



A Computational Framework for Controlling the Self-Restorative Brain Based on the Free Energy and Degeneracy Principles

Hae-Jeong Park^{1,2,3,4,5*} and Jiyoung Kang^{1,2*}

¹ Center for Systems and Translational Brain Science, Institute of Human Complexity and Systems Science, Yonsei University, Seoul, South Korea, ² Department of Nuclear Medicine, Yonsei University College of Medicine, Seoul, South Korea,

³ Department of Psychiatry, Yonsei University College of Medicine, Seoul, South Korea, ⁴ Brain Korea 21 Project, Graduate School of Medical Science, Yonsei University College of Medicine, Seoul, South Korea, ⁵ Department of Cognitive Science, Yonsei University, Seoul, South Korea

OPEN ACCESS

Edited by:

Tatyana Sharpee,
Salk Institute for Biological Studies,
United States

Reviewed by:

Fei He,
Coventry University, United Kingdom
Sungho Hong,
Okinawa Institute of Science and
Technology Graduate
University, Japan
Adeel Razi,
Monash University, Australia

*Correspondence:

Hae-Jeong Park
parkhj@yonsei.ac.kr
Jiyoung Kang
jykang@yuhs.ac

[†]These authors have contributed
equally to this work

Received: 31 July 2020

Accepted: 18 March 2021

Published: 14 April 2021

Citation:

Park H-J and Kang J (2021) A
Computational Framework for
Controlling the Self-Restorative Brain
Based on the Free Energy and
Degeneracy Principles.
Front. Comput. Neurosci. 15:590019.
doi: 10.3389/fncom.2021.590019

The brain is a non-linear dynamical system with a self-restoration process, which protects itself from external damage but is often a bottleneck for clinical treatment. To treat the brain to induce the desired functionality, formulation of a self-restoration process is necessary for optimal brain control. This study proposes a computational model for the brain's self-restoration process following the free-energy and degeneracy principles. Based on this model, a computational framework for brain control is established. We posited that the pre-treatment brain circuit has long been configured in response to the environmental (the other neural populations') demands on the circuit. Since the demands persist even after treatment, the treated circuit's response to the demand may gradually approximate the pre-treatment functionality. In this framework, an energy landscape of regional activities, estimated from resting-state endogenous activities by a pairwise maximum entropy model, is used to represent the pre-treatment functionality. The approximation of the pre-treatment functionality occurs via reconfiguration of interactions among neural populations within the treated circuit. To establish the current framework's construct validity, we conducted various simulations. The simulations suggested that brain control should include the self-restoration process, without which the treatment was not optimal. We also presented simulations for optimizing repetitive treatments and optimal timing of the treatment. These results suggest a plausibility of the current framework in controlling the non-linear dynamical brain with a self-restoration process.

Keywords: free energy principle, resting state, brain dynamics, energy landscape, self-restoration, maximum entropy model, degeneracy

INTRODUCTION

The goal of clinical treatment for the brain is to modify the brain circuit to yield a desirable brain function. Since the brain is a complex non-linear dynamic system, clinical treatment can be considered a control problem. For the clinical treatment to the human brain, various methods have been developed, such as thermal ablation with the high intensity focused ultrasound (Park et al., 2017), deep brain stimulation (DBS) (Park et al., 2015), vagus nerve stimulation (Yu et al., 2018),

and transcranial magnetic stimulation (TMS) (Park et al., 2013; Kar, 2019) as well as conventional medications and traditional surgical operations (Schreglmann et al., 2018). Despite the remarkable advancement of these therapeutic techniques, the optimal control of the brain by treatments has many practical challenges due to the complexity of the brain and ethical issues. Since various experiments are not allowed in the human brain, establishing optimal control procedures for the brain is very slow and limited. Therefore, optimal brain control remains mostly theoretical and based on computational models.

The brain control studies with computational models have been conducted mainly in two approaches: characterization of the network controllability and prediction of the network behavior based on the state dynamic equation of the brain. To characterize the brain network, the controllability of a brain system has been evaluated in the graph-theoretic perspective (Liu et al., 2011; Gu et al., 2017; Tang et al., 2017; Cornblath et al., 2019; Lee et al., 2019; Stiso et al., 2019; Karrer et al., 2020). This approach optimizes input signals to increase or decrease activity at some brain network nodes to induce the desired brain activity at all the brain nodes. The other approach is to predict a brain system's behavior by altering some nodes or edges of the brain. In this approach, the optimal control is determined by evaluating the outcome after removing nodes or edges or changing the system's parameters in the virtual brain model (Falcon et al., 2016; Jirsa et al., 2017; Proix et al., 2017; An et al., 2019; Olmi et al., 2019). Those two types of computational approaches on brain control have primarily focused on the immediate changes in the brain network's activity or function. Those studies, however, did not consider the fact that the brain has self-restorative plasticity, making the system resilient to external treatments or perturbations.

Self-restoration capacity in the brain has been found after damage or stress via neural, molecular, and hormonal mechanisms (Russo et al., 2012; King, 2016; Murrugh and Russo, 2019). At the macroscopic level, the self-restoration process toward the initial functionality has widely been reported in clinical neuroscience, for e.g., functional recovery after stroke (Murphy and Corbett, 2009; Malone and Felling, 2020), recovery of the language capacity (Saur et al., 2006), recovery of the vision after surgery (Mikellidou et al., 2019). This self-restorative property of the brain is advantageous in protecting the brain after various external attacks (Glassman, 1987). In terms of clinical treatment, however, this self-restoration process is a bottleneck as it tends to recover the initial abnormal functionality, acting against the aim of any treatment. Examples of this anti-treatment action can be found with neurological or antipsychotic medication (Abbott, 2010) showing drug resistance in most brain disease such as schizophrenia (Potkin et al., 2020), depression (Bennabi et al., 2019), Parkinson's diseases (Vorovenci et al., 2016), and epilepsy (Lee et al., 2017). Goellner et al. (2013) showed that the late seizure recurrence after temporal lobe epilepsy surgery was as much as 48.9%. Despite the anatomical alteration by resection, the treated brains, initially showing free or reduction of abnormal function (seizure behavior) after surgical dissection, returned to the initial state of abnormal functionality in a certain period after treatment.

In this respect, the brain's self-restoration process may well be included as an essential part of the computational model of brain control. Despite the criticality of the brain's self-restoration process, research that has a self-restoration process in the control problem is hardly found. This may be partly attributable to the difficulty in defining the driving force of the self-restoration process and its mathematical formulation. In contrast to the microscopic level, where the mechanism of the self-recovery process has actively been researched in terms of neurogenesis (e.g., Mattson, 2008), a systematic understanding of the self-restoration process at the macroscopic level is still lacking. Are there any principles that we may refer to formulate the brain's self-restoration process at the macroscopic level?

In the current study, as an extension of our previous study (Kang et al., 2021), we propose a computational framework for controlling the self-restorative brain by formulating the driving force of self-restoration based on the free-energy principle (Friston et al., 2006; Friston, 2010) and the degeneracy nature of a non-linear complex system (Glassman, 1987). According to the free-energy principle, the brain acts based on a model established to minimize long-term average surprise from the external environment (Friston et al., 2006; Friston, 2010). The brain network and its behavior can be considered a result of long-term adjustment to meet environmental demands (Park and Friston, 2013). For a neural population at any level of the information hierarchy, neural populations that send signals to and receive signals from the neural population are environment to the neural population. For a neural circuit of the brain, the circuit's environment can involve lower-level and higher-level neural populations, affecting the neural circuit by sensation from the lower-level neural populations and regulation from the higher-level neural populations (Friston, 2008). The long-term demands of the environment to a neural circuit or system can be represented by the statistics of bottom-up (from the lower-level neural populations) and top-down (from the higher-level neural populations) signals. Although any clinical treatment may alter a neural circuit, the circuit's environmental demands persist and do not change rapidly even after alteration in the circuitry. Since the functionality before treatment has been developed as an optimal solution to environmental demands, we posit that the altered neural circuit gradually approximates the pretreated neural circuit's functionality while adjusting itself to meet the environmental demands after treatment.

In this framework, the pre-treatment state the brain circuit tends to recover is not the same circuit, but the functionality that the circuit has established to satisfy the external demands. In most cases, the functionality for ongoing environmental demands has to be approximated via reallocation of the reduced circuit resources after treatment (e.g., after removing a node of the circuit). In this respect, the recovery of functionality via an altered circuit can be referred to as a well-known property of the complex biological system called "degeneracy" (Glassman, 1987; Edelman and Gally, 2001). Degeneracy of a system indicates a function (or behavior) can be implemented with different network configurations (Friston and Price, 2003). Degeneracy of the self-reorganizing biological system is essential to manage and protect its functionality from damage (Marder and Goaillard,

2006; Marder et al., 2017). This study is based on the degeneracy principle as it focuses on the restoration of the pre-treatment functionality, not the pre-treatment circuit.

The final component of the current framework for brain control is to define the pre-treatment functionality of the brain. As a representation of the brain's functionality, we used the non-linear dynamics of regional activity while the brain is at rest (called resting-state). A strong relationship exists between the resting-state connectivity (or distributed patterns of endogenous activity) with task-related brain activation or connectivity (Biswal et al., 1995; Smith et al., 2009; Cole et al., 2014, 2016; Krienen et al., 2014; Park et al., 2014; Yeo et al., 2015; Tavor et al., 2016; Jung et al., 2018). Fox and Raichle (2007) argued that resting-state connectivity might serve as a potential scaffold that supports diverse configurations subserving the functional elements of a given task (Fox and Raichle, 2007). In this respect, a stabilized brain network's resting-state dynamics were considered to summarize the environmental demands established by long-term interactions with the environment.

The resting-state brain network may behave as a non-linear dynamical system with its microstate (defined by distributed regional activity pattern) transitioning over the energy landscape of multistable microstates (attractors) (Freyer et al., 2011, 2012; Rabinovich and Varona, 2011; Deco and Jirsa, 2012; Kelso, 2012; Cabral et al., 2014; Tognoli and Kelso, 2014; Deco et al., 2015; Breakspear, 2017). The non-linear dynamics of the brain circuit can be modeled in terms of non-linear interactions among nodes in the network using a pairwise maximum entropy model (MEM) (Watanabe et al., 2013, 2014a,b,c; Kang et al., 2017, 2019; Ezaki et al., 2018; Gu et al., 2018). From the pairwise MEM, we can infer the probability distribution of each microstate (an activation pattern), the microstates' energy landscape. In the energy landscape, a microstate's energy is the minus (scaled) log of its probability of occurrence. In this model, the microstate dynamics (represented in microstates' energy landscape) are emergent from the underlying complex network (or circuitry). They are considered to represent the gross functionality of a brain.

In summary, our proposal for the recovery process can be explained by the free energy principle to satisfy environmental demands by reconfiguring the remained resources after treatment according to the degeneracy principle of the complex brain. Utilizing this self-restoration model, we could develop a strategy to identify the optimal treatment target region (nodes or edges in the network) and the amount of treatment strength within a source system to be treated (e.g., a disease system) to induce microstate dynamics of the desired goal system (a healthy system). We call this procedure optimal brain control throughout the paper. In the conventional control theory problems, control signals are inputs to the system to achieve the desired system's state without changing the system parameters. Meanwhile, the optimal brain control in this study refers to adjusting the source system's model parameters to approximate the desired functionality of the goal system. In this study based on the pairwise MEM, the model parameters include the sensitivity of a brain region and interaction among brain regions, which indicate neurobiological connectivity or synaptic efficacy that

modulate the input and output relationship, i.e., the functionality of the brain.

We used the term "optimal brain control" in consideration of clinical treatment settings that call for optimal selection of treatment target gray matter regions (nodes) or white matter regions (edges) and treatment strength, within limited access to the brain circuit at a time. For example, the circuit that generates epileptic seizures is the source system, and the goal system is a healthy functioning state without a seizure. The treatment target can be multiple nodes in the medication. For example, lorazepam enhances the effect of the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) receptors distributed in multiple brain areas. The target can be a single node by temporal lobectomy, which removes a part of the anterior temporal lobe. Callosotomy, which dissects interhemispheric fibers, is an example of targeting edges in the network.

The current paper is composed in the following order. It begins with a mathematical description of pairwise MEM and its energy landscape analysis. We then formulated a self-restoration process and optimal control in a non-linear dynamical system. Based on this formulation, we present diverse simulations to illustrate the self-restoration process and show the effect of modeling the self-restoration process in brain control. We also present simulations for optimizing repetitive treatment strategy and its timing in consideration of the clinical practice, where any treatment is highly restricted. Using these simulations, we expected to show the construct validity of the current framework in the brain's control.

