Quantifying brain state transition cost via Schrödinger's bridge

Genji Kawakita[1,2], Shunsuke Kamiya[2], Shuntaro Sasai[3,4], Jun Kitazono[2], and Masafumi Oizumi[2]

[1]Swarthmore College, Swarthmore, PA, USA
[2]Graduate School of Arts and Scence, The University of Tokyo, Tokyo, Japan
[3]Araya Inc, Tokyo, Japan
[4]University of Wisconsin-Madison, WI, USA

Abstract

Quantifying brain state transition cost is a fundamental problem in systems neuroscience. Previous studies utilized network control theory to measure the cost by considering a neural system as a deterministic dynamical system. However, this approach does not capture the stochasticity of neural systems, which is important for accurately quantifying brain state transition cost. Here, we propose a novel framework based on optimal control in stochastic systems. In our framework, we quantify the transition cost as the Kullback-Leibler divergence from an uncontrolled transition path to the optimally controlled path, which is known as Schrödinger's bridge. To test its utility, we applied this framework to functional magnetic resonance imaging data from the Human Connectome Project and computed the brain state transition cost in cognitive tasks. We demonstrate correspondence between brain state transition cost and the difficulty of tasks. The results suggest that our framework provides a general theoretical tool for investigating cognitive functions from the viewpoint of transition cost.

Introduction

The brain is considered a dynamical system that flexibly transitions through various states[1–3]. Depending on the properties of a dynamical system (e.g., the biophysical properties of neurons and the connectivity between neurons), some transitions are difficult to realize, i.e., they require a high cost while others are easy to realize. It appears reasonable that a difficult transition requiring a high cost in brain dynamics corresponds to effortful tasks or functions while an easy transition corresponds to effortless tasks. Thus, characterizing the dynamical properties of brain state transition would be important for understanding various brain functions[4], including decision-making[5], motor control[6], and working memory[7], with potential applications in the diagnosis and clinical treatment of disease[8–10]. To date, however, no unified framework for quantifying brain state transition cost from brain activity data has been available.

One promising framework for quantifying brain state transition cost is the network control-theoretic framework[11]. Control theory provides useful perspectives for measuring the cost required for controlling a dynamical system to reach a desirable state. Considering the brain as a dynamical system, control-theoretic approaches enable us to quantify the cost of transitioning to a brain state that produces desirable behavior. Recently, the network control theoretic framework was proposed for study of the control property of the brain by viewing the brain as a networked dynamical system[11–14]. Although the framework provides an important new perspective for fundamentally

understanding brain state transition, it has two major limitations. First, it does not capture the stochasticity of brain activity, which is ubiquitous in brain activity and is essential for accurately describing brain dynamics [15–17]. Disregarding stochasticity may result in an inaccurate estimation of transition cost. Second, the model obtained from structural connectivity, which is static over time, may not be able to capture change in the functional dynamics of the brain [4], such as while performing tasks, for instance. Recently, alternative models using functional and effective connectivity has been proposed [18,19], but these models still do not capture stochasticity. Thus, no unified framework able to take account of the key properties of brain dynamics is available.

Here, by employing control-theoretic approaches, we propose a novel framework for measuring brain state transition cost that can account for stochasticity. In our framework, we consider transition from a probability distribution of brain states to another distribution, rather than a transition from one brain state to another brain state (ie. a point-to-point transition in a state-space). To transition from an initial distribution to a target distribution, the brain needs to modulate (control) its baseline transition probability. Although there are many possible ways to reach the target distribution, in this study we consider the optimally controlled path only and estimate the lower bound of brain state transition cost. We propose defining the minimum brain state transition cost as the Kullback-Leibler (KL) divergence from the baseline uncontrolled path to the optimally controlled path, i.e., the closest path to the original path, with the fixed initial and target distributions. The problem of finding the closest path to the original path connecting the initial and target distribution is known as Schrödinger's bridge problem [20], which has been studied in the fields of stochastic process and optimal transport [21–25].

Here, as proof of concept, we apply the proposed framework based on Schrödinger's bridge problem to evaluate the cost of task switching [27], an executive function for moving from one task to another. Specifically, we address two questions. First, is the cost of transitioning to a more difficult task larger? A previous study [28] reported that performing effortful tasks drives larger reconfiguration of functional brain networks. We therefore hypothesized that transitioning to a more difficult task required a larger cost. Second, is the brain state transition cost asymmetric? Specifically, is the transition cost from an easier task to a more difficult task larger than the cost accompanying the reverse transition?

