation  $R = 0.995$  [Eq. (6)]) with green to blue indicating nodes with high to low weighted degree. The black lines show the uncontrolled activities ( $R = 0.289$ ). (b) Synchronization task at point S2 (high bifurcation) in Fig. 3(a) (controlled:  $R = 0.996$ , uncontrolled:  $R = 0.368$ ). (c, d) Corresponding optimal control input  $\bar{u}_{k1}$  to each node. The red bar indicates the time interval during which the control is active (from  $t = 0$  to  $t = 500$ ). The vertical black dashed line indicates the critical time  $t_c$  (see text). Parameters were: Point S1  $\mu = 0.7$  and  $\sigma = 0.025$ , point S2  $\mu = 1.3$  and  $\sigma = 0.025$ , other parameters:  $\eta = 0$ ,  $I_p = 0.1$ ,  $I_e = 1.0$ ,  $I_s = 0$ .

is achieved most efficiently with one strong, biphasic pulse. This deflects the system from its initial mixLC trajectory and causes it to drop to the attractor with a lower amplitude.

### B. Synchronizing the network dynamics

In this section, we apply the nonlinear control method to find the optimal inputs to synchronize the dynamics of the individual nodes. For this application we use the state dependent cost functionals given in Eq. (4) to penalize the deviation from the target cross correlation ( $R_T = 1$ , fully synchronous state) and in Eq. (7) to penalize the energy of the control and enforce its sparsity in space.

We parametrize the system to be in the two asynchronous regions marked by the labels S1 and S2 in Fig. 3(a), close to the low and high bifurcation lines. The optimal control method is then applied to synchronize the dynamics for the noise-free (Fig. 8) and noisy (Fig. 10) case. Both figures show the controlled (synchronous) and uncontrolled (asynchronous) time series and the corresponding optimal control inputs. Since the dynamics in points S1 and S2 are monostable, the synchronous state is not a stable solution and the system returns to its original state as soon as the control is switched off. The network dynamics and the optimal control strategies are best seen in Videos 5–8 (in Ref. [78]).

Figure 8 shows how the optimal control inputs synchronize the oscillation of all nodes in the network. In the uncontrolled state [in gray, Figs. 8(a) and 8(b)] the nodes at both bifurcations oscillate with similar frequencies (mean and standard deviation of oscillation frequencies across nodes at S1:  $\langle f \rangle_N = 32.2$ ,  $\sigma_{\langle f \rangle_N} = 1.1$ , and at S2:  $\langle f \rangle_N = 27.0$ ,  $\sigma_{\langle f \rangle_N} = 0.9$ ). Close inspection of the control inputs in Figs. 8(c) and 8(d) show that the optimal control acts in two phases. First, the control aligns the phases of all oscillators until all nodes

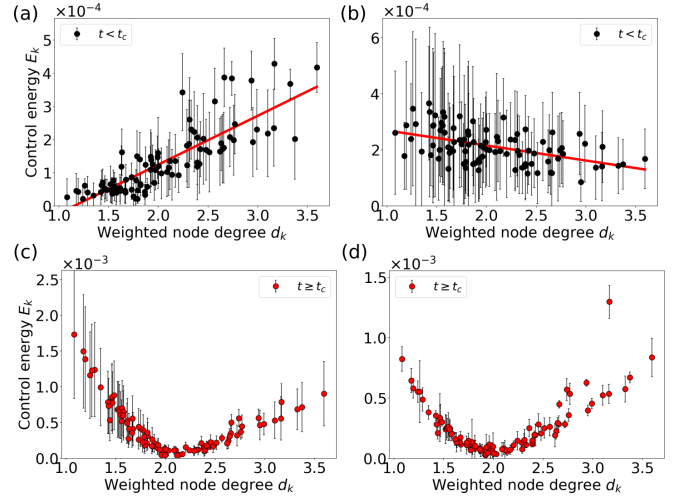


FIG. 9. Mean energies  $E_k$  of the control inputs [Eq. (12)] as a function of the weighted node degree  $d_k$  for the noise-free synchronization task. The error bars show the standard deviation over 10 different initial conditions of the network dynamics. (a, b) The control signal for time  $t \in [0, t_c]$  is considered. Linear regression (red line) coefficients are  $r = 0.83$ ,  $p < 10^{-24}$  and  $r = -0.49$ ,  $p < 10^{-6}$ . (c, d) The control signal for time  $t \in [t_c, T]$  is considered. (a, c) Point S1 (low bifurcation) in Fig. 3(a). (b, d) Point S2 (high bifurcation) in Fig. 3(a). All parameters are as in Fig. 8.

are synchronized for the first time, which we defined as the critical time  $t_c$  (please refer to Ref. [78] for details about the computation of  $t_c$ ). Second, a periodic input maintains the synchronous oscillation during the period the control is active.

For times  $t < t_c$ , Fig. 9(a) shows that at the low bifurcation the energies  $E_k$  of the control inputs to the nodes are positively correlated with the weighted node degrees. Thus, it is beneficial to focus the control input on the network hubs for alignment. A possible interpretation for this strategy is that the optimal control utilizes the influence of the network hubs on the remaining nodes to force them on the synchronous limit cycle trajectory. However, we again observe a different behavior at the high bifurcation, as shown in Fig. 9(b). Here,  $E_k$  is negatively correlated with the weighted node degree for times  $t < t_c$ . In this case, the network coupling is much smaller compared to the background input  $\mu$  and the node's phase space oscillations are similar to the trajectory of the limit cycle of an uncoupled node (cf. Video 6 in Ref. [78]). The negative correlation in Fig. 9(b) suggests that if the dynamics of the nodes are similar in the first place, it is beneficial to focus the control on more weakly coupled nodes and align their phase space trajectory with the trajectory of the network hubs. Similar to the attractor switching at the high bifurcation [cf. Figs. 7(b) and 7(d)], the control of the low degree nodes drives the network oscillation toward the target state.

In the second phase, for  $t \geq t_c$ , the high degree nodes without control would have a higher (or lower) phase velocity than the nodes with intermediate degree, while the opposite is true for nodes with low degree. The optimal control strategy is thus to act on the low and high degree nodes in opposite directions, which can be observed as antiphase control inputs shown in Figs. 8(c) and 8(d), where nodes with intermediate

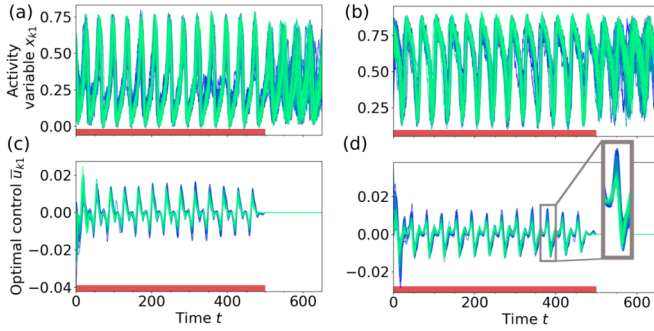


FIG. 10. Synchronizing the noisy network with optimal control inputs. (a) Synchronization task at point S1 (low bifurcation) in Fig. 3(a). Activities  $x_{k1}$  of each node  $k$  are shown over time with green to blue indicating nodes with high to low weighted degree (average cross-correlation  $R = 0.845$  [Eq. (6)], uncontrolled:  $R = 0.210$ ). (b) Synchronization task at point S2 (high bifurcation) in Fig. 3(a) (controlled:  $R = 0.853$ , uncontrolled:  $R = 0.321$ ). (c, d) Corresponding optimal control input  $\bar{u}_{k1}$  to each node. The red bar indicates the time interval during which the control is active (from  $t = 0$  to  $t = 500$ ). Parameters were  $\eta = 0.024$ ,  $I_p = 0.1$ ,  $I_e = 1.0$ ,  $I_s = 0$ , all other parameters are as in Fig. 8.

degree receive only small control inputs. (Videos 5 and 6 in Ref. [78] show how the control inputs keep the nodes together in phase space.) This is reflected in the arclike form of the mean energies  $E_k$  of the control inputs at both bifurcations, as seen in Figs. 9(c) and 9(d), with high control energies for nodes with low or high degree. As a result, the optimal control achieves a mean network cross-correlation of  $R = 0.995$  in the time interval  $t \in [t_c, T]$ , compared to the uncontrolled scenario with  $R = 0.289$  at point S1 [Fig. 8(a)], and  $R = 0.996$  compared to  $R = 0.368$  at point S2 [Fig. 8(b)].

Figure 10 shows results for the application of optimal control to the synchronization task for the case of finite noise. We obtain a mean network cross-correlation of  $R = 0.83$  (analysis of the deviation from  $R_T$  in Ref. [78], uncontrolled:  $R = 0.25$ ) at point S1 [Fig. 10(a)] and  $R = 0.85$  (uncontrolled:  $R = 0.35$ ) at point S2 [Fig. 10(b)]. We conclude that the optimal control input successfully synchronizes the network dynamics, just as in the noise-free case.