BACKGROUND

Dynamic Properties of the Brain Using Pairwise Maximum Entropy Model

To define the dynamic properties of a system, we used the energy landscape of microstate established on pairwise MEM. Here, we briefly explain the pairwise MEM model. The details for deriving the MEM of the resting-state functional magnetic resonance imaging (rsfMRI) can be found elsewhere (Watanabe et al., 2013, 2014a; Kang et al., 2017, 2019).

In the pairwise MEM model of a brain with N regions, a brain state V_k , is defined as an N -dimensional binary vector;

$$V_k = (\sigma_1, \dots, \sigma_N), \quad (1)$$

where $\sigma_i = 1$ for an activated state and $\sigma_i = 0$ for an inactivated state at region i . Thus, totally 2^N states exist. An energy $E(V_k)$ of a state V_k is defined as

$$E(V_k) = - \sum_{i=1}^N H_i \sigma_i(V_k) - \sum_{i=1}^{N-1} \sum_{j=i+1}^N J_{ij} \sigma_i(V_k) \sigma_j(V_k), \quad (2)$$

where H_i and J_{ij} are model parameters that represent weights for independent activation of region i and pairwise interaction (coactivation) between regions i and j , respectively. For simplicity, we used $A = \{H_i, J_{ij}\}_{i=1, \dots, N, j=1, \dots, N}$ to express all the model parameters. These model parameters were estimated using

the maximum likelihood estimation approach. For a detailed mathematical review, see Yeh et al. (2010).

The probability of a state V_k is given by the Boltzmann distribution $p(V_k)$,

$$p(V_k) = \frac{\exp(-E(V_k))}{\sum_{j=1}^{2^N} \exp(-E(V_j))}. \quad (3)$$

To analyze the energy landscape of state dynamics, we defined local minima (attractors, LM) and occupation time ratio of each local minimum ($OCR(LM_i)$) as below.

Local Minimum

A local minimum (LM_i) of an energy landscape of a system with parameters A is a state with lower energy than its neighbor states. Neighbor states are defined as states that differ from each other by only one element (one region) of the activation state.

Occupation Time Ratio

$OCR(LM_i)$: Occupation time ratio of LM_i is the sum of probabilities of all states in the basin region of LM_i . The basin region of LM_i is the set of states that belong to the LM_i . To determine whether a state belongs to the LM_i , each element of the state is gradually changed along the energy gradient until it reaches one of the local minima.

Functional Distance Between Energy Landscapes

To measure the functional distance between the target and source systems, A^t and A^s , in the energy landscape, we defined a distance function between states in terms of dynamic properties (energy landscape) of the two systems, governed by the system's network parameters.

To focus on the functional distance between major attractors and their properties in the optimal treatment, we use a partial Kullback–Leibler (KL)-divergence defined as follows.

$$\mathcal{D}(A^t, A^s) = \sum_{k \in \mathbb{R}} p(V_k | A^t) \ln \frac{p(V_k | A^t)}{p(V_k | A^s)}, \quad (4)$$

where \mathbb{R} represents a set of states that belong to basin regions of major attractors.

We also defined the similarity between two systems in terms of system parameters by the root-mean-square deviation (RMSD) of the two systems' parameter vectors.

$$RMSD(A^t, A^s) = \|A^t - A^s\|. \quad (5)$$

Recovery Process

We modeled the recovery process based on three assumptions: (1) recovery occurs by adjusting the network connectivity (interactions) of the neighbors of the treated node or edge; (2) adjustment of connectivity is performed within a range of its flexibility, and (3) recovery occurs to meet the external demands, which were represented in the state dynamics of the pretreated stabilized system.

The treatment at region m (a node or an edge, for simplicity, we call it "region") is denoted by changing the element A_m^p in the pretreated network parameters A^p with

$$A_m^{tr} \leftarrow A_m^p + \alpha \quad (6)$$

where α is the amount of treatment. The system parameters right after treatment can be expressed as $A^t = \{A_m^{tr}, A_{\setminus m}^p\}$, where m and $\setminus m$ represent the treated and untreated regions, respectively (Figure 2D). m can be multiple nodes or edges. In this theoretical study, we assumed that we know how to achieve the desired level α and achieve the desired parameter A_m^{tr} . The treated state A^t of a system is the starting point of the recovery A_0^r .

The system proceeds with its recovery to minimize the functional distance between state dynamics before treatment and recovery (Figure 2B). When we decompose the network elements (nodes or edges) into recovery regions (strongly connected neighbors of the treated node or edge), \mathbb{R}_m , and unchanged (weakly or unconnected) regions, $\setminus \mathbb{R}_m$, for a treated region (node or edge) m with a treatment strength A_m^{tr} , the network state of the treated system just after treatment, $A^t = \{A_m^{tr}, A_{\setminus m}^p\}$, can be written as $A^t = \{A_{\mathbb{R}_m}^t, A_{\setminus \mathbb{R}_m}^t\}$. The recovery then begins from A^t and the recovery regions \mathbb{R}_m cooperate to find an optimal parameter set $A_{\mathbb{R}_m}^*$ within a constrained bound \mathbb{C} to return to the pre-treatment state A^p . This recovery process can be written as below:

$$A_{\mathbb{R}_m}^* = \arg \min_{A_{\mathbb{R}_m}^{t'}, |A_{\setminus \mathbb{R}_m}^{t'}| \leq \mathbb{C}} \mathcal{D}_r(A^p, A^t | A_{\mathbb{R}_m}^{t'}) \quad (7)$$

$$A^t = \{A_{\mathbb{R}_m}^{t'}, A_{\setminus \mathbb{R}_m}^t\}, A^r = \{A_{\mathbb{R}_m}^*, A_{\setminus \mathbb{R}_m}^t\},$$

where \mathcal{D}_r indicates the distance function between the stabilized state before treatment A^p and a plausible treatment solution A^t by adjusting parameters $A_{\mathbb{R}_m}^{t'}$ in the recovery region \mathbb{R}_m while keeping the other region $\setminus \mathbb{R}_m$ unchanged after treatment A^t . The final recovered state A^r is composed of the optimal parameter set within the recovered regions $A_{\mathbb{R}_m}^*$ and the unchanged regions of the treated system $A_{\setminus \mathbb{R}_m}^t$. Considering the limited capacity of the biological change, we restricted maximum changes at the recovered regions $A_{\mathbb{R}_m}^*$ to be <20% of those of the previous step.

In this study, we define the pre-treatment network parameter A^p as a stable state after a long period of adaptation to the environment. The optimally recovered state A^r can be a new pre-treatment network state A^p for a subsequent treatment, after stabilization, e.g., $A_\infty^r = A^p$, where ∞ indicates a sufficient time for stabilization. Since the treated system may not revert completely to the pre-treatment network state by utilizing the constrained resources of the recovery regions, A^p is a function of trial number or time, moving toward the target system over a very long time scale. It should also be noted that a subsequent treatment can be applied to a system before the system is stabilized. We refer to treatment before stabilization as the treatment at the transient network stage. We considered that the transient state does not satisfy the environment's demands. In this case, the pre-treatment network parameter A^p was not

updated with the state right before treatment, but instead referred to the recent stabilized state. We utilized functional distance \mathcal{D}_r to generate similar dynamics instead of generating similar network parameters for the recovering and initial systems.

The reorganization is performed by modifying the network parameters of neighbors \mathbb{R}_m but within a plausible range of each parameter (connectivity)'s flexibility. In Eq. 7, we denote this with $|A'_{\mathbb{R}_m}| \leq \mathbb{C}$, indicating plausible parameters within a constrained bound \mathbb{C} . The amount of change at each node or edge in the recovery regions can be defined proportionally to the treated system's baseline (pretreated state). For the treated region m , we can assign less flexibility than for neighbors or assign inflexibility (i.e., no change) after treatment during the recovery process.

To measure effects of a treatment on brain dynamics, we define a recovery capacity as the difference between functional distance (partial KL divergence) of the treated state A^t and pretreated state A^p , $D(A^p, A^t)$, and the functional distance between the recovered state A^r and pretreated state A^p , $D(A^p, A^r)$ as follows,

$$\text{Recovery capacity} = \Delta D = D(A^p, A^t) - D(A^p, A^r).$$

Optimal Control

A self-restorative system A^s is decomposed into the region that requires treatment m and the unaffected (untreated) region $\setminus m$ and is represented as $A^s = \{A^s_m, A^s_{\setminus m}\}$. When a target system A^g is given as the goal to achieve for a source system A^s , the optimal treatment is to find a region m and its treatment level A^{t*}_m to minimize the distance function \mathcal{D} between the goal system A^g and recovered system $A^{s' r}$ that develops following the self-restoration process of the source system A^s in response to the treatment. To differentiate this from the distance function \mathcal{D} between A^g and $A^{s' r}$, we use \mathcal{D}^+ to indicate the functional distance \mathcal{D} between A^g and A^s . The optimal control is defined as below,

$$A^{t*}_m = \arg \min \mathcal{D}^+ (A^g, A^s | A^{s' r}_m) \quad (8)$$

$$= \arg \min \mathcal{D} (A^g, A^{s' r} | A^{s' r}_m), \quad (9)$$

$$A^{s' r} = \{A^{s' r}_m, A^s_{\setminus m'}\},$$

$$A^{s' r} = \{A^{s' r}_{\mathbb{R}_{m'}}, A^{s' r}_{\setminus \mathbb{R}_{m'}}\},$$

$$A^t = \{A^{t*}_m, A^s_{\setminus m}\},$$

where the recovered system $A^{s' r}$ is achieved following a self-restoration process after changes in the neighbors $\mathbb{R}_{m'}$ of the treated region m' , according to Equation 7. The optimal control is conducted by searching for the best solution to achieve the goal system's dynamics by adjusting the parameter $A^{s' r}_m$ in the treated region m' while maintaining the other parameters $A^s_{\setminus m'}$ in untreated regions m' unchanged. The final treated system A^t is composed of the optimal treatment region with its strength A^{t*}_m and the unaffected parameters of the treated system $A^s_{\setminus m}$.

Note that the distance function \mathcal{D} is defined in functional space (between energy landscapes), not in parameter space. In

other words, \mathcal{D} indicates a distance between the source dynamics that emerge from the source system with a parameter A^s , and the target dynamics that emerge from the goal system with a parameter A^g . From the perspective of degeneracy, the minimal distance function \mathcal{D} in the dynamics space does not necessarily indicate the closeness in the network parameter space. Even though the two parameter sets, A^g and A^s , are distant in the parameter space, they can be close in the dynamics space.

Strategy for Iterative Optimal Treatment

Optimal treatment is a recursive procedure between treatment planning by the operator and the restoration process in the treated system (Figures 1, 2C). The target region (node and edge) to be treated and the strength of treatment was chosen using a grid search algorithm in this study. In practice, the treatment to the system was performed by altering the MEM parameter A_i (an activity of a region H_i or a pairwise interaction J_{ij}) by an amount of α . We assumed that only neighboring nodes and edges of the treated node participate in the recovery process to return the brain dynamics to the pre-treatment state (Figures 1B,C). The treatment strength induces changes in the energy landscape in a non-linear manner (Figure 2A). When a node is altered (i.e., H_i is changed), edges that are strongly connected with the node (Figure 1B) undergo self-restoration steps, gradually changing the energy landscape (Figure 2B). When an edge is treated (i.e., J_{ij} is selected for treatment), two nodes that are connected with the treated edge and the strongly connected edges of the two nodes undergo self-restoration (Figure 1C). A threshold ($|J_{ij}| \geq 0.1$) was used to determine strongly connected edges. If we applied treatments multiple times, the energy landscape evolved as the iteration of treatment and restoration (Figure 2C).

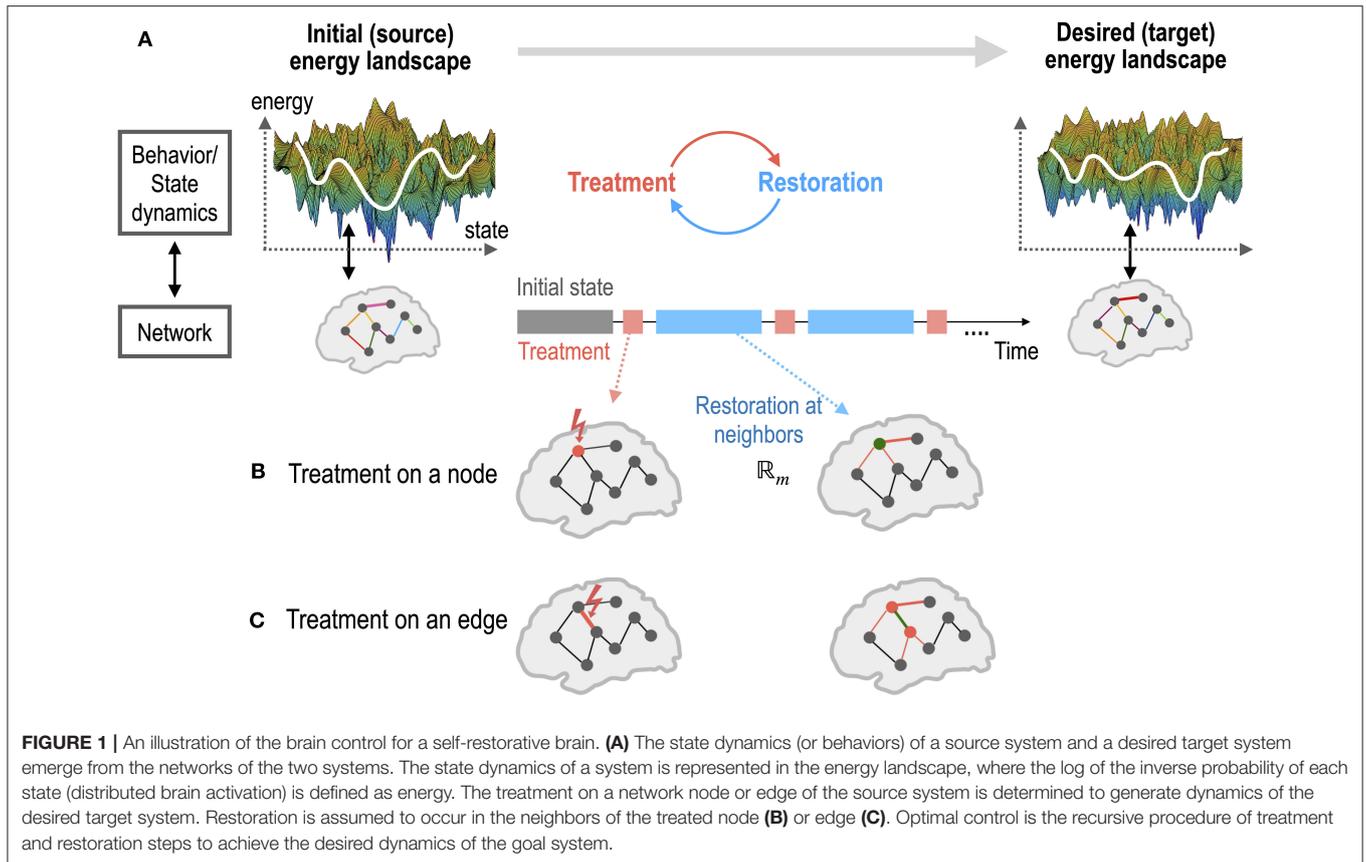
The restoration process is an optimization procedure with reference to the stable pre-treatment state (Figure 2B) as shown in Equation 7. When the restoration process is saturated (no significant improvement in minimizing the functional distance by changing parameters), the saturated network becomes a new pre-treatment state for a new restoration step, i.e., $A_{\infty}^r = A^p$ (Figure 2C).

To implement the recovery process in Equation 7, we adopted the gradient ascent method, which is generally used to estimate the pairwise MEM model parameters from the experimental data by maximizing the log-likelihood (Watanabe et al., 2013, 2014a; Kang et al., 2017, 2019). To maximize the log-likelihood, model parameters, H_i and J_{ij} , are updated iteratively according to differences between data-driven and model-driven expectations of activations and coactivations, as shown below.

$$H_i(t+1) \leftarrow H_i(t) + \alpha_g (\log \langle \sigma_i \rangle - \log \langle \sigma_i \rangle_A), \quad (10)$$

$$J_{ij}(t+1) \leftarrow J_{ij}(t) + \alpha_g (\log \langle \sigma_i \sigma_j \rangle - \log \langle \sigma_i \sigma_j \rangle_A), \quad (11)$$

where α_g is a learning rate, $\langle \sigma_i \rangle$ and $\langle \sigma_i \sigma_j \rangle$ are expectations of activations and coactivations of the brain regions evaluated using the empirical data. From the pairwise MEM parameter A , probability $p(V_k|A)$ for each state V_k can be derived using Equation 3, based on which the expected activations and coactivations of the brain regions are derived using the



following equation.

$$\langle \sigma_i \rangle_A = \sum_{k=1}^{2^N} \sigma_i(V_k) p(V_k|A), \quad (12)$$

$$\langle \sigma_i \sigma_j \rangle_A = \sum_{k=1}^{2^N} \sigma_i \sigma_j(V_k) p(V_k|A). \quad (13)$$

If the dynamics of a pretreated system A^p follows the pairwise MEM (we assumed in this study), the expected activations $\langle \sigma_i \rangle$ and coactivations $\langle \sigma_i \sigma_j \rangle$ of sufficiently large samples from the stabilized system A^p , equal to the model-driven expectations of the activation $\langle \sigma_i \rangle_{A^p}$ and coactivation $\langle \sigma_i \sigma_j \rangle_{A^p}$ of the brain regions. Then, a recovery process can be explained as follows:

$$H_i(t+1) \leftarrow H_i(t) + \alpha_g (\log \langle \sigma_i \rangle_{A^p} - \log \langle \sigma_i \rangle_{A^r}), \quad (14)$$

$$J_{ij}(t+1) \leftarrow J_{ij}(t) + \alpha_g (\log \langle \sigma_i \sigma_j \rangle_{A^p} - \log \langle \sigma_i \sigma_j \rangle_{A^r}), \quad (15)$$

$$A^r = \{H_i, J_{ij} \mid i=1, \dots, N, j=1, \dots, N\}$$

The recovery is proceeded by adjusting the network parameters of neighbors ($A_{\mathbb{R}_m}^t$) of the treated node or edge m .

In general, a treatment is applied to a system when the recovery process is saturated for a sufficient time after each treatment. However, we also presented a simulation of treatment in the transient state before full saturation. We denote the

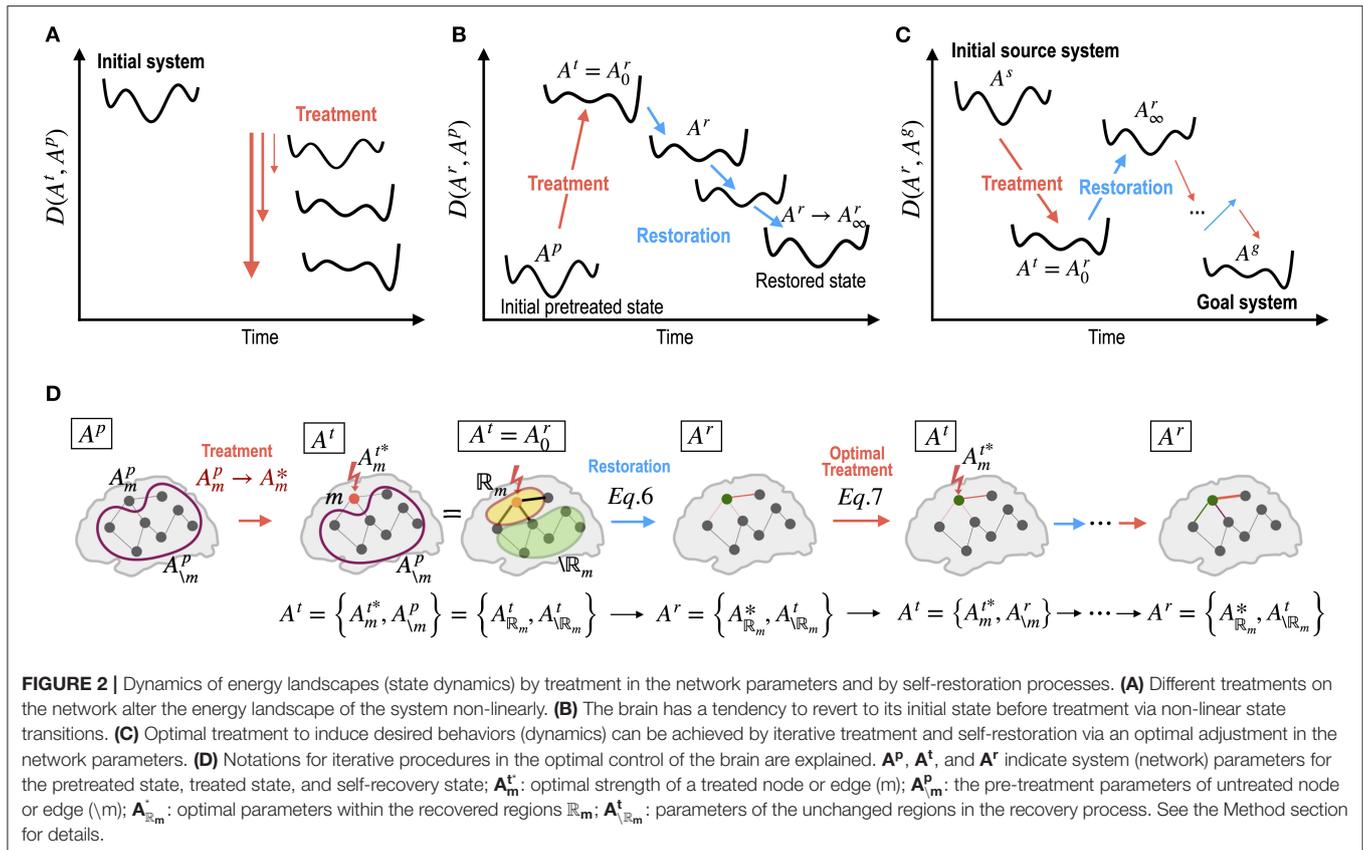
transition state as the proportion of time relative to the time for full recovery. In this case, we used a grid search method to determine the optimal treatment time without waiting for full recovery, as well as the target region and its scale. In this situation, we denote the functional distance \mathcal{D}^+ as a function of treatment timing T .

$$A_m^{t*} = \arg \min \mathcal{D}^+ (A^g, A^{s'} | A_m^{s'}, T) \quad (16)$$

MATERIALS AND METHODS

A Test System: the Subcortical Limbic Brain

As a test system for optimal control, we reused the MEM for the subcortical human brain (Kang et al., 2017). Briefly, the system consists of 15 subcortical regions of interests (ROIs): the hippocampus (HIP), amygdala (AMY), caudate (CAU), putamen (PUT), pallidum (PAL), thalamus (THL), nucleus accumbens (NACC) of the left (L), and right (R) hemispheres, and the brainstem (BSTEM). The MEM parameters were estimated from the resting-state fMRI data of 470 participants in the HCP database (Van Essen et al., 2012). The estimated parameters are presented in **Figures 3A,B**. The subcortical-limbic system has highly symmetric interactions across hemispheres and appears to be modular. We sorted local minima (LMs) with their occupation time ratios and selected



the five top local minima that have the highest occupation time ratios. Among 18 LMs, the five major LMs, i.e., LM1(8000), LM2(24769), LM3(32768), LM4(1), and LM5(25286), occupied 83.5% of all possible states (Figure 3E). The brain activation patterns for the two major local minima are displayed in Figure 3F. In this study, we set this system as a goal of the control for the virtual abnormal systems. By controlling the regional activity parameter H_i and pairwise interaction parameter J_{ij} (node and edge of the MEM parameters) of a virtual abnormal system, the abnormal system is expected to be guided to have dynamics of this goal system.