The control strategy that leads to the respective target state, however, changes substantially when noise is added to the system. The control input in the noise-free case is tailored individually for each node, slowing the phase space velocity of some nodes down and speeding others up at the same time [cf. blue inset in Fig. 8(d), mean cross-correlation of the control inputs during this control period is 0.03]. In contrast to that, the control input for the noisy network dynamics has a similar shape for all nodes (cf. blue inset in Fig. 10(d), mean cross-correlation of the control inputs during control period is 0.88) with varying amplitude depending on the weighted node degree  $d_k$ .

This change in control strategy is also reflected in Fig. 11, which shows the control energy  $E_k$  per node for different noise levels. The control energy increases for all nodes with increasing noise strength, both at the low [Fig. 11(a)] and the high bifurcation [Fig. 11(b)]. With increasing noise level  $\eta$ , we observe a gradual transition of the optimal control strategy. In-

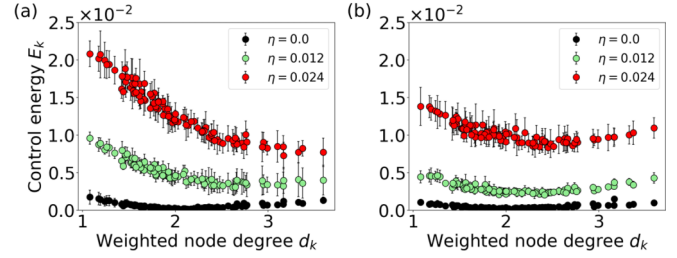


FIG. 11. Mean energies  $E_k$  of the optimal control inputs [Eq. (12)] as a function of the weighted node degree  $d_k$  for different noise levels  $\eta$  for the synchronization task. The error bars show the standard deviation over 10 different initial conditions of the network dynamics with 20 independent noise realizations of the optimization each.  $E_k$  is, in contrast to Fig. 9, calculated over the whole interval during which the control is active. (a) Point S1 (low bifurcation) in Fig. 3(a). (b) Point S2 (high bifurcation) in Fig. 3(a). Parameters were  $I_p = 0.1$ ,  $I_e = 1.0$ ,  $I_s = 0$ , all other parameters are as in Fig. 8.

stead of balancing out the phase velocity differences between nodes toward the phase of nodes with intermediate degree (cf. Videos 5 and 7 in Ref. [78]), the control in the noisy case forces the system on a limit cycle with a slightly higher oscillation amplitude (cf. Videos 6 and 8 in Ref. [78]). This strategy requires a higher control energy but is independent of the specific realization of the noise.

By imposing sparsity constraints on the system, we can investigate to what extent synchronization can be achieved with fewer control sites. Figure 12 shows the relation of the synchrony in the network, measured by the average cross-correlation  $R$  [Eq. (6)], to the sparsity parameter  $I_s$ . The average cross-correlation  $R$  decreases with increasing sparsity parameter  $I_s$  at both locations in state space and for all noise levels. This shows that, under the imposed constraints, it is not possible to fully synchronize the network with sparse control. Optimally synchronizing aperiodic states is a collective effort. Unlike state switching it cannot be achieved by controlling only a few sites, because the reduced number of control sites cannot sufficiently be compensated for by a higher control energy. Especially in the case of noisy network dynamics, where the control must drive all nodes, the network's synchrony

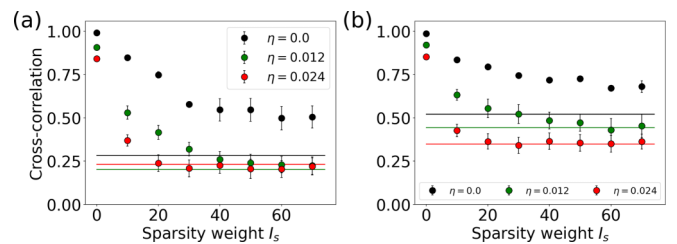


FIG. 12. Average network cross-correlation  $R$  [Eq. (6)] with sparse optimal control as a function of the sparsity parameter  $I_s$  [Eq. (7)] for different noise levels  $\eta$ . The mean and standard deviation are shown for 5 different initial conditions of the network dynamics, each with 20 independent noise realizations. Horizontal lines indicate the cross correlation  $R$  for the uncontrolled case. (a) Point S1 (low bifurcation) in Fig. 3(a). (b) Point S2 (high bifurcation) in Fig. 3(a). Parameters were  $I_p = 0.1$ ,  $I_e = 1.0$ , all other parameters are as in Fig. 8.

quickly drops to the value of the uncontrolled network when  $I_s$  is increased. For fixed  $I_s$  the number of controlled nodes and the control sites optimal for synchronizing the network changes for each initial condition  $\mathbf{x}_0$  (cf. Fig. S1 in Ref. [78]), as it depends on details of the exact network dynamics.

## V. COMPARISON WITH CONTROLLABILITY MEASURES DERIVED FROM LINEAR CONTROL THEORY

When relating functional properties of a neural system to properties of the underlying connectome, neural activity is often approximated by linear dynamical systems [23,60,61]. This has obvious benefits for analyzing the effects of perturbations. Calculating optimal control inputs for linear systems, as in Refs. [12,22], can be done analytically and with little computational effort, and conclusions about the effects of external inputs can be drawn from controllability measures, which depend on network topology only [21].

Two of these measures have previously been applied to quantify the impact of perturbations in a whole-brain network setting [22,23,25]. *Modal controllability* refers to the ability of a node to control each evolutionary mode of a dynamical network [62], and the *average controllability* is given by the average control input energy to the respective node over all possible target states [23,25] (mathematical description in Appendix C). Nodes with high average controllability require only low energy input to move a linear system into “easy-to-reach” states, while nodes with high modal controllability require a high control energy input to have an effect on the dynamics but are crucial when the target state of the system is “difficult-to-reach” [21,23,63]. It has also been shown that the average controllability of a node is strongly correlated with its weighted degree  $d_k$  ( $r = 0.85$ ,  $p < 10^{-26}$  for our structural connectivity matrix) while the modal controllability is known to have a strong inverse correlation ( $r = -0.82$ ,  $p < 10^{-23}$ , see Fig. S7 in Ref. [78]).

We now apply the diagnostics from linear control theory to the brain network model, Eq. (14). Figure 13 shows the correlations of the energies  $E_k$  of the optimal control inputs with the average and modal controllability for the attractor switching tasks. At the low bifurcation A1 [Figs. 13(a) and 13(c)] we find, that the energy of the control input for a node is positively correlated with its average controllability and negatively correlated with its modal controllability, both irrespective of the switching direction. Therefore, the most efficient strategy for switching between the attractors is to control the network hubs with high average controllability which then force the other nodes toward the target state. These results can be considered consistent with predictions from linear control theory [23] and previous results on the global impact of stimulation [25].

At the high bifurcation at location B2 [Figs. 13(b) and 13(d)], however, we observe the opposite trend, with the control energy being negatively correlated with the average and positively correlated with the modal controllability. Predictions based on linear control theory alone are indifferent to the actual dynamics and therefore not able to distinguish between the two cases. Yet, nodes with high modal controllability play a much more important role in affecting the global dynamics at this location in state space.

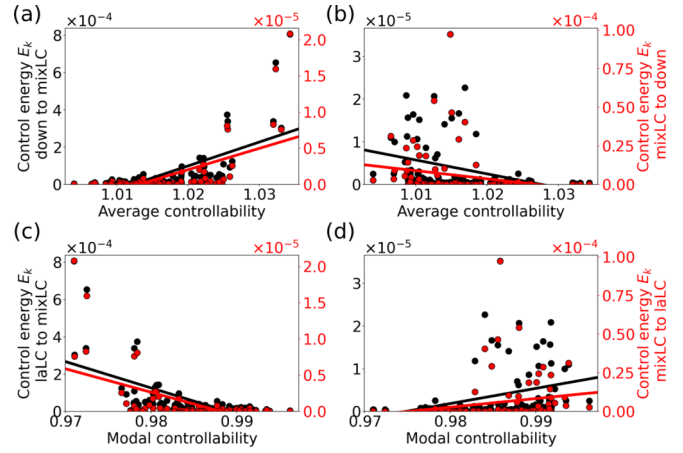


FIG. 13. Mean energies  $E_k$  of the optimal control inputs [Eq. (12)] plotted against measures from linear control theory for state switching tasks from down state / low amplitude oscillation toward high-amplitude oscillation (black) and the opposite direction (red). (a)  $E_k$  vs. average controllability when switching at the low bifurcation [Point A1 in Fig. 3(a)]. (b) Same as (a) but at the high bifurcation [Point B2 in Fig. 3(a)]. (c)  $E_k$  vs. modal controllability at A1 and (d) at B2. Solid lines denote the results of a linear regression with the following obtained coefficients: (a)  $r = 0.66$ ,  $p < 10^{-12}$  for the down-to-up switch (black) and  $r = 0.61$ ,  $p < 10^{-10}$  for the up-to-down switch (red). (b)  $r = -0.33$ ,  $p = 0.001$  (black) and  $r = -0.24$ ,  $p = 0.02$  (red). (c)  $r = -0.62$ ,  $p < 10^{-10}$  (black) and  $r = -0.57$ ,  $p < 10^{-8}$  (red). (d)  $r = 0.31$ ,  $p = 0.003$  (black) and  $r = 0.21$ ,  $p = 0.04$  (red). Same parameters as in Figs. 5 and 6, for low and high bifurcation, respectively.