Overview of Simulations

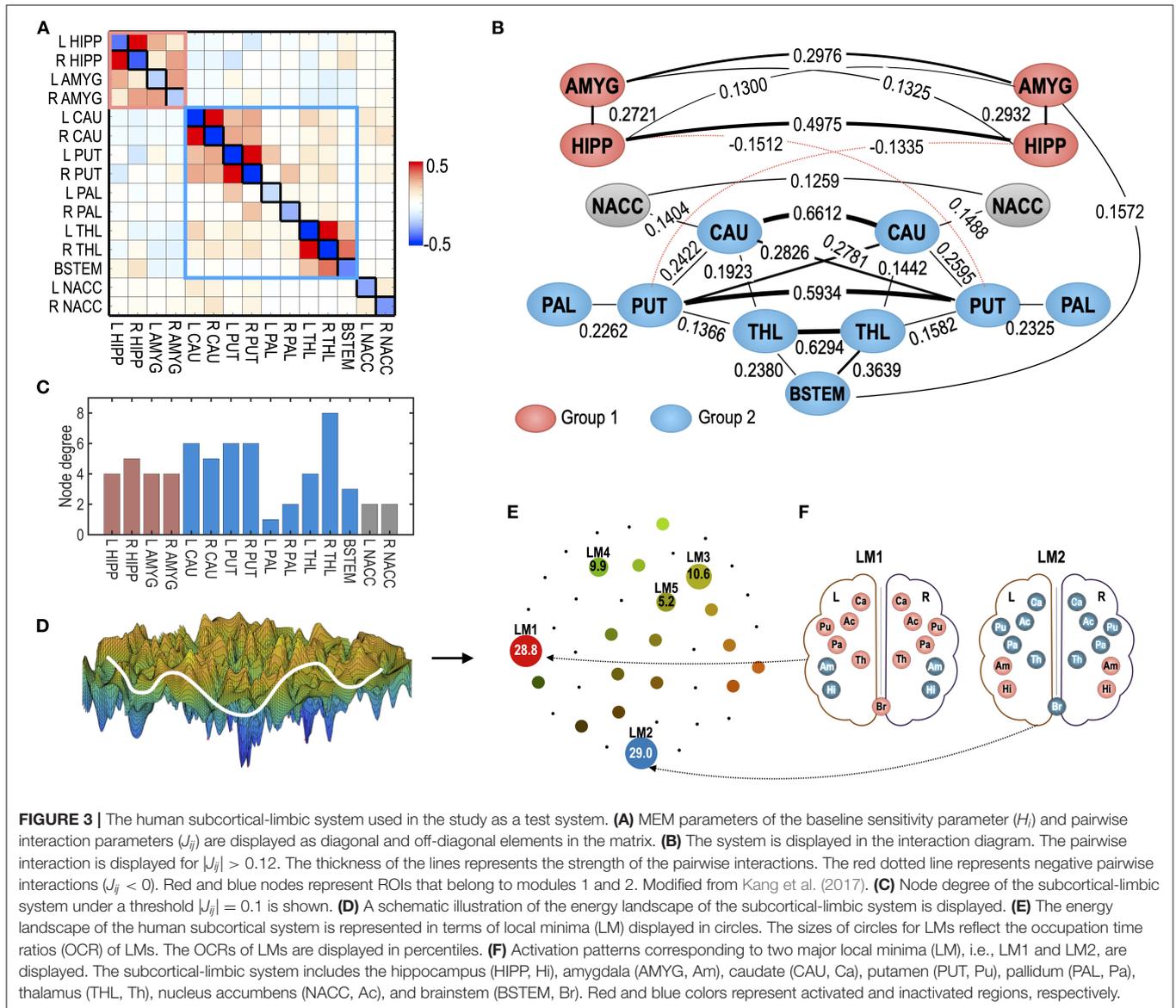
We conducted six simulation experiments to show the construct validity of the proposed framework. In simulation 1, we present an example to show the self-restoration process in the energy landscape of the brain network after treatment (or damage) in a region. In simulation 2, we present the effects of treatment on nodes and edges according to the number of neighbors to show the advantage of more neighboring edges in the restoration process. In simulation 3, we show the need for a self-restoration process in the control model by comparing treatment effects with and without considering the system's self-restoration. In simulation 4, we show the effects of repetitive treatment on each node of a source system to induce the desired dynamics. In simulation 5, we further control the timing of subsequent treatments before full recovery when treating a system. In simulation 6, we optimize the dissection of interhemispheric

connectivity to simulate a corpus callosotomy for epilepsy surgery. All of these simulations were conducted to show the effects of the self-restoration process and how to treat the system to achieve the desired behaviors.

Simulation 1: The System's Self-Restoration Process After Treatment

To illustrate the self-restoration process, we presented a perturbation simulation of the right thalamus (R THL) by adding 0.5 to its H_i parameter. After this perturbation, the neighboring edges connected to the node were gradually reconfigured to generate state dynamics similar to those of the initial pre-treatment state (Figure 4A). For a treatment that induced a deviation of state dynamics from the pre-treatment state, the self-restoration procedure gradually moved the system toward the pre-treatment state, which induced a shorter functional distance \mathcal{D} (partial KL-divergence) between the recovering and pre-treatment states (Figure 4B), and receded its parameters from those of the pre-treatment state (Figures 4B,C). Despite the increasing distance in the parameter space (RMSD, Figure 4C), the distance in the state dynamics from the pre-treatment state is reduced (\mathcal{D} , Figure 4B). This is an example of degeneracy, which refers to the phenomenon where similar behaviors can be formulated using different networks.

Treatment of the right thalamus changed the energy landscape significantly from that of the pre-treatment state (Figure 4D); the OCR of LM1 increased from 28.8 to 40.0, and the OCR of



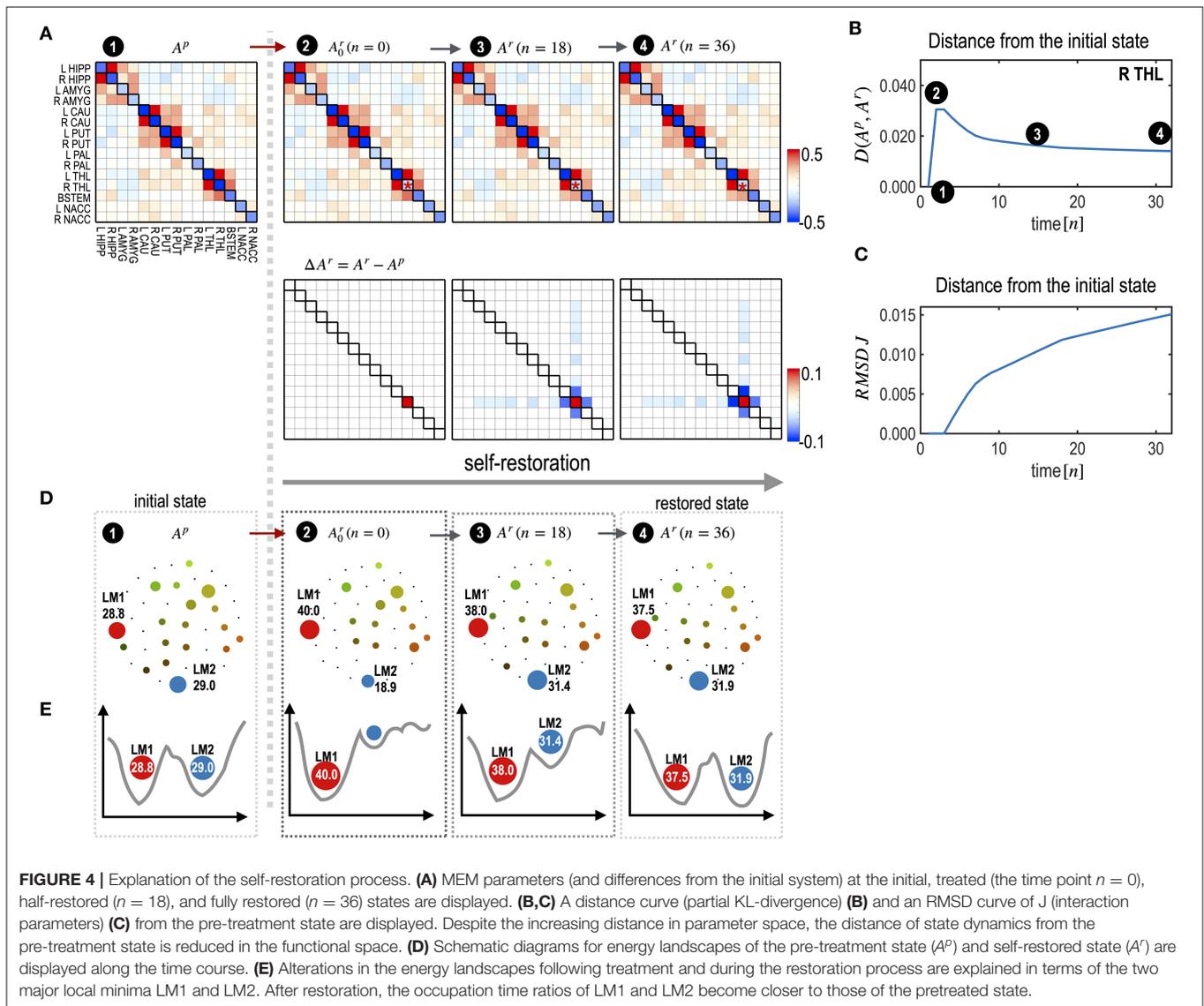
LM2 decreased from 29.02 to 18.9. During self-restoration, this asymmetric energy landscape (between LM1 and LM2) gradually recovered (Figure 4D). In the final stage of the self-restoration, the OCRs of LM1 and LM2 were 37.5 and 31.9, similar to those of the initial system. Full recovery was not achieved in this system as it utilized only the limited resources of neighboring edges. Figure 4E shows the changes in the major two local minima (LM1 and LM4) along with the treatment and transient states in the recovery process. This suggests that the system recovers similar energy landscapes after recovery.

Simulation 2: Region-Specific Self-Restoration Capacity

To test the node- or edge-specific recovery capacity, we evaluated the self-restoration process after treating each node

(Figure 5) and each edge (Figure 6) one by one. The degree of freedom was defined by the number of neighboring edges that participated in the self-restoration process. In the node's treatment, the neighboring edges were strongly (threshold $(|J_{ij}| \geq 0.1)$) connected with the treated node (Figure 1B). In the treatment of an edge, neighboring nodes connected to the treated edge and strongly connected edges connected to these neighboring nodes were considered to participate in the self-restoration process (Figure 1C).

Figure 5 presents the node-specific restoration process in the real subcortical-limbic system. Nodes with diverse node degrees (Figure 5B) have different recovery capacities (Figure 5A) and recovery curves (Figure 5C). The finally recovered network parameters and energy landscapes differed from each other (Figures 5C,D). The recovery capacity depended highly on the number of neighbors (or node degrees); nodes with more



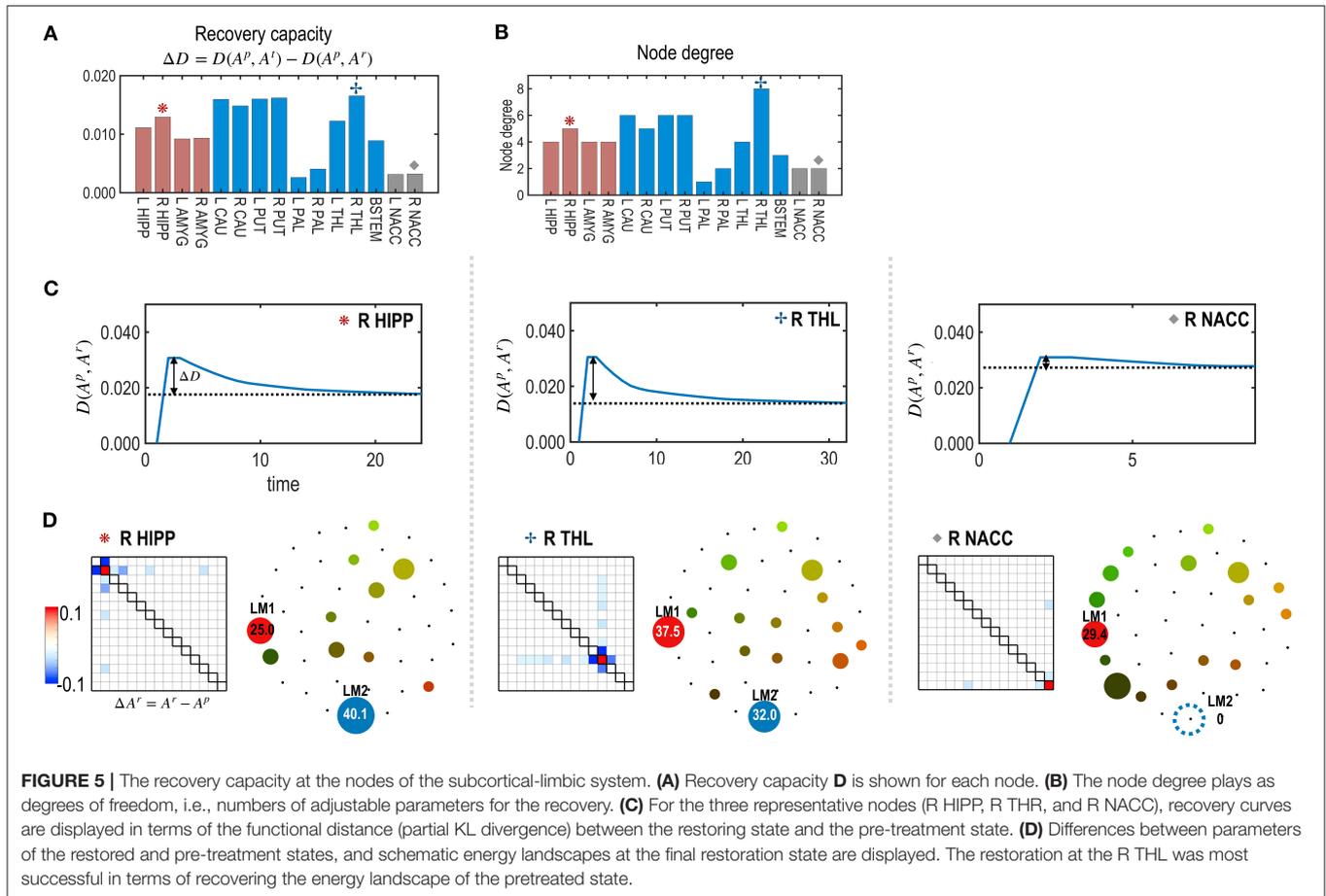
neighbors had a higher recovery capacity. In this case, the node degree acted as a degree of freedom of the system. This was also found in the treatment of edges shown in **Figure 6**. In this case, the number of neighboring edges of the system again explained the recovery capacity. The edges with a higher number of neighboring edges (size of neighbors \mathbb{R}_m in **Figure 6B**) showed better restoration (**Figure 6C**). For example, greater self-restoration occurred after treatment in the edge between R CAU and L THL compared to the edge between L PAL and L NACC (**Figure 6D**). The latter utilizes adjustments of more edges than the former one in the self-restoration (**Figure 6E**).

Simulation 3. Effects of the Self-Restoration Process in Controlling the Brain System

We simulated treatments with and without considering the self-restorative properties of the system. In this study, we generated a

virtual system by adding a Gaussian random noise $\sim N(0, 0.1)$ to the parameters of the human subcortical-limbic system presented in **Figure 3**. We considered the virtual system as a source system and the human subcortical system as a goal system (**Figure 7A**).

For each node, the best strength of treatment ($|\alpha|$) was identified using a grid search method among a set of α , 0.05, 0.10, 0.15, ..., and 0.5. We selected the best treatment strength that minimizes the functional distance between the final restoration state and goal state based on Equations 8, 9. **Figure 7B** shows the expected treatment effects without considering the restoration process, while **Figure 7C** shows the actual treatment results for the self-restorative system. Discrepancies between the expected and treatment effects occurred when the self-restoration process was not considered. In contrast, when the recovery process was considered, we obtained increased treatment effects (**Figure 7D**), with functional distances from the goal system shorter than that of treatments without considering the system's recovery process (**Figure 7C**). Optimal nodes differed according to how the nodes



were chosen with or without a self-restorative model. When we determine the optimal node and its treatment strength to treat without considering the recovery process, an optimal treatment target was chosen in the left putamen (L PUT) with a treatment strength of 0.45 (Figure 7B) but the optimal treatment did not effectively change the system to the desired goal after restoration (Figures 7C,E). When we consider the restoration's effects, the best treatment was selected on BSTEM with -0.5 (Figure 7D). For this treatment, functional distance from the desired state decreased right after treatment, followed by an increase during the restoration process (Figure 7G). The final treatment effects by considering the restorative process are better in this treatment than the treatment without considering restoration (Figure 7E).

Simulation 4. Repetitive Treatment at a Single Node

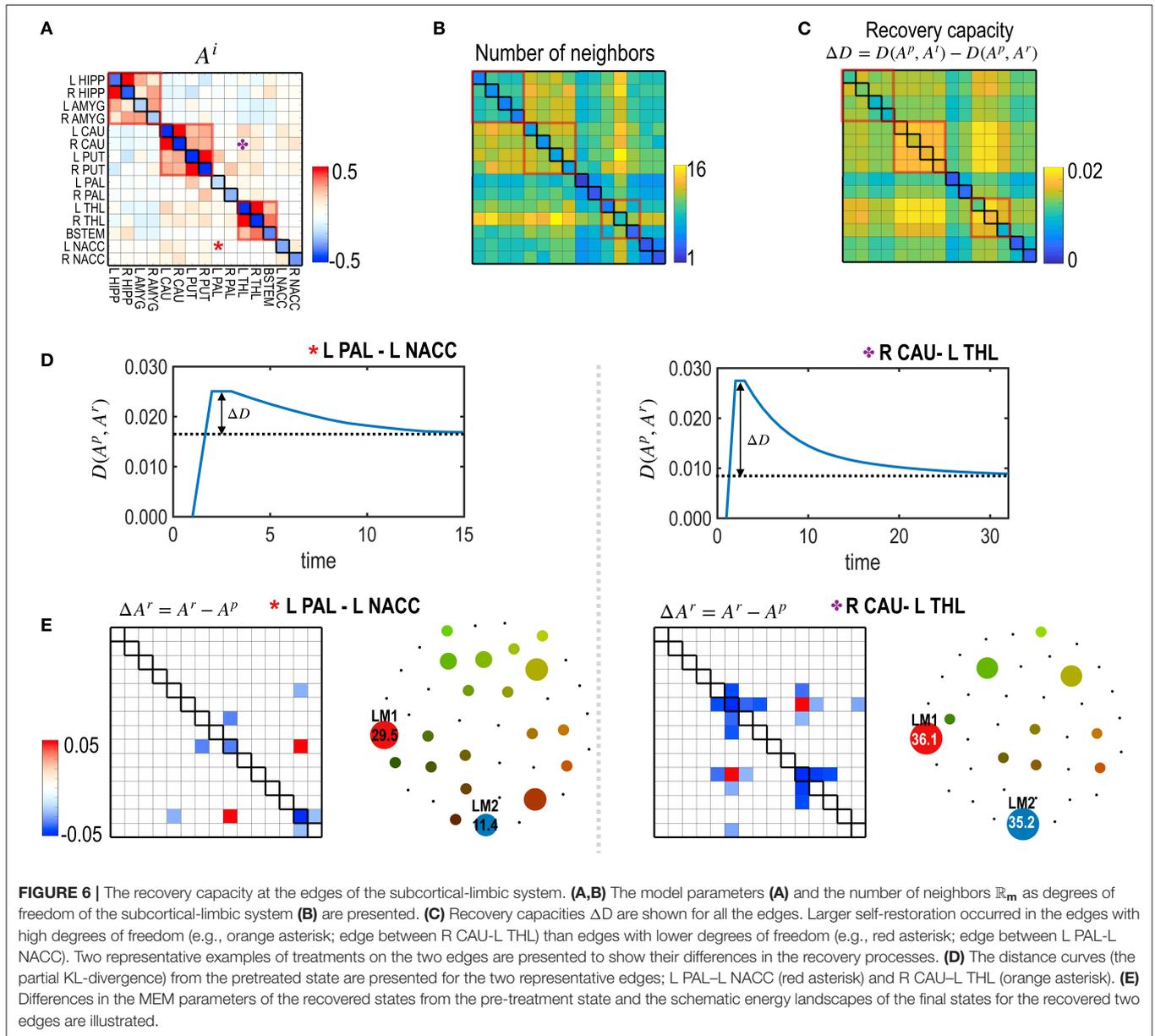
In the clinic, most treatments are repetitive, particularly concerning medications. We simulated repetitive treatments without changing the treated node. In the repetitive treatment simulation, the subsequent treatment was applied after the effects of the previous treatment had become saturated (i.e., reached an equilibrium state). Figure 8 shows the results of the repetitive treatment at each node. Compared to the single treatment shown

in Figure 7D, repetitive treatments with a sequence of different strengths (Figures 8C,F) made the system closer to the desired goal (Figure 8B). After repetitive treatment, the treated system's final energy landscape gets closer to that of the target goal system (Figure 8H).

In Figure 8E, the state change due to the second treatment suggests the non-linear nature of the treatment vs. the behavioral response (dynamics). The optimal treatment was not always chosen to minimize the distance to the goal system in the early stages of the time curves, as shown in the first treatment effect (Figure 8E). Instead, the optimal treatment made the system deviate from the desired goal system but eventually get closer to the desired system than other treatments that are initially effective but finally ineffective (Figure 8E).

Simulation 5. Repetitive Treatment at a Single Node With Flexible Timing

Most previous studies did not consider the timing of the treatment under the dynamically responding brain. When a treatment is applied, the brain gradually recovers and transitions to an equilibrium state. Considering the restoration process, one may apply the subsequent treatment at transient states without waiting for the equilibrium state. We simulated

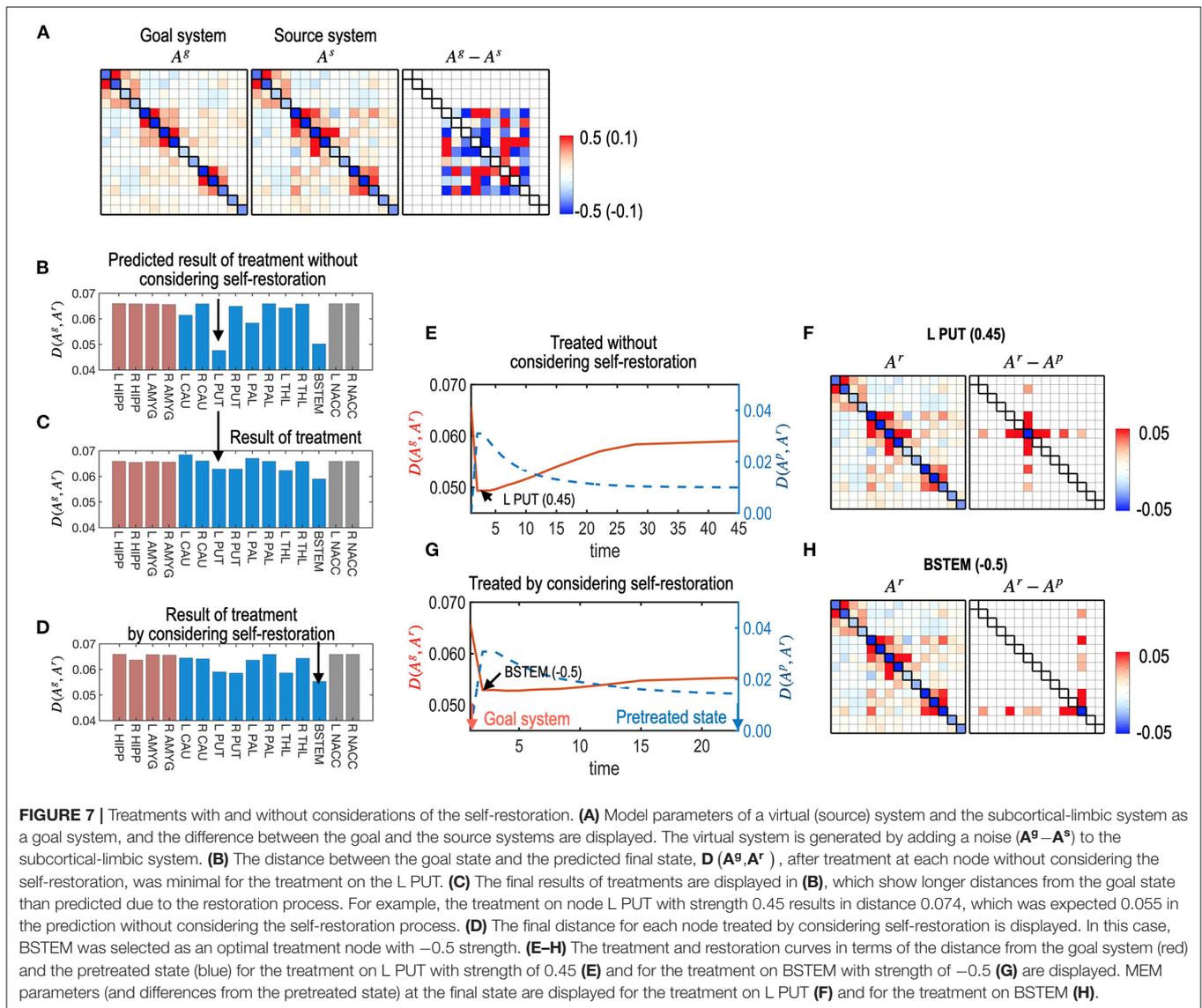


optimal repetitive treatment by optimizing the timing of subsequent treatments. In this simulation, we used the same simulation setting as simulation 3, except for the flexible timing of the treatment. Compared to simulation 4, we explored the best strength of the treatment and the best timing of subsequent treatments for each node. As shown in **Figure 9**, the treatment with the best timing increased the treatment effect compared to treatment after full recovery for each treatment (**Figure 8**).