For the synchronization task, we observe a similar pattern. While there is a clear correlation between the optimal nodewise control energies  $E_k$  with the average controllability when aligning the phases (i.e.,  $t < t_c$ ) at the low bifurcation [Fig. 14(a)], no correlation is observed at the high bifurcation [Fig. 14(b)]. (The corresponding results involving modal controllability are shown in Fig. S8 in Ref. [78].) In the case of maintaining the synchronous network dynamics [ $t \geq t_c$ , Figs. 14(c) and 14(d)], we observe—at both bifurcations—a similar but less obvious arclike shape as in Figs. 9(c) and 9(d). For noisy network dynamics (cf. Fig. S9 in Ref. [78]), however,  $E_k$  is negatively correlated with the average controllability at both bifurcations (cf. Fig. S9 in Ref. [78]).

This shows that in our nonlinear setting, linear controllability measures do not provide additional insights compared to the weighted node degree, and that intuitions based on these measures can be misleading. Diagnostics from linear control theory have previously been claimed to be predictive for a node’s role in driving brain state transitions [22,23] or for the global impact of local brain stimulation [25]. Our results however show, that the optimal control inputs and sites in nonlinear systems not only depend on the structural network connectivity, but also on the location in state space, the control task, and other factors like the amount of noise.

## VI. DISCUSSION

In this contribution we apply techniques from the optimal control of nonlinear dynamical systems to the dynamics of



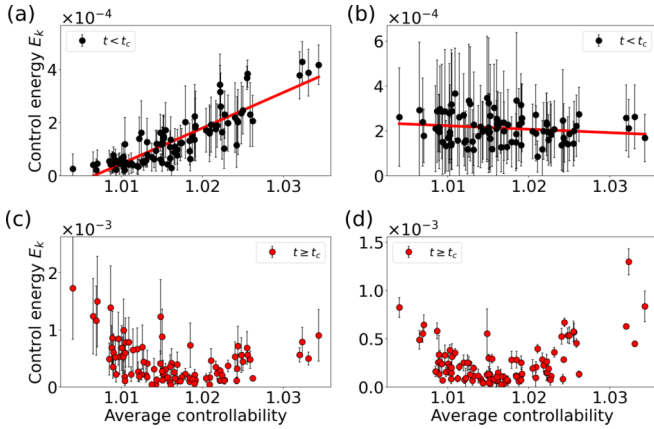


FIG. 14. Mean energies  $E_k$  of the optimal control inputs [Eq. (12)] plotted against the average controllability for the synchronization task (cf. Fig. 9). (a)  $E_k$  vs. average controllability, with control signal considered during the initial time interval  $t \in [0, t_c]$ , at the low bifurcation [Point S1 in Fig. 3(a)]. (b) Same as (a) but at the high bifurcation [Point S2 in Fig. 3(a)]. Solid red lines denote the linear regression results with the following obtained coefficients: (a)  $r = 0.87$ ,  $p < 10^{-29}$ , (b)  $r = -0.16$ ,  $p = 0.12$  (not significant). (c)  $E_k$  vs. average controllability, with control signal considered for time  $t \in [t_c, T]$ , at S1. (d) Same as (c) but at S2. All parameters as in Fig. 9.

brain network models. Nodes were equipped with FitzHugh-Nagumo oscillators, since they are simple and well studied nonlinear models for neural dynamics. Changing the background input for the FHN nodes or the global coupling strength of the network can both lead to transitions between two different stable fixed points and two different limit cycle attractors (laLC and haLC). The interaction between nodes in different oscillation states (mixLC) can lead to an asynchronous network dynamics. At the bifurcations, we also find different coexisting stable states.

The general mathematical framework of the optimal control of partial differential equations [29] is adapted for noisy dynamical systems on graphs, where the local network dynamics, the noise level, the network connectivity, and the local coupling schemes [Eq. (1)] can be freely chosen. The state dependent part of the cost functional is averaged over multiple noise realizations and penalizes the deviation of the network dynamics from a task dependent control target [Eq. (3) for attractor switching, Eq. (4) for synchronization]. The part of the cost functional which depends on the control input [Eq. (7)] penalizes the control energy and nonspare solutions. The presented method is applicable to any network that can be described in the form of Eq. (1), including models of power grids, social networks, or climate dynamics.

A common problem for gradient based methods are potential local optima of the cost functional which may prevent the convergence to a globally optimal solution. To alleviate this problem we performed the minimization as described in Sec. II C with different initial conditions  $\mathbf{u}_0$  for the control time-series and choose the result with minimal cost. The set of initial conditions included  $\mathbf{u}_0 = 0$  as well as valid control time-series taken from different parametrizations.

We use the optimal control method to cause targeted attractor switching between previously identified coexisting stable

states. When no sparsity is enforced, we show that it is optimal to resonantly drive all nodes to transition to the high-amplitude oscillatory state. When the task is to switch to a state with lower amplitude or no oscillation, the optimal strategy is to apply a precisely timed biphasic pulse. The nodes that receive the largest control energy are the same for both of these switching directions and are also the ones that are still controlled when we enforce sparsity in space. Depending on the location in state space, either nodes with high degree (at the low bifurcation) or low degree (at the high bifurcation) are the ones that most efficiently drive the network dynamics from the initial to the target state. When sparsity is enforced, controlling only a small number of these driving nodes with increased control energies  $E_k$  is sufficient to switch from one attractor to another in an optimal way.

In the second application, we show that our method can also be used to control global properties of the network dynamics, such as the average cross-correlation between node activities in the oscillating regime. Immediately with control onset, the control signal acts on all nodes to align their phases. As soon as this is achieved, the control maintains the synchronous oscillation with periodic control signals. Which nodes receive larger control input for the initial alignment again depends on the location in the state space. Both at the low and high bifurcation, individually adapted control inputs lead to a successful synchronization of the network dynamics. While the average cross-correlation  $R$  is increased also with sparse control, the synchronous target state is only achieved in an optimal way, when all nodes in the network receive a finite control input. This suggests that synchronizing all nodes needs collective intervention, while attractor switching can be caused by controlling a few selected nodes only. The introduction of noise to the system makes the dynamics of each node less predictable, resulting in a loss in specificity and more similar control inputs to the nodes. Consequently, the optimal control of noisy network dynamics requires higher total control energy  $E$  (cf. Fig. 11) while resulting in a lower precision (cf. Fig. 12 for  $I_s = 0$ ).

The information on the different states and bifurcations, which we show to be a decisive factor for choosing the optimal control sites in our applications, is lost when techniques from linear control theory are applied. While predictions do qualitatively agree for certain dynamical systems, control tasks, and locations in state space (cf. Ref. [25]), this agreement does not hold in a general setting. It would, however, be worthwhile investigating for what classes of dynamical systems and control tasks “controllability measures” can be defined, which only depend on the properties of the connectome.

In this contribution, the techniques of nonlinear optimal control were applied to a simplified model of the global brain dynamics. The bifurcations in our FHN model (cf. Fig. 1) phenomenologically capture the state transitions found in more complex, biophysically motivated network models [55] as we have adapted the FHN parameters to resemble their dynamics. Consequently, the activity variables of the FHN nodes can be interpreted as relative output firing rates of cortical nodes, and their total input  $\mu + \sigma \sum_{i=1}^N A_{ki} x_{i1}(t)$  can – at least qualitatively – be interpreted as a proxy of the local field potential [64]. Given this interpretation, model down-(up-)states are



related to down-(up-)states in cortical physiology, while an oscillatory regime of the whole-brain model corresponds to a brain state with oscillations present. Locations at the lower bifurcation of the network model were implicated as proper “operating points” for modeling the brain’s resting state activity [17–19], while locations at the high bifurcation can be considered a cartoon model for the global brain dynamics during non-Rapid Eye Movement (non-REM) sleep (see Ref. [20] for a biophysically more detailed model).

In this work we consider a network with instantaneous couplings, although coupling delays in human white matter are finite. The interpretation of results regarding the model’s state space should thus be limited to cortical states whose dynamics is slow compared to these delays. Given that delays are of the order of approx. 5–15 ms [65] brain oscillations with periods of approx. 1 s or less would qualify. These include the so-called slow oscillations, which are the prominent brain rhythm during non-REM sleep and which have already been subject to external perturbation experiments in human neuroscience and clinical settings [66,67].

In lieu of a biologically more detailed model of the effects of control inputs on whole-brain dynamics, delays must be included as soon as brain oscillations with frequencies above a few Hz are of interest. Relative delays can be computed from DTI-based estimates of the length of white matter fiber tracts and included in the coupling terms of Eqs. (1) and (14).

When relative delays are scaled by a global delay constant, the ratio of the coupling delays to the FHN oscillation period can be adapted to match the physiologically observed ratios for any given brain rhythm of interest. The formalism of nonlinear optimal control summarized in Sec. II must be extended, though, to cover finite delays by modifying the coupling term in Eq. (1) to, e.g.,  $\sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x}(t - d)$  for the case of constant delays. The adapted coupling term reappears when computing the adjoint state with the differential equation [Eq. (10)], and thus the iterative algorithm for the numerical optimization (cf. Sec. II C) has to be adapted accordingly in the calculation of step 3.

Here we used a highly simplified model of the global brain dynamics to showcase the applicability of nonlinear optimal control and its added value beyond connectome-based diagnostics derived from linear control theory. When applied to biophysically more realistic models of whole-brain activity (cf. Ref. [20]), nonlinear optimal control may serve as a tool to evaluate the impact of external stimulation and to facilitate the design of new brain stimulation protocols. Noninvasive brain stimulation like transcranial current stimulation [68–70] is a highly promising technique to perturb the global brain dynamics with the goal to improve sensory [71], motor [72], and cognitive abilities [66] of human subjects. Using a cost functional which penalizes spatial sparseness and control energy may—for example—help reducing the required current applied to a subject’s brain as well as focusing the electrical perturbation on the relevant brain areas only. While exact target trajectories are unlikely to be available in a practical setting, state-dependent cost functionals which refer to global quantities [cf. Eq. (4) for the synchronization task] may well be. Particularly, reformulating the control formalism in frequency space will be beneficial here. It would for example allow for computing optimal interventions for changing the

power of certain brain rhythms, which can be monitored by electroencephalography and which are common control targets in clinical settings (cf. Ref. [67]). As shown in other computational and physiological studies [55,73] and emphasized again by our results, the timing of the control inputs may also be crucial for successful interventions. Here we see a high potential of nonlinear optimal control in guiding electrical brain stimulation under electro- or magnetoencephalography [74,75], a setting which allows for such precisely timed control inputs.

For quantitative predictions, an application of nonlinear optimal control to biophysically grounded network models (cf. Ref. [20]) is desirable, where the coupling of neurons to externally applied electric fields is included in a biophysically realistic way. Still, results obtained with simplified brain-network models will strongly facilitate biologically detailed but computationally expensive *in silico* experiments and may already inspire physiological studies on brain-stimulation to explore new control paradigms for clinical stimulation protocols with potentially better control results, reduced energy expenditure, and higher spatial sparseness.

## ACKNOWLEDGMENTS

We thank Dr. Michael Scholz, Cristiana Dimulescu, and Lena Salfenmoser for scientific discussions, and Prof. Dr. Eckehard Schöll for his comments on the manuscript. This work was funded by the German Research Foundation (Grants No. 163436311–SFB 910, No. 327654276–SFB 1315, and under Germany’s Excellence Strategy–EXC 2002/1 “Science of Intelligence,” No. 390523135). Human connectivity data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; Grant No. 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

## APPENDIX A: MATHEMATICAL FORMULATION OF THE OPTIMAL CONTROL PROBLEM

### 1. Optimization problem: General formulation

In this section the mathematical foundations for controlling the network dynamics are laid. We use nonlinear sparse optimal control. All derivations below are similar to the ones in Refs. [29,31], where sparse optimal control was applied to a system of partial differential equations with diffusive coupling.

We consider a network of  $N$  nodes with  $d$ -dimensional dynamics each. Its dynamics is defined by a system of  $N$  stochastic differential equations,

$$\begin{aligned} \frac{d}{dt}\mathbf{x}(t) = & \mathbf{h}[\mathbf{x}(t)] + \sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x}(t) \\ & + (\mathbf{B} \otimes \mathbf{K})\mathbf{u}(t) + \eta(\mathbf{I}_N \otimes \mathbf{D})\boldsymbol{\xi}(t), \end{aligned} \quad (\text{A1})$$

with the state vector  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$  and  $\mathbf{x}_i = (x_{i1}, \dots, x_{id})$ .  $\otimes$  denotes the Kronecker product. The local node dynamics is governed by  $\mathbf{h}(\mathbf{x}) = (\mathbf{h}(\mathbf{x}_1), \dots, \mathbf{h}(\mathbf{x}_N))$  with  $\mathbf{h}(\mathbf{x}_i) = [h_1(\mathbf{x}_i), \dots, h_d(\mathbf{x}_i)]$ . The coupling term consists of the

$N \times N$ -dimensional adjacency matrix  $\mathbf{A}$  and the  $d \times d$ -dimensional local coupling scheme  $\mathbf{G}$ .  $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_N)$  with  $\mathbf{u}_i = (u_{i1}, \dots, u_{id})$  denotes the vector of control inputs,  $\mathbf{B}$  is a diagonal  $N \times N$ -dimensional control matrix, and  $\mathbf{K}$  the  $d \times d$ -dimensional local control scheme. The noise term consists of the Kronecker product of an  $N$ -dimensional identity matrix  $\mathbf{A}$  and the  $d \times d$ -dimensional local noise scheme  $\mathbf{G}$ . The independent white Gaussian stochastic process is given by  $\xi$  with standard deviation  $\eta$ . Boundary conditions are given by

$$\mathbf{x}(t=0) = \mathbf{x}_0. \quad (\text{A2})$$

Since our system is stochastic while the control is deterministic, the optimal control must minimize a cost functional which is an expectation over many realizations of the noise. It is defined as

$$\langle F[\mathbf{x}(\mathbf{u}), \mathbf{u}] \rangle = \int_0^T \langle f[\mathbf{x}(\mathbf{u}), \mathbf{u}] \rangle dt, \quad (\text{A3})$$

where the angle brackets denote the expectation.

Our goal is to find the optimal control  $\bar{\mathbf{u}}$  that minimizes the above mean cost functional. Hence, the variational inequality

$$\langle F[\mathbf{x}(u_i), u_i] \rangle - \langle F[\mathbf{x}(\bar{u}_i), \bar{u}_i] \rangle \geq 0 \quad (\text{A4})$$

holds when regarding every dimension  $i$  separately. We linearize around the optimal solution and write  $\delta u_i = (u_i - \bar{u}_i)$  to obtain

$$\begin{aligned} & \langle F[\mathbf{x}(u_i), u_i] \rangle - \langle F[\mathbf{x}(\bar{u}_i), \bar{u}_i] \rangle \\ & \approx \left. \frac{\partial}{\partial u_i} \langle F[\mathbf{x}(u_i), u_i] \rangle \right|_{\bar{u}_i} \delta u_i \geq 0. \end{aligned} \quad (\text{A5})$$

Since inequality Eq. (A5) has to hold for all  $\delta u_i$  and  $-\delta u_i$ , we find the stronger necessary condition

$$\langle F[\mathbf{x}(u_i), u_i] \rangle - \langle F[\mathbf{x}(\bar{u}_i), \bar{u}_i] \rangle = 0 \quad (\text{A6})$$

for the optimal solution.