Simulation 6. Optimal Removal of Edges

We simulated dissections of interhemispheric connections to imitate a corpus callosotomy for epilepsy surgery. We generated an abnormal brain that has stronger connections between the left and right hemispheres (**Figure 10A**). This was

performed by increasing positive inter-hemispheric connectivity and decreasing the negative interhemispheric connectivity of the subcortical-limbic system by adding a Gaussian noise $\sim N(0, 0.1)$ according to the polarity of the initial connectivity. We tested the optimal treatment strategy for different numbers and targets of interhemispheric edges (one, two, and three) to be dissected (**Figures 10B,C**). For each number of edges (49 single interhemispheric edges, i.e., left $7 \times$ right 7 , 2,352 combinations for two edges, and 110,544 combinations for three edges), we evaluated the best edges to remove. For the source brain, dissection of LAMYG-RCAU (one edge); LAMYG-RCAU and LAMYG-RPAL (two edges); LAMYG-RCAU, LAMYG-RPAL, and LHIPP-RCAU (three edges) were the best combinations to decrease the functional distance from the goal system (**Figure 10D**).

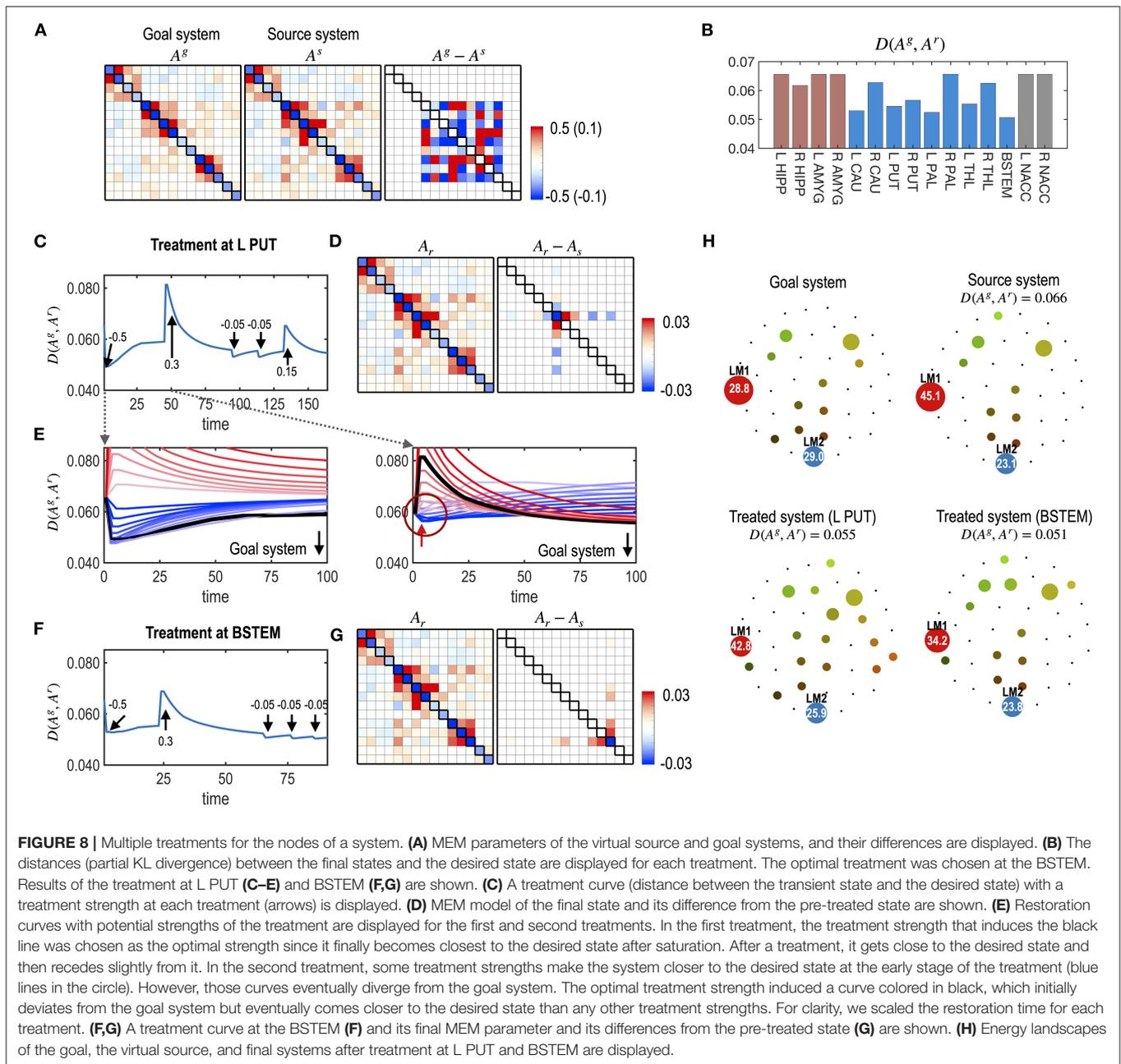


DISCUSSION

Although brain control has garnered increasing interest, brain control research has mainly been conducted based on theoretical and computational models as the practical control of the brain has many challenges due to the complexity of the brain and ethical issues. Several computational models to control the brain network have been proposed to characterize the graph-theoretic properties of the system (Tang et al., 2017; Lee et al., 2019; Stiso et al., 2019; Karrer et al., 2020) or a purpose of predicting outcomes after treatment (Falcon et al., 2016; Jirsa et al., 2017; Proix et al., 2017; An et al., 2019; Olmi et al., 2019). These previous studies have assumed the brain as a dynamic system, immediately responding to the incoming treatment. However, the system's self-restoration process after the cessation of the treatment has not been fully considered, without which the brain control may not be optimal. Compared to the brain circuit's

various dynamic state equations, the formulation of the self-restoration process has been rarely researched.

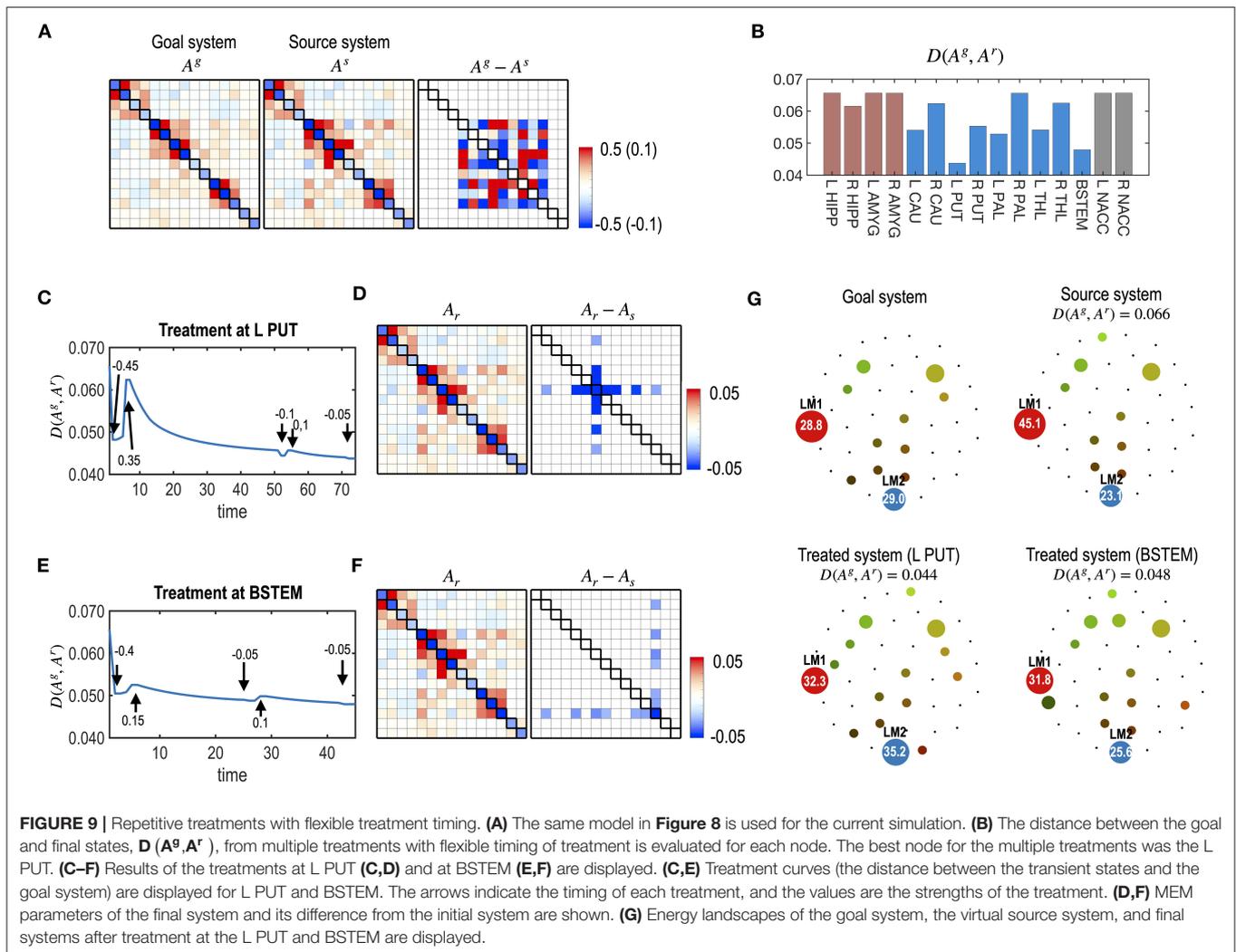
To account for the effect of the restoration process on brain control, we propose a formulation of the brain's recovery process that drives the system to perform the function before treatment. The driving force of this self-restoration process is based on the free-energy principle (Friston et al., 2006; Friston, 2010) over a non-linear complex system, with degeneracy in terms of generating the same behaviors from diverse network configurations. According to the free-energy principle (Friston et al., 2006; Friston, 2010), the network is configured to respond or predict the environment's statistical demands, making the system energy-efficient. As long as the environment's statistics do not change, the treated or partially lesioned system may well-adjust its remaining subnetwork (neighbors of the treated node) to satisfy those demands. Since the altered (treated) node cannot participate in the organized work of the



subnetwork at the same performance level as a pre-treatment state, the system tries to compensate for the role of the altered node by reorganizing interactions with its neighbors. This restoration process in the brain can be called a type of optimization process in that the system tries to adjust itself and gradually approximates the desired functionality of the pre-treatment state by interaction with the environment under biological constraints.

The other central concept of the current study is the redundant nature of the non-linear brain (Glassman, 1987; Edelman and Gally, 2001). A complex system has degeneracy, i.e., the same or similar functions (behavior) can be achieved

using different configurations of networks (or connectivity). Since it is complicated to restore all connectivity after damage, optimal control utilizes non-linearity between networks and behaviors by reconfiguring networks among neighbors within limited ranges to approximate the goal system dynamics. In this non-linear relationship, the closeness in the system parameters (e.g., connectivity) does not necessarily indicate closeness in behaviors. Instead of matching network connectivity, the current framework fits behaviors (i.e., microstate dynamics) by modulating a smaller number of network parameters. This is possible as the brain is a complex non-linear system from which non-linear microstate dynamics emerge.