Inserting the expression from the right-hand side of Eq. (A3) yields

$$\begin{aligned} & \int_0^T [\langle f[\mathbf{x}(u_i), u_i] \rangle - \langle f[\mathbf{x}(\bar{u}_i), \bar{u}_i] \rangle] dt \\ & \approx \int_0^T \left. \frac{\partial}{\partial u_i} \langle f[\mathbf{x}(u_i), u_i] \rangle \right|_{\bar{u}_i} (u_i - \bar{u}_i) dt \\ & = 0. \end{aligned} \quad (\text{A7})$$

Since Eq. (A7) must hold for all components  $i$ , we write

$$\int_0^T \nabla_{\mathbf{u}} \langle f[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \rangle \circ (\mathbf{u} - \bar{\mathbf{u}}) dt = 0, \quad (\text{A8})$$

where  $\circ$  denotes the elementwise multiplication (Schur product). Here we assume that the cost functional can be separated into a part  $F^u(\mathbf{u})$ , which depends on control inputs only and which is deterministic, and a part  $\langle F^x[\mathbf{x}(\mathbf{u})] \rangle$ , which is a function of the network state  $\mathbf{x}$  and only indirectly depends on the applied control, i.e.,

$$\langle F[\mathbf{x}(\mathbf{u}), \mathbf{u}] \rangle = F^u(\mathbf{u}) + \langle F^x[\mathbf{x}(\mathbf{u})] \rangle. \quad (\text{A9})$$

Equivalently, one can write

$$\int_0^T \langle f[\mathbf{x}(\mathbf{u}), \mathbf{u}] \rangle dt = \int_0^T [f^u(\mathbf{u}) + \langle f^x[\mathbf{x}(\mathbf{u})] \rangle] dt. \quad (\text{A10})$$

Applying the chain rule in Eq. (A8) yields

$$\begin{aligned} & \int_0^T \nabla_{\mathbf{u}} \langle f[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \rangle \circ (\mathbf{u} - \bar{\mathbf{u}}) dt \\ & = \int_0^T [\nabla_{\mathbf{u}} f^u(\bar{\mathbf{u}}) + \langle D_{\mathbf{u}}^T[\mathbf{x}(\bar{\mathbf{u}})] \nabla_{\mathbf{x}} f^x(\mathbf{x}(\bar{\mathbf{u}})) \rangle] \circ (\mathbf{u} - \bar{\mathbf{u}}) dt \\ & = 0, \end{aligned} \quad (\text{A11})$$

where  $D_{\mathbf{u}}[\mathbf{x}(\bar{\mathbf{u}})]$  is the Jacobian matrix of the state  $\mathbf{x}$  with respect to the control vector  $\mathbf{u}$ , evaluated at the optimal control  $\bar{\mathbf{u}}$ .

To calculate the above derivative we must find an expression for the Jacobian matrix  $D_{\mathbf{u}}[\mathbf{x}(\bar{\mathbf{u}})]$ . However, the dependence of the network state on the control inputs cannot be provided in closed form. The Lagrange Formalism enables us to derive an alternative expression for the second term on the right-hand side of Eq. (A11), which can be evaluated to find the optimal control.

## 2. The method of Lagrange multipliers

We can use the formalism of Lagrange multipliers to formulate conditions for the optimization of the cost functional, Eq. (A3), that is subject to the constraint given in Eq. (A1) and the boundary conditions given in Eq. (A2). Our approach is similar to the method used in Ref. [68], and we will obtain a result that resembles what is found in Ref. [31] for a system reaction-diffusion equations.

We define the mean Lagrange function  $\langle \mathcal{L}[\mathbf{x}(t), \mathbf{u}(t)] \rangle$  as

$$\begin{aligned} & \langle \mathcal{L}[\mathbf{x}(t), \mathbf{u}(t)] \rangle \\ & = \left\langle F(\mathbf{x}, \mathbf{u}) - \int_0^T \left[ \frac{d}{dt} \mathbf{x} - \mathbf{h}(\mathbf{x}) - \sigma(\mathbf{A} \otimes \mathbf{G}) \mathbf{x} \right. \right. \\ & \quad \left. \left. - (\mathbf{B} \otimes \mathbf{K}) \mathbf{u} - (\mathbf{I}_N \otimes \mathbf{G}) \xi \right]^T \boldsymbol{\phi}(t) dt \right\rangle \\ & = \left\langle \int_0^T \left\{ f(\mathbf{x}, \mathbf{u}) - \left[ \frac{d}{dt} \mathbf{x} - \mathbf{h}(\mathbf{x}) - \sigma(\mathbf{A} \otimes \mathbf{G}) \mathbf{x} \right. \right. \right. \\ & \quad \left. \left. - (\mathbf{B} \otimes \mathbf{K}) \mathbf{u} - (\mathbf{I}_N \otimes \mathbf{G}) \xi \right]^T \boldsymbol{\phi}(t) \right\} dt \right\rangle. \end{aligned} \quad (\text{A12})$$

Equation (A1) for the dynamics of the network state  $\mathbf{x}$  has to be satisfied for all times  $0 \leq t \leq T$ , which is ensured by the time integration of the constraint.  $\boldsymbol{\phi}(t) = [\boldsymbol{\phi}_1(t), \dots, \boldsymbol{\phi}_N(t)]$ , with  $\boldsymbol{\phi}_i(t) = [\phi_{i1}(t), \dots, \phi_{id}(t)]$  is the  $d \times N$ -dimensional vector of time-dependent Lagrange multipliers.

Linearizing around the optimal solution [cf. Eqs. (A4) and (A8)], we obtain the optimality conditions

$$\left\langle \int_0^T \nabla_{\mathbf{u}} \mathcal{L}[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \circ \delta \mathbf{u} dt \right\rangle = 0 \quad (\text{A13})$$

and

$$\left\langle \int_0^T \nabla_{\mathbf{x}} \mathcal{L}[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \circ \delta \mathbf{x} dt \right\rangle = 0, \quad (\text{A14})$$

with  $\delta \mathbf{u} = \mathbf{u} - \bar{\mathbf{u}}$  and  $\delta \mathbf{x} = \mathbf{x}(\mathbf{u}) - \mathbf{x}(\bar{\mathbf{u}})$ .

After inserting the Lagrange function, Eq. (A12), into Eq. (A14), we apply partial integration and the chain rule, and obtain

$$\begin{aligned}
 & \langle \nabla_x \mathcal{L}[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \circ \delta \mathbf{x} \rangle \\
 &= \left\langle \int_0^T \nabla_x f[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \circ \delta \mathbf{x} dt \right. \\
 &\quad - \int_0^T \nabla_x \left[ \frac{d}{dt} \mathbf{x}(\bar{\mathbf{u}}) \right]^T \phi \circ \delta \mathbf{x} dt \\
 &\quad - \int_0^T \nabla_x \{ (-\mathbf{h}[\mathbf{x}(\bar{\mathbf{u}})] - \sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x}(\bar{\mathbf{u}}) \\
 &\quad - (\mathbf{B} \otimes \mathbf{K})\bar{\mathbf{u}} - (\mathbf{I}_N \otimes \mathbf{G})\xi \}^T \phi \circ \delta \mathbf{x} dt \left. \right\rangle \\
 &= \left\langle \int_0^T \left\{ [\nabla_x f^x(\mathbf{x})] + \frac{\partial}{\partial t} \phi \right. \right. \\
 &\quad \left. \left. + [D_x(\mathbf{h}) + \sigma(\mathbf{A} \otimes \mathbf{G})]^T \phi \right\} \circ \delta \mathbf{x} dt - \phi(T) \circ \delta \mathbf{x}(T) \right\rangle \\
 &= \mathbf{0}. \tag{A15}
 \end{aligned}$$

The boundary term

$$\phi(T) \circ \delta \mathbf{x}(T) = \mathbf{0} \tag{A16}$$

must hold for every realization. Since  $\delta \mathbf{x}(T)$  could be finite, the boundary condition  $\phi(T)$  for the vector of Lagrange multipliers is

$$\phi(T) = \mathbf{0}. \tag{A17}$$

Equation (A15) is fulfilled if the integrand vanishes. We thus obtain a  $d \times N$ -dimensional system of linear ordinary differential equations for the so-called adjoint states  $\phi$ ,

$$-\frac{d}{dt} \phi(t) = [D_x(\mathbf{h}) + \sigma(\mathbf{A} \otimes \mathbf{G})]^T \phi(t) + \nabla_x f^x(\mathbf{x}), \tag{A18}$$

where  $D_x(\mathbf{h})$  denotes the Jacobian matrix of the local dynamics with respect to the network state  $\mathbf{x}$ . The adjoint states are obtained for every realization by solving Eq. (A18) backwards in time obeying the boundary conditions Eq. (A17).