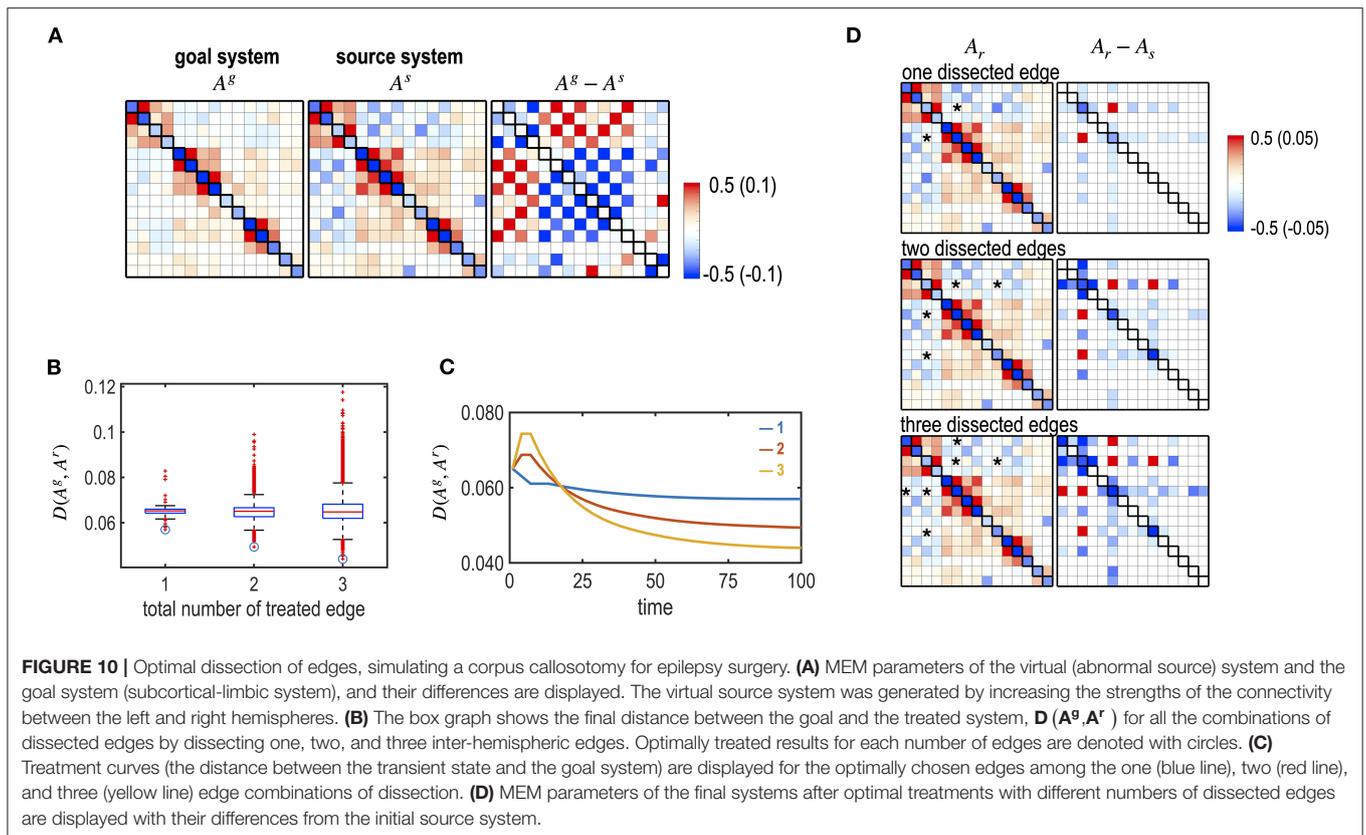


The non-linear microstate dynamics of a brain are represented in the microstate's energy landscape, the microstate of which is often defined by a distributed activity pattern over the temporal scale of a second. Energy landscape analysis has been applied to explore dynamics in large-scale functional brain networks, such as the default mode and pre-frontal networks, on resting-state fMRI (Watanabe et al., 2013, 2014a; Kang et al., 2017, 2019) and in sleep (Watanabe et al., 2014b). In our previous study (Kang et al., 2017), the energy landscape analysis revealed that the subcortical brain at rest exhibits the maximal number of stable states and small sets of stable states account for most of the occupation time. Furthermore, a graph theory analysis of the energy landscape revealed a hub-like state transition organization embedded in the resting-state human brain (Kang et al., 2019). The energy landscape of brain states is governed by a set of network parameters in the pairwise MEM, upon which treatment is imposed.

The brain control extends the energy landscape concept at the temporal scale of a second (microstates) to the energy landscape over a more extended period. Over a longer period of years, a

brain can be considered in a network state of the macroscopic energy landscape. For example, the brain develops from one network state (a set of network parameters) to another (another set of network parameters). A network state in the macroscopic energy landscape is defined by a network parameter, which differs from the definition of a microstate in the microscopic energy landscape by a distributed activity pattern. In this respect, the brain control problem is to choose an optimal way to guide a brain network to a desired network along the macroscopic energy landscape.

We assumed that the restoration process is conducted by the cooperative activity among neural populations within the brain network, which tries to generate similar functionality established before treatment. In this process, the brain network's modularity, an essential property to protect against damage to a complex brain (Park and Friston, 2013; Sporns and Betzel, 2016), would be crucial for the recovery. By rearranging resources within a module (e.g., altering connectivity within a biological range), the modularity actuates the system's reorganization to construct a similar behavior. There is plentiful evidence of modular-based reorganization in brain diseases (Balenzuela



et al., 2010; Chen et al., 2013; Siegel et al., 2018). In this study, nodes with more (stronger) connections (functional neighbors) play a more efficient role in restoration than nodes with fewer connections. Since the current test system has a relatively small network size (nodes = 15), we did not restrict the neighbors within a module but functionally close regions (Figure 6). Even though not strictly the same as the modularity concept in systems science, the functional neighbors work as a functional module in terms of cooperation within the module. Consideration of neighbors restricted within a functional module and within an anatomical limit of a larger network would be more realistic in modeling the recovery process.

Clinical treatment is generally exerted on the brain multiple times. After treatment, e.g., antipsychotic medication, clinicians wait to stabilize the brain to avoid transient states. However, one may consider applying subsequent treatment before stabilization. Until fully stabilized, the system has multiple transient states for network parameters. Some transient states may be more efficient in achieving a desired goal than the stabilized state. However, the transient state may be unstable, and finding an optimal strategy may be unpredictable. In the current study, we showed a possibility to optimize the right timing without waiting for complete stabilization when we have a model for self-restorative process.

The current framework as computational modeling takes advantage of prediction capacity by simulation. It is theoretically

possible that some treatment parameters may lead the treated system unstable, generating abnormal functionality. The self-restoration process may also cause the malformation of the function. By evaluating treatment outcomes for all possible ranges of parameters, we may check unstable points before deciding the treatment. The model-based prediction could also be used in evaluating the treatment effect due to noise in the restoration process as a type of Monte-Carlo simulation (See the **Supplementary Material**). As the noise effects differ across brain regions, one may choose a reliable target that is less sensitive to noise in the restoration process. The treatment could benefit from evaluating the treatment outcomes with noise in any parameters or any updating rules besides the restoration process. This Monte-Carlo simulation may complement the limitation of the current deterministic approach. We used a simple deterministic model and its solver for the control problem to explain the basic framework of brain control and show the current framework's construct validity. More sophisticated models based on more advanced control theory methods, such as a stochastic model proposed by Todorov (2009), could be further researched.

In this study, we showed the construct validity of the proposed framework using various simulations to consider the clinical environment. A simulation suggests that the optimal brain control should include the system's self-restoration process, without which a (so called optimal) treatment is not optimal. Using simulation, we also proposed how to control the

self-restoration process by choosing the optimal target region and treatment strength. We then presented simulations for optimizing repetitive treatments and the optimal timing of treatment. We found that some treatment choices led to a degraded performance at an early stage but eventually showed a better treatment effect (**Figure 8**). This is a typical example of the non-linear property of the self-restoration system that should be considered in optimal control. All of these simulations suggest the plausibility and rationale of the proposed brain control framework.

The current study is theoretical, and we acknowledge all possible limitations of the theoretical framework. The current brain control framework will be more practical when we know more about the system's reorganization mechanisms. Empirical experiments and validation are most demanding. Determining the means of achieving the desired treatment level at the right target for each treatment is one of the fundamental challenges. The details of the restoration process require extensive research and experiments. There exist many challenges before brain control can be applied to actual experiments. However, the current conceptual framework with the self-restorative process in the treatment is highly needed in clinical practices, which calls for personalized treatments based on individualized self-restoration systems and basic neuroscience to understand how the brain works.

In summary, we propose an optimal brain control framework by introducing self-restoration processes in the brain after treatment. Simulation results showing the responses and movement of a source system toward the desired system in diverse testing sets suggest the framework's plausibility in

optimal brain control within a restricted treatment environment. Although further research with experimental data should be conducted, we believe the proposed computational framework would help attain optimal brain control of the dynamic self-restorative brain after treatment.

DATA AVAILABILITY STATEMENT

The MATLAB codes used in the current study are available by request or online from a public drive, <http://bit.ly/monet-optimal-control-MEM>.

AUTHOR CONTRIBUTIONS

H-JP developed the idea and JK performed experiments. Both wrote the manuscript.

FUNDING

This research was supported by Brain Research Program and the Brain Pool Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017M3C7A1049051 and NRF-2017H1D3A1A01053094).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2021.590019/full#supplementary-material>

REFERENCES

- Abbott, A. (2010). Schizophrenia: the drug deadlock. *Nature* 468, 158–159. doi: 10.1038/468158a
- An, S., Bartolomei, F., Guye, M., and Jirsa, V. (2019). Optimization of surgical intervention outside the epileptogenic zone in the virtual epileptic patient (VEP). *PLoS Comput. Biol.* 15:e1007051. doi: 10.1371/journal.pcbi.1007051
- Balenzuela, P., Cherno moretz, A., Fraiman, D., Cifre, I., Sitges, C., Montoya, P., et al. (2010). Modular organization of brain resting state networks in chronic back pain patients. *Front. Neuroinform* 4:116. doi: 10.3389/fninf.2010.00116
- Bennabi, D., Charpeaud, T., Yrondi, A., Genty, J. B., Destouches, S., Lancrenon, S., et al. (2019). Clinical guidelines for the management of treatment-resistant depression: french recommendations from experts, the french association for biological psychiatry and neuropsychopharmacology and the fondation fondamentale. *BMC Psychiatry* 19:262. doi: 10.1186/s12888-019-2237-x
- Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409
- Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nat. Neurosci.* 20, 340–352. doi: 10.1038/nn.4497
- Cabral, J., Kringelbach, M. L., and Deco, G. (2014). Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* 114, 102–131. doi: 10.1016/j.pneurobio.2013.12.005
- Chen, G., Zhang, H. Y., Xie, C., Chen, G., Zhang, Z. J., Teng, G. J., et al. (2013). Modular reorganization of brain resting state networks and its independent validation in Alzheimer's disease patients. *Front. Hum. Neurosci.* 7:456. doi: 10.3389/fnhum.2013.00456
- Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., and Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron* 83, 238–251. doi: 10.1016/j.neuron.2014.05.014
- Cole, M. W., Ito, T., Bassett, D. S., and Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nat. Neurosci.* 19, 1718–1726. doi: 10.1038/nn.4406
- Cornblath, E. J., Tang, E., Baum, G. L., Moore, T. M., Adebimpe, A., Roalf, D. R., et al. (2019). Sex differences in network controllability as a predictor of executive function in youth. *Neuroimage* 188, 122–134. doi: 10.1016/j.neuroimage.2018.11.048
- Deco, G., and Jirsa, V. K. (2012). Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J. Neurosci.* 32, 3366–3375. doi: 10.1523/JNEUROSCI.2523-11.2012
- Deco, G., Tononi, G., Boly, M., and Kringelbach, M. L. (2015). Rethinking segregation and integration: contributions of whole-brain modelling. *Nat. Rev. Neurosci.* 16, 430–439. doi: 10.1038/nrn3963
- Edelman, G. M., and Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13763–13768. doi: 10.1073/pnas.231499798
- Ezaki, T., Sakaki, M., Watanabe, T., and Masuda, N. (2018). Age-related changes in the ease of dynamical transitions in human brain activity. *Hum. Brain Mapp.* 39, 2673–2688. doi: 10.1002/hbm.24033
- Falcon, M. I., Riley, J. D., Jirsa, V., Mcintosh, A. R., Elinor Chen, E., and Solodkin, A. (2016). Functional mechanisms of recovery after chronic stroke: modeling with the virtual brain. *eneuro* 3, 1–14. doi: 10.1523/ENEURO.0158-15.2016
- Fox, M. D., and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance

- imaging. *Nat. Rev. Neurosci.* 8, 700–711. doi: 10.1038/nrn2201
- Freyer, F., Roberts, J. A., Becker, R., Robinson, P. A., Ritter, P., and Breakspear, M. (2011). Biophysical mechanisms of multistability in resting-state cortical rhythms. *J. Neurosci.* 31, 6353–6361. doi: 10.1523/JNEUROSCI.6693-10.2011
- Freyer, F., Roberts, J. A., Ritter, P., and Breakspear, M. (2012). A canonical model of multistability and scale-invariance in biological systems. *PLoS Comput. Biol.* 8:e1002634. doi: 10.1371/journal.pcbi.1002634
- Friston, K. (2008). Hierarchical models in the brain. *PLoS Comput. Biol.* 4:e1000211. doi: 10.1371/journal.pcbi.1000211
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J. Physiol. Paris* 100, 70–87. doi: 10.1016/j.jphysparis.2006.10.001
- Friston, K. J., and Price, C. J. (2003). Degeneracy and redundancy in cognitive anatomy. *Trends Cogn. Sci.* 7, 151–152. doi: 10.1016/S1364-6613(03)00054-8
- Glassman, R. B. (1987). An hypothesis about redundancy and reliability in the brains of higher species: analogies with genes, internal organs, and engineering systems. *Neurosci. Biobehav. Rev.* 11, 275–285. doi: 10.1016/S0149-7634(87)80014-3
- Goellner, E., Bianchin, M. M., Burneo, J. G., Parrent, A. G., and Steven, D. A. (2013). Timing of early and late seizure recurrence after temporal lobe epilepsy surgery. *Epilepsia* 54, 1933–1941. doi: 10.1111/epi.12389
- Gu, S., Betzel, R. F., Mattar, M. G., Cieslak, M., Delio, P. R., Grafton, S. T., et al. (2017). Optimal trajectories of brain state transitions. *Neuroimage* 148, 305–317. doi: 10.1016/j.neuroimage.2017.01.003
- Gu, S., Cieslak, M., Baird, B., Muldoon, S. F., Grafton, S. T., Pasqualetti, F., et al. (2018). The energy landscape of neurophysiological activity implicit in brain network structure. *Sci. Rep.* 8:2507. doi: 10.1038/s41598-018-20123-8
- Jirsa, V. K., Proix, T., Perdikis, D., Woodman, M. M., Wang, H., Gonzalez-Martinez, J., et al. (2017). The virtual epileptic patient: individualized whole-brain models of epilepsy spread. *Neuroimage* 145, 377–388. doi: 10.1016/j.neuroimage.2016.04.049
- Jung, K., Friston, K. J., Pae, C., Choi, H. H., Tak, S., Choi, Y. K., et al. (2018). Effective connectivity during working memory and resting states: a DCM study. *Neuroimage* 169, 485–495. doi: 10.1016/j.neuroimage.2017.12.067
- Kang, J., Eo, J., Lee, D. M., and Park, H. J. (2021). A computational framework for optimal control of a self-adjustive neural system with activity-dependent and homeostatic plasticity. *Neuroimage* 230:117805. doi: 10.1016/j.neuroimage.2021.117805
- Kang, J., Pae, C., and Park, H. J. (2017). Energy landscape analysis of the subcortical brain network unravels system properties beneath resting state dynamics. *Neuroimage* 149, 153–164. doi: 10.1016/j.neuroimage.2017.01.075
- Kang, J., Pae, C., and Park, H. J. (2019). Graph-theoretical analysis for energy landscape reveals the organization of state transitions in the resting-state human cerebral cortex. *PLoS ONE* 14:e0222161. doi: 10.1371/journal.pone.0222161
- Kar, S. K. (2019). Predictors of response to repetitive transcranial magnetic stimulation in depression: a review of recent updates. *Clin. Psychopharmacol. Neurosci.* 17, 25–33. doi: 10.9758/cpn.2019.17.1.25
- Karrer, T. M., Kim, J. Z., Stiso, J., Kahn, A. E., Pasqualetti, F., Habel, U., et al. (2020). A practical guide to methodological considerations in the controllability of structural brain networks. *J. Neural Eng.* 17:026031. doi: 10.1088/1741-2552/ab6e8b
- Kelso, J. A. (2012). Multistability and metastability: understanding dynamic coordination in the brain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 906–918. doi: 10.1098/rstb.2011.0351
- King, A. (2016). Neurobiology: rise of resilience. *Nature* 531, S18–S19. doi: 10.1038/531S18a
- Krienen, F. M., Yeo, B. T. T., and Buckner, R. L. (2014). Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos. Trans. R. Soc. B Biol. Sci.* 369. doi: 10.1098/rstb.2013.0526
- Lee, I. C., Li, S. Y., and Chen, Y. J. (2017). Seizure recurrence in children after stopping antiepileptic medication: 5-year follow-up. *Pediatr. Neonatol.* 58, 338–343. doi: 10.1016/j.pedneo.2016.08.005
- Lee, W. H., Rodrigue, A., Glahn, D. C., Bassett, D. S., and Frangou, S. (2019). Heritability and cognitive relevance of structural brain controllability. *Cereb. Cortex.* 30, 3044–3054. doi: 10.1093/cercor/bhz293
- Liu, Y. Y., Slotine, J. J., and Barabasi, A. L. (2011). Controllability of complex networks. *Nature* 473, 167–173. doi: 10.1038/nature10011
- Malone, L. A., and Felling, R. J. (2020). Pediatric stroke: unique implications of the immature brain on injury and recovery. *Pediatr. Neurol.* 102, 3–9. doi: 10.1016/j.pediatrneurol.2019.06.016
- Marder, E., and Goaillard, J. M. (2006). Variability, compensation and homeostasis in neuron and network function. *Nat. Rev. Neurosci.* 7, 563–574. doi: 10.1038/nrn1949
- Marder, E., Gutierrez, G. J., and Nusbaum, M. P. (2017). Complicating connectomes: electrical coupling creates parallel pathways and degenerate circuit mechanisms. *Dev. Neurobiol.* 77, 597–609. doi: 10.1002/dneu.22410
- Mattson, M. P. (2008). Glutamate and neurotrophic factors in neuronal plasticity and disease. *Ann. N. Y. Acad. Sci.* 1144, 97–112. doi: 10.1196/annals.1418.005
- Mikellidou, K., Arrighi, R., Aghakhanyan, G., Tinelli, F., Frijia, F., Crespi, S., et al. (2019). Plasticity of the human visual brain after an early cortical lesion. *Neuropsychologia* 128, 166–177. doi: 10.1016/j.neuropsychologia.2017.10.033
- Murphy, T. H., and Corbett, D. (2009). Plasticity during stroke recovery: from synapse to behaviour. *Nat. Rev. Neurosci.* 10, 861–872. doi: 10.1038/nrn2735
- Murrough, J. W., and Russo, S. J. (2019). The neurobiology of resilience: complexity and hope. *Biol. Psychiatry* 86, 406–409. doi: 10.1016/j.biopsych.2019.07.016
- Olmí, S., Petkoski, S., Guye, M., Bartolomei, F., and Jirsa, V. (2019). Controlling seizure propagation in large-scale brain networks. *PLoS Comput. Biol.* 15:e1006805. doi: 10.1371/journal.pcbi.1006805
- Park, B., Kim, D.-S., and Park, H.-J. (2014). Graph independent component analysis reveals repertoires of intrinsic network components in the human brain. *PLoS ONE* 9:e82873. doi: 10.1371/journal.pone.0082873
- Park, H. J., and Friston, K. (2013). Structural and functional brain networks: from connections to cognition. *Science* 342:1238411. doi: 10.1126/science.1238411
- Park, H. J., Pae, C., Friston, K., Jang, C., Razi, A., Zeidman, P., et al. (2017). Hierarchical dynamic causal modeling of resting-state fMRI reveals longitudinal changes in effective connectivity in the motor system after thalamotomy for essential tremor. *Front. Neurol.* 8:346. doi: 10.3389/fneur.2017.00346
- Park, H. J., Park, B., Kim, H. Y., Oh, M. K., Kim, J. I., Yoon, M., et al. (2015). A network analysis of (1)(5)O-H(2)O PET reveals deep brain stimulation effects on brain network of parkinson's disease. *Yonsei Med. J.* 56, 726–736. doi: 10.3349/yjmj.2015.56.3.726
- Park, S., Park, H. J., Kyeong, S. H., Moon, I. S., Kim, M., Kim, H. N., et al. (2013). Combined rTMS to the auditory cortex and prefrontal cortex for tinnitus control in patients with depression: a pilot study. *Acta Otolaryngol.* 133, 600–606. doi: 10.3109/00016489.2012.763181
- Potkin, S. G., Kane, J. M., Correll, C. U., Lindenmayer, J. P., Agid, O., Marder, S. R., et al. (2020). The neurobiology of treatment-resistant schizophrenia: paths to antipsychotic resistance and a roadmap for future research. *NPJ Schizophr.* 6:1. doi: 10.1038/s41537-019-0090-z
- Proix, T., Bartolomei, F., Guye, M., and Jirsa, V. K. (2017). Individual brain structure and modelling predict seizure propagation. *Brain* 140, 641–654. doi: 10.1093/brain/awx004
- Rabinovich, M. I., and Varona, P. (2011). Robust transient dynamics and brain functions. *Front. Comput. Neurosci.* 5:24. doi: 10.3389/fncom.2011.00024
- Russo, S. J., Murrough, J. W., Han, M. H., Charney, D. S., and Nestler, E. J. (2012). Neurobiology of resilience. *Nat. Neurosci.* 15, 1475–1484. doi: 10.1038/nn.3234
- Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., et al. (2006). Dynamics of language reorganization after stroke. *Brain* 129, 1371–1384. doi: 10.1093/brain/awl090
- Schreglmann, S. R., Krauss, J. K., Chang, J. W., Martin, E., Werner, B., Bauer, R., et al. (2018). Functional lesional neurosurgery for tremor: back to the future? *J. Neurol. Neurosurg. Psychiatr.* 89, 727–735. doi: 10.1136/jnnp-2017-316301
- Siegel, J. S., Seitzman, B. A., Ramsey, L. E., Ortega, M., Gordon, E. M., Dosenbach, N. U. F., et al. (2018). Re-emergence of modular brain networks in stroke recovery. *Cortex* 101, 44–59. doi: 10.1016/j.cortex.2017.12.019
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045. doi: 10.1073/pnas.0905267106
- Sporns, O., and Betzel, R. F. (2016). Modular brain networks. *Annu. Rev. Psychol.* 67, 613–640. doi: 10.1146/annurev-psych-122414-033634

- Stiso, J., Khambhati, A. N., Menara, T., Kahn, A. E., Stein, J. M., Das, S. R., et al. (2019). White matter network architecture guides direct electrical stimulation through optimal state transitions. *Cell Re.* 28, 2554–2566.e2557. doi: 10.1016/j.celrep.2019.08.008
- Tang, E., Giusti, C., Baum, G. L., Gu, S., Pollock, E., Kahn, A. E., et al. (2017). Developmental increases in white matter network controllability support a growing diversity of brain dynamics. *Nat. Commun.* 8:1252. doi: 10.1038/s41467-017-01254-4
- Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., and Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science* 352, 216–220. doi: 10.1126/science.aad8127
- Todorov, E. (2009). Efficient computation of optimal actions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11478–11483. doi: 10.1073/pnas.0710743106
- Tognoli, E., and Kelso, J. A. (2014). The metastable brain. *Neuron* 81, 35–48. doi: 10.1016/j.neuron.2013.12.022
- Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., et al. (2012). The human connectome project: a data acquisition perspective. *Neuroimage* 62, 2222–2231. doi: 10.1016/j.neuroimage.2012.02.018
- Vorovenci, R. J., Biundo, R., and Antonini, A. (2016). Therapy-resistant symptoms in Parkinson's disease. *J. Neural Transm.* 123, 19–30. doi: 10.1007/s00702-015-1463-8
- Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., et al. (2013). A pairwise maximum entropy model accurately describes resting-state human brain networks. *Nat. Commun.* 4:1370. doi: 10.1038/ncomms2388
- Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., et al. (2014a). Energy landscapes of resting-state brain networks. *Front. Neuroinform* 8:12. doi: 10.3389/fninf.2014.00012
- Watanabe, T., Kan, S., Koike, T., Misaki, M., Konishi, S., Miyauchi, S., et al. (2014b). Network-dependent modulation of brain activity during sleep. *Neuroimage* 98, 1–10. doi: 10.1016/j.neuroimage.2014.04.079
- Watanabe, T., Masuda, N., Megumi, F., Kanai, R., and Rees, G. (2014c). Energy landscape and dynamics of brain activity during human bistable perception. *Nat. Commun.* 5:4765. doi: 10.1038/ncomms5765
- Yeh, F. C., Tang, A. N., Hobbs, J. P., Hottoway, P., Dabrowski, W., Sher, A., et al. (2010). Maximum entropy approaches to living neural networks. *Entropy* 12, 89–106. doi: 10.3390/e12010089
- Yeo, B. T. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., et al. (2015). Functional specialization and flexibility in human association cortex. *Cerebr. Cortex* 25, 3654–3672. doi: 10.1093/cercor/bhu217
- Yu, R., Park, H. J., Cho, H., Ko, A., Pae, C., Oh, M. K., et al. (2018). Interregional metabolic connectivity of 2-deoxy-2[(18) F]fluoro-D-glucose positron emission tomography in vagus nerve stimulation for pediatric patients with epilepsy: a retrospective cross-sectional study. *Epilepsia* 59, 2249–2259. doi: 10.1111/epi.14590

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Park and Kang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.