Inserting the Lagrange function [Eq. (A12)] into Eq. (A13), we obtain

$$\begin{aligned}
 & \langle \nabla_u \mathcal{L}[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \circ \delta \mathbf{u} \rangle \\
 &= \left\langle \int_0^T \nabla_u f[\mathbf{x}(\bar{\mathbf{u}}), \bar{\mathbf{u}}] \circ \delta \mathbf{u} dt \right. \\
 &\quad - \int_0^T \frac{d}{dt} \{ D_u[\mathbf{x}(\bar{\mathbf{u}})] \}^T \phi \circ \delta \mathbf{u} dt \\
 &\quad - \int_0^T \nabla_u \{ \mathbf{h}[\mathbf{x}(\bar{\mathbf{u}})] - \sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x}(\bar{\mathbf{u}}) \\
 &\quad - (\mathbf{B} \otimes \mathbf{K})\bar{\mathbf{u}} \}^T \phi \circ \delta \mathbf{u} dt \left. \right\rangle \\
 &= \mathbf{0}. \tag{A19}
 \end{aligned}$$

The first term in Eq. (A19) vanishes [see Eq. (A8)]. Furthermore, after applying partial integration and the chain rule,

above Eq. (A19) simplifies to

$$\begin{aligned}
 \langle \nabla_u \mathcal{L} \circ \delta \mathbf{u} \rangle &= \left\langle \int_0^T \{ D_u[\mathbf{x}(\bar{\mathbf{u}})] \}^T \frac{d}{dt} \phi \circ \delta \mathbf{u} dt \right. \\
 &\quad + \int_0^T \{ [D_u[\mathbf{x}(\bar{\mathbf{u}})] \}^T (\nabla_x \{ \mathbf{h}[\mathbf{x}(\bar{\mathbf{u}})] \\
 &\quad + \sigma(\mathbf{A} \otimes \mathbf{G}) \}^T \phi) + (\mathbf{B} \otimes \mathbf{K})^T \phi \} \circ \delta \mathbf{u} dt \left. \right\rangle \\
 &= \left\langle \int_0^T \{ [D_u[\mathbf{x}(\bar{\mathbf{u}})] \}^T \left[ \frac{d}{dt} \phi + D_x(\mathbf{h}) \right. \right. \\
 &\quad \left. \left. + \sigma(\mathbf{A} \otimes \mathbf{G}) \right]^T \phi \right. \\
 &\quad \left. + (\mathbf{B} \otimes \mathbf{K})^T \phi \right\} \circ \delta \mathbf{u} dt \left. \right\rangle \\
 &= \mathbf{0}, \tag{A20}
 \end{aligned}$$

where the boundary term of the partial integration was set to zero because of the boundary conditions imposed on  $\phi$  and  $\mathbf{x}$ .

Inserting the differential Eq. (A18) that governs the dynamics of the adjoint states into Eq. (A20), we obtain

$$\left\langle \int_0^T \{ -D_u^T[\mathbf{x}(\bar{\mathbf{u}})] \nabla_x f^x[\mathbf{x}(\bar{\mathbf{u}})] + (\mathbf{B} \otimes \mathbf{K})^T \phi \} \circ \delta \mathbf{u} dt \right\rangle = \mathbf{0}. \tag{A21}$$

The first term corresponds to the expression found in Eq. (A11). We can solve Eq. (A21) for this term and insert it into Eq. (A11) to obtain the optimality condition

$$\int_0^T [\nabla_u f^u(\bar{\mathbf{u}}) + (\mathbf{B} \otimes \mathbf{K})^T \langle \phi \rangle] \circ \delta \mathbf{u} dt = \int_0^T \mathbf{g}(t) dt = \mathbf{0}. \tag{A22}$$

Equation (A22) is fulfilled if the stronger optimality condition

$$\mathbf{g}(t) = \nabla_u f^u(\bar{\mathbf{u}}) + (\mathbf{B} \otimes \mathbf{K})^T \langle \phi \rangle = \mathbf{0} \tag{A23}$$

holds for all times  $t \in [0, T]$ . This equation is similar to the equation derived by Casas *et al.* [31] for a system of reaction-diffusion equations.

## APPENDIX B: NUMERICAL INTEGRATION METHOD AND LOCAL OPTIMA

We integrate the equations for the network dynamics [Eq. (1)] and the adjoint state [Eq. (10)] using the fourth order Runge-Kutta (RK4) method. A small step size of  $\Delta t = 0.1$  arb. units is used to ensure numeric stability and small truncation errors. The noise term is scaled with the factor  $\frac{1}{\sqrt{\Delta t}}$ , such that the effective noise strength does not depend on the integration step size.

When solving the minimization problem for the optimal control numerically (see Sec. II C) using gradient based optimization, we may converge to a local minimum of the cost functional. To alleviate this problem, we choose multiple initial conditions  $\mathbf{u}_0$  for the conjugate gradient algorithm. In the case of synchronizing the network dynamics, the control

scheme was robust to changes in  $\mathbf{u}_0$ . The coexistence of multiple local minima of the cost functional was observed only in the case of switching between network states using sparse control (cf. Fig. 7). For larger values for  $I_s$  [cf. Eq. (7)] we find one minimum at  $\mathbf{u}(t) = \mathbf{0}$  (smaller cost for the sparsity term) coexisting with another minimum corresponding to a control input  $\mathbf{u}(t)$  affecting a small number of nodes (smaller cost for the precision term). Which of these local optima has minimal cost depends on the value of  $I_s$ .

To ensure that the global optimum was found in the latter case for a given value of  $I_s$ , we systematically varied the initial conditions. Starting with  $I_s = 0$ , we first calculated the optimal control signal, then increased the value of  $I_s$ , and used the previous optimal control as the new initial condition (*continuation*). Likewise, starting with a high enough value for  $I_s$ , this process is repeated, but for decreasing values of  $I_s$ . For a given value of  $I_s$  the control input with the lower cost was chosen for further analysis.

**APPENDIX C: DEFINITIONS OF AVERAGE AND MODAL CONTROLLABILITY**

We consider a simplified linear network dynamics of the form

$$\mathbf{x}(t + 1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{C1}$$

with activity vector  $\mathbf{x}$ , adjacency matrix  $\mathbf{A}$ , control matrix  $\mathbf{B}$ , and control inputs  $\mathbf{u}$ . To calculate the average controllability of node  $k$  in the network, we choose the input matrix  $\mathbf{B}^{(k)}$  to target a single node, i.e.,  $B_{ij}^{(k)} = 1$  for  $i, j = k$  and  $B_{ij}^{(k)} = 0$ , otherwise. The controllability Gramian of this system is defined as

$$\mathbf{W}^{(k)} = \sum_{\tau=0}^{\infty} \mathbf{A}^{\tau} \mathbf{B}^{(k)} \mathbf{B}^{(k)T} (\mathbf{A}^{\tau})^T \tag{C2}$$

and the average controllability as its trace (cf. Ref. [23])

$$c_k^{av} = \text{Tr} \mathbf{W}^{(k)}. \tag{C3}$$

Network nodes with high average controllability have a large impact on the network dynamics. This makes them important control sites since by controlling these nodes a large number of possible activity vectors  $\mathbf{x}$  can be reached with little control energy. They are consequently said to be able to move the system to many *easy-to-reach states* [23].

The modal controllability is defined based on an eigenvalue decomposition of the network adjacency matrix  $\mathbf{A}$ ,

$$c_k^{\text{mod}} = \sum_{j=1}^N (1 - \lambda_j^2) v_{kj}^2, \tag{C4}$$

where  $v_{kj}$  are the elements of the eigenvector matrix  $\mathbf{V} = [v_{kj}]$  and  $\lambda_1 \dots \lambda_N$  are the corresponding eigenvalues [23].

Nodes with high modal controllability tend to be sparsely connected. They are capable of pushing the network dynamics toward activity vectors  $\mathbf{x}$  which require substantial control energy to be reached. Thus, nodes with high modal controllability are said to be capable of steering the network into *difficult-to-reach states* [23].

For a detailed description and derivation of these regional controllability measures please refer to Refs. [23,25,63]. We use the matlab code provided with the publication of Ref. [23] to compute the controllability measures [79]. For their calculation, only the adjacency matrix  $\mathbf{A}$  (cf. Fig. 2) is required.

**APPENDIX D: DIFFUSION TENSOR IMAGING**

**1. Subjects**

We use the Diffusion Tensor Imaging (DTI) data of 12 human subjects from the Human Connectome Project (Young Adults HCP) [56] with the following IDs: 101309, 121416, 211215, 211619, 212116, 213522, 219231, 220721, 268749, 284646, 303119, 329844. All subjects are male, healthy, and 26–30 years old.

**2. Data processing pipeline**

Data acquisition and preprocessing is described in Ref. [56]. We used BEDPOSTX (Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques) from the FSL toolbox [57] for building up distributions on diffusion parameters, which automatically determines the number of crossing fibres per voxel. Based on these diffusion parameters at each voxel, we then applied the PROBTRACKX (probabilistic tractography) algorithm from the same toolbox with 5000 samples per voxel. The resulting connectivity matrix  $A^s$  for each subject  $s$  was normalized by dividing the connections between any two regions by the number of voxels in the source region multiplied by 5000. Self connections were deleted,  $A_{ii}^s \stackrel{\text{def}}{=} 0, \forall i$ , and we averaged the structural connectivity matrices of the individual subjects to obtain the adjacency matrix  $A = \frac{1}{12} \sum_{s=1}^{N=12} A^s$ . Probabilistic fiber tracking does not provide information about directionality; hence the resulting graph must be undirected. Following standard procedures (cf. Ref. [76]), the structural connectivity matrix was symmetrized ( $A \leftarrow \frac{A+A^T}{2}$ ), as any deviations from symmetry are artifacts of the processing method. Since application of the PROBTRACKX algorithm typically results in a certain number of false positive fibers and thus assigns nonzero connection strengths to all pairs of brain regions, we enforced a sparsity of 20% [77] by discarding all connections with a relative connection strength smaller than 0.00071 (0.15% of strongest relative connection strength).

[1] T. Zaehle, S. Rach, and C. Herrmann, Transcranial alternating current stimulation enhances individual alpha activity in human EEG, *PLoS One* **5**, e13766 (2010).

[2] T. Neuling, S. Rach, and C. Herrmann, Orchestrating neuronal networks: sustained after-effects of transcranial alternating current stimulation depend upon brain states, *Front. Hum. Neurosci.* **7**, 161 (2013).



- [3] R. Polanía, M. A. Nitsche, C. Korman, G. Batsikadze, and W. Paulus, The importance of timing in segregated theta phase-coupling for cognitive performance, *Curr. Biol.* **22**, 1314 (2012).
- [4] D. Strüber, S. Rach, S. Trautmann-Lengsfeld, A. Engel, and C. Herrmann, Antiphase 40 Hz oscillatory current stimulation affects bistable motion perception, *Brain Topogr.* **27**, 158 (2014).
- [5] J. Ladenbauer, J. Ladenbauer, N. Külzow, R. de Boer, E. Avramova, U. Grittner, and A. Flöel, Promoting sleep oscillations and their functional coupling by transcranial stimulation enhances memory consolidation in mild cognitive impairment, *J. Neurosci.* **37**, 7111 (2017).
- [6] W. Sun, W. Fu, W. Mao, D. Wang, and Y. Wang, Low-frequency repetitive transcranial magnetic stimulation for the treatment of refractory partial epilepsy, *Clin. EEG Neurosci.* **42**, 40 (2011).
- [7] F. Fregni, P. Otachi, A. Valle, P. Boggio, G. Thut, S. Rigonatti, A. Pascual-Leone, and K. Valente, A randomized clinical trial of repetitive transcranial magnetic stimulation in patients with refractory epilepsy, *Ann. Neurol.* **60**, 447 (2006).
- [8] A. Hasan, P. Falkai, and T. Wobrock, Transcranial brain stimulation in schizophrenia: targeting cortical excitability, connectivity and plasticity, *Curr. Med. Chem.* **20**, 405 (2013).
- [9] J. Cole, C. Bernacki, A. Helmer, N. Pinninti, and J. O’rardon, Efficacy of transcranial magnetic stimulation (TMS) in the treatment of schizophrenia: A review of the literature to date, *Innov. Clin. Neurosci.* **12**, 12 (2015).
- [10] I. Fried, Brain stimulation in Alzheimer’s disease, *J. Alzheimer’s Dis.* **54**, 789 (2016).
- [11] M. S. George, J. J. Taylor, and E. B. Short, The expanding evidence base for rTMS treatment of depression, *Curr. Opin. Psych.* **26**, 13 (2013).
- [12] D. S. Bassett and O. Sporns, Network neuroscience, *Nat. Neurosci.* **20**, 353 (2017).
- [13] K. Amunts, M. Hawrylycz, D. Van Essen, J. Van Horn, N. Harel, J.-B. Poline, F. De Martino, J. Bjaalie, G. Dehaene-Lambertz, S. Dehaene, P. Valdes-Sosa, B. Thirion, K. Zilles, S. Hill, M. Abrams, P. Tass, W. Vanduffel, A. Evans, and S. Eickhoff, Interoperable atlases of the human brain, *Neuroimage* **99**, 525 (2014).
- [14] E. T. Rolls, M. Joliot, and N. Tzourio-Mazoyer, Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas, *Neuroimage* **122**, 1 (2015).
- [15] S. B. Eickhoff, B. T. Yeo, and S. Genon, Imaging-based parcellations of the human brain, *Nat. Rev. Neurosci.* **19**, 672 (2018).
- [16] M. Breakspear, Dynamic models of large-scale brain activity, *Nat. Neurosci.* **20**, 340 (2017).
- [17] G. Deco and V. K. Jirsa, Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors, *J. Neurosci.* **32**, 3366 (2012).
- [18] G. Deco, A. Ponce-Alvarez, D. Mantini, G. L. Romani, P. Hagmann, and M. Corbetta, Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations, *J. Neurosci.* **33**, 11239 (2013).
- [19] M. Demirtaş, J. B. Burt, M. Helmer, J. L. Ji, B. D. Adkinson, M. F. Glasser, D. C. Van Essen, S. N. Sotiropoulos, A. Anticevic, and J. D. Murray, Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics, *Neuron* **101**, 1181 (2019).
- [20] C. Cakan, C. Dimulescu, L. Khakimova, D. Obst, A. Flöel, and K. Obermayer, A deep sleep model of the human brain: How slow waves emerge due to adaptation and are guided by the connectome, [arXiv:2011.14731](https://arxiv.org/abs/2011.14731) (2020).
- [21] E. Tang and D. S. Bassett, Colloquium: Control of dynamics in brain networks, *Rev. Mod. Phys.* **90**, 031003 (2018).
- [22] S. Gu, R. F. Betzel, M. G. Mattar, M. Cieslak, P. R. Delio, S. T. Grafton, F. Pasqualetti, and D. S. Bassett, Optimal trajectories of brain state transitions, *Neuroimage* **148**, 305 (2017).
- [23] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, B. Y. Alfred, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton *et al.*, Controllability of structural brain networks, *Nat. Commun.* **6**, 8414 (2015).
- [24] M. Golos, V. Jirsa, and E. Dauce, Multistability in large scale models of brain activity, *PLoS Comput. Biol.* **11**, e1004644 (2015).
- [25] S. F. Muldoon, F. Pasqualetti, S. Gu, M. Cieslak, S. T. Grafton, J. M. Vettel, and D. S. Bassett, Stimulation-based control of dynamic brain networks, *PLoS Comput. Biol.* **12**, e1005076 (2016).
- [26] L. D. Berkovitz and N. G. Medhin, *Nonlinear Optimal Control Theory* (CRC Press, Boca Raton, FL, 2012).
- [27] R. Herzog, G. Stadler, and G. Wachsmuth, Directional sparsity in optimal control of partial differential equations, *SIAM J. Control Optim.* **50**, 943 (2012).
- [28] E. Casas, R. Herzog, and G. Wachsmuth, Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations, *ESAIM: Control Optimis. Calc. Var.* **23**, 263 (2017).
- [29] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, Vol. 112 (American Mathematical Society, Philadelphia, PA, 2010).
- [30] W. Stannat and L. Wessels, Deterministic control of stochastic reaction-diffusion equations, *Evol. Equ. Control Theory* (2019).
- [31] E. Casas, C. Ryll, and F. Tröltzsch, Sparse optimal control of the Schlögl and FitzHugh–Nagumo systems, *Comput. Methods Appl. Math.* **13**, 415 (2013).
- [32] R. FitzHugh, Impulses and physiological states in theoretical models of nerve membrane, *Biophys. J.* **1**, 445 (1961).
- [33] J. Nagumo, S. Arimoto, and S. Yoshizawa, An active pulse transmission line simulating nerve axon, *Proc. IRE* **50**, 2061 (1962).
- [34] T. Kostova, R. Ravindran, and M. Schonbek, FitzHugh–Nagumo revisited: Types of bifurcations, periodical forcing and stability regions by a Lyapunov functional, *Int. J. Bifurcat. Chaos* **14**, 913 (2004).
- [35] E. Schöll, G. Hiller, P. Hövel, and M. A. Dahlem, Time-delayed feedback in neurosystems, *Philos. Trans. R. Soc. A* **367**, 1079 (2009).
- [36] S. Eydam, I. Franović, and M. Wolfrum, Leap-frog patterns in systems of two coupled FitzHugh–Nagumo units, *Phys. Rev. E* **99**, 042207 (2019).
- [37] I. Shepelev and T. Vadivasova, Variety of spatio-temporal regimes in a 2D lattice of coupled bistable FitzHugh–Nagumo oscillators. Formation mechanisms of spiral and double-well chimeras, *Commun. Nonlin. Sci. Numer. Simul.* **79**, 104925 (2019).
- [38] S. A. Plotnikov, J. Lehnert, A. L. Fradkov, and E. Schöll, Synchronization in heterogeneous FitzHugh–Nagumo networks with hierarchical architecture, *Phys. Rev. E* **94**, 012203 (2016).
- [39] J. Lehnert, T. Dahms, P. Hövel, and E. Schöll, Loss of synchronization in complex neuronal networks with delay, *Europhys. Lett.* **96**, 60013 (2011).

- [40] C. Cakan, J. Lehnert, and E. Schöll, Heterogeneous delays in neural networks, *Eur. Phys. J. B* **87**, 54 (2014).
- [41] D. Nikitin, I. Omelchenko, A. Zakharova, M. Avetyan, A. L. Fradkov, and E. Schöll, Complex partial synchronization patterns in networks of delay-coupled neurons, *Philos. Trans. R. Soc., A* **377**, 20180128 (2019).
- [42] I. Omelchenko, T. Hülser, A. Zakharova, and E. Schöll, Control of chimera states in multilayer networks, *Front. Appl. Math. Stat.* **4**, 67 (2019).
- [43] G. Ruzzene, I. Omelchenko, J. Sawicki, A. Zakharova, E. Schöll, and R. G. Andrzejak, Remote pacemaker control of chimera states in multilayer networks of neurons, *Phys. Rev. E* **102**, 052216 (2020).
- [44] T. Chouzouris, I. Omelchenko, A. Zakharova, J. Hlinka, P. Jiruska, and E. Schöll, Chimera states in brain networks: Empirical neural vs. modular fractal connectivity, *Chaos* **28**, 045112 (2018).
- [45] M. Gerster, R. Berner, J. Sawicki, A. Zakharova, A. Škoch, J. Hlinka, K. Lehnertz, and E. Schöll, FitzHugh–Nagumo oscillators on complex networks mimic epileptic-seizure-related synchronization phenomena, *Chaos* **30**, 123130 (2020).
- [46] L. Ramlow, J. Sawicki, A. Zakharova, J. Hlinka, J. C. Claussen, and E. Schöll, Partial synchronization in empirical brain networks as a model for unihemispheric sleep, *Europhys. Lett.* **126**, 50007 (2019).
- [47] A. Ghosh, Y. Rho, A. McIntosh, R. Kötter, and V. Jirsa, Cortical network dynamics with time delays reveals functional connectivity in the resting brain, *Cogn. Neurodynam.* **2**, 115 (2008).
- [48] V. Vuksanović and P. Hövel, Dynamic changes in network synchrony reveal resting-state functional networks, *Chaos* **25**, 023116 (2015).
- [49] A. Messé, M.-T. Hütt, P. König, and C. C. Hilgetag, A closer look at the apparent correlation of structural and functional connectivity in excitable neural networks, *Sci. Rep.* **5**, 7870 (2015).
- [50] R. Buchholz, H. Engel, E. Kammann, and F. Tröltzsch, On the optimal control of the Schlögl-model, *Comput. Optimiz. Appl.* **56**, 153 (2013).
- [51] E. Polak and G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique* **3**, 35 (1969).
- [52] J. Nocedal and S. Wright, *Numerical Optimization* (Springer Science & Business Media, Berlin, 2006).
- [53] R. Fletcher and C. M. Reeves, Function minimization by conjugate gradients, *Comput. J.* **7**, 149 (1964).
- [54] E. M. Izhikevich, *Dynamical Systems in Neuroscience* (MIT Press, Cambridge, MA, 2007).
- [55] C. Cakan and K. Obermayer, Biophysically grounded mean-field models of neural populations under electrical stimulation, *PLoS Comput. Biol.* **16**, e1007822 (2020).
- [56] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium *et al.*, The WU-Minn human connectome project: An overview, *Neuroimage* **80**, 62 (2013).
- [57] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, FSL, *Neuroimage* **62**, 782 (2012).
- [58] Noise-induced transitions in physics, chemistry, and biology, in *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology* (Springer, Berlin, 1984), pp. 164–200.
- [59] S. Schmidt, M. Scholz, K. Obermayer, and S. A. Brandt, Patterned brain stimulation, what a framework with rhythmic and noisy components might tell us about recovery maximization, *Front. Hum. Neurosci.* **7**, 325 (2013).
- [60] C. Honey, O. Sporns, L. Cammoun, X. Gigandet, J.-P. Thiran, R. Meuli, and P. Hagmann, Predicting human resting-state functional connectivity from structural connectivity, *Proc. Natl. Acad. Sci. USA* **106**, 2035 (2009).
- [61] R. F. Galán, On how network architecture determines the dominant patterns of spontaneous neural activity, *PLoS One* **3**, e2148 (2008).
- [62] A. Hamdan and A. Nayfeh, Measures of modal controllability and observability for first-and second-order linear systems, *J. Guid. Control Dynam.* **12**, 421 (1989).
- [63] E. Tang, H. Ju, G. L. Baum, D. R. Roalf, T. D. Satterthwaite, F. Pasqualetti, and D. S. Bassett, Control of brain network dynamics across diverse scales of space and time, *Phys. Rev. E* **101**, 062301 (2020).
- [64] A. Mazzoni, H. Lindén, H. Cuntz, A. Lansner, S. Panzeri, and G. T. Einevoll, Computing the local field potential (LFP) from integrate-and-fire network models, *PLoS Comput. Biol.* **11**, e1004584 (2015).
- [65] R. Caminiti, F. Carducci, C. Piervincenzi, A. Battaglia-Mayer, G. Confalone, F. Visco-Comandini, P. Pantano, and G. M. Innocenti, Diameter, length, speed, and conduction delay of callosal axons in macaque monkeys and humans: Comparing data from histology and magnetic resonance imaging diffusion tractography, *J. Neurosci.* **33**, 14501 (2013).
- [66] L. Marshall, M. Mölle, M. Hallschmid, and J. Born, Transcranial direct current stimulation during sleep improves declarative memory, *J. Neurosci.* **24**, 9985 (2004).
- [67] J. Ladenbauer, N. Külzow, S. Passmann, D. Antonenko, U. Grittner, S. Tamm, and A. Flöel, Brain stimulation during an afternoon nap boosts slow oscillatory activity and memory consolidation in older adults, *Neuroimage* **142**, 311 (2016).
- [68] M. A. Nitsche and W. Paulus, Transcranial direct current stimulation—update 2011, *Restor. Neurol. Neurosci.* **29**, 463 (2011).
- [69] A. Antal, K. Boros, C. Poreisz, L. Chaieb, D. Terney, and W. Paulus, Comparatively weak after-effects of transcranial alternating current stimulation (tACS) on cortical excitability in humans, *Brain Stimul.* **1**, 97 (2008).
- [70] D. Terney, L. Chaieb, V. Moliadze, A. Antal, and W. Paulus, Increasing human brain excitability by transcranial high-frequency random noise stimulation, *J. Neurosci.* **28**, 14147 (2008).
- [71] J. R. Behrens, A. Kraft, K. Irlbacher, H. Gerhardt, M. C. Olma, and S. A. Brandt, Long-lasting enhancement of visual perception with repetitive noninvasive transcranial direct current stimulation, *Front. Cell. Neurosci.* **11**, 238 (2017).
- [72] M. Moisa, R. Polania, M. Grueschow, and C. C. Ruff, Brain network mechanisms underlying motor enhancement by transcranial entrainment of gamma oscillations, *J. Neurosci.* **36**, 12053 (2016).
- [73] S. Alagapan, S. L. Schmidt, J. Lefebvre, E. Hadar, H. W. Shin, and F. Fröhlich, Modulation of cortical oscillations by low-frequency direct cortical stimulation is state-dependent, *PLoS Biol.* **14**, e1002424 (2016).
- [74] G. Thut, T. O. Bergmann, F. Fröhlich, S. R. Soekadar, J.-S. Brittain, A. Valero-Cabré, A. T. Sack, C. Miniussi, A. Antal,

- H. R. Siebner *et al.*, Guiding transcranial brain stimulation by EEG/MEG to interact with ongoing brain activity and associated functions: A position paper, *Clin. Neurophysiol.* **128**, 843 (2017).
- [75] T. O. Bergmann, Brain state-dependent brain stimulation, *Front. Psychol.* **9**, 2108 (2018).
- [76] G. Deco, J. Cabral, M. W. Woolrich, A. B. Stevner, T. J. Van Hartevelt, and M. L. Kringelbach, Single or multiple frequency generators in on-going brain activity: A mechanistic whole-brain model of empirical meg data, *Neuroimage* **152**, 538 (2017).
- [77] G. Gong, P. Rosa-Neto, F. Carbonell, Z. J. Chen, Y. He, and A. C. Evans, Age-and gender-related differences in the cortical anatomical network, *J. Neurosci.* **29**, 15684 (2009).
- [78] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.104.024213> for additional information on the state space exploration of the whole-brain network (limit cycle criterion, aperiodic behavior in mixed states, influence of noise on the network dynamics), the optimal control of the brain network dynamics (recovery of laLC target state, computation of critical times  $t_c$ , controlled nodes in case of sparse control, videos of the node dynamics in phase space with optimal control), and linear network control theory (correlation of average and modal controllability with weighted degree, modal controllability evaluated for the synchronization task in the noise free case, controllability measures for the synchronization task for noisy network dynamics).
- [79] <https://complexsystemsupenn.com/cod> edata.