To address these questions, we apply our framework to functional magnetic resonance imaging (fMRI) data from the Human Connectome Project (HCP) [29]. We use fMRI data of $n = 937$ subjects in the resting state and in seven cognitive tasks. After preprocessing and parcellation, we computed the probability distributions of coarse-grain brain activity patterns for the rest and cognitive tasks [14,30]. We then calculated the brain state transition cost by finding the Schrödinger's bridge, i.e., the optimally controlled path [26] between the initial and target distributions of brain states. We found that the transition cost to a more difficult task carried a larger transition cost. We also observed that the transition cost between an easy and a difficult task is asymmetric. Overall, our findings provide a new perspective on the investigation of brain state transition, which may facilitate our understanding of cognitive functions.

Results

Quantification of brain state transition cost from Schrödinger's bridge problem

In this study, we propose a novel framework to quantify state transition cost in a stochastic neural system, building on the formulation of Schrödinger's bridge problem [26]. We consider brain state transition to be the transition from an initial probability distribution of brain states to a target probability distribution. In order to reach the target probability distribution, the brain is assumed
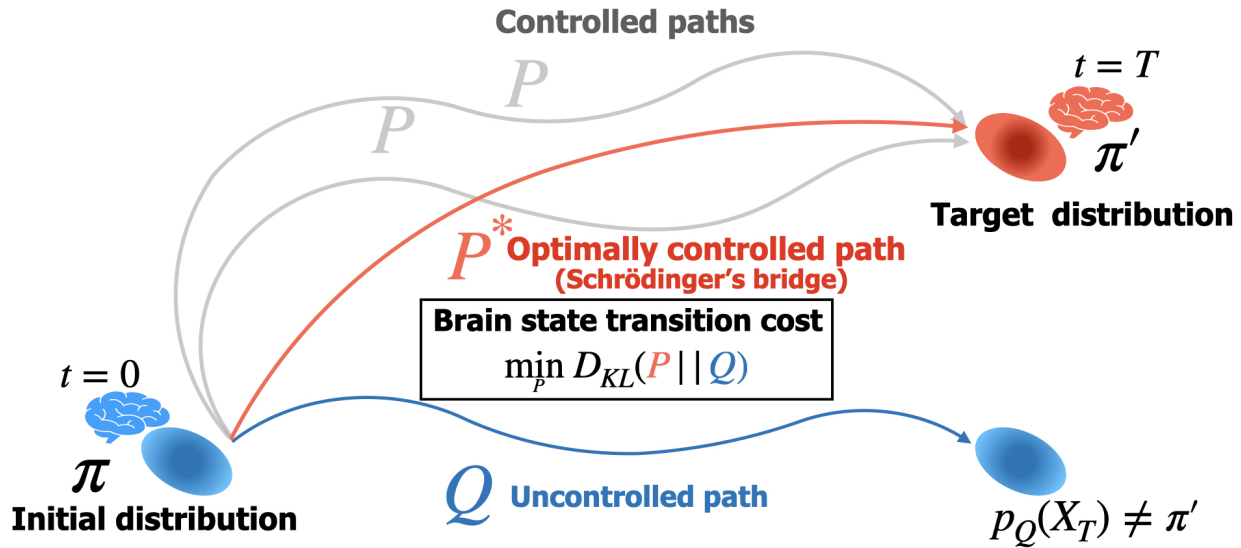
Figure 1: Schematic of brain state transition reframed as Schrödinger's bridge problem. We consider brain state transition as transition from an initial probability distribution of brain states, π, to a target probability distribution, π'. The brain follows an uncontrolled baseline transition probability, Q, which does not lead to the target distribution but to $p_Q(X_T) \neq \pi'$, where $p_Q(X_T)$ represents the probability distribution at $t = T$ under transition probability, Q. In order to reach the target distribution, the brain needs to follow some controlled transition probability, P. The brain state transition cost is defined as the minimum Kullback Leibler divergence between the controlled and uncontrolled transition probabilities (paths), P and Q, respectively. Optimally controlled transition probability, P^* is equivalent to Schrodinger's bridge. It was shown in the previous work[26] that P^* can be numerically obtained.

to follow some controlled paths. Although there are many possible paths that bridge the initial and target probability distributions, we look for the optimally controlled path which minimizes the Kullback-Leibler (KL) divergence from an uncontrolled to a controlled path. Here, we define brain state transition cost as the minimum KL divergence from an uncontrolled to a controlled path that bridges the initial and target probability distributions (Fig. 1).

We can mathematically formulate brain state transition as follows. Let $(X_0, X_1, ..., X_T)$ be a sequence of random variables that forms a first-order Markov chain on a finite set of brain activity states $\mathcal{S}$, where $|\mathcal{S}| = k$ is the number of brain states. Let Q be a transition probability matrix, where $Q_{ij} = \Pr(X_{t+1} = j | X_t = i)$ for all $i, j \in \mathcal{S}$. We consider Q as a baseline or an uncontrolled transition probability matrix. We have $p_Q(X_0) = \pi$ and $p_Q(X_T)$, which are k-dimensional row vectors. Here, p_Q represents a probability distribution of random variables $X_0, X_1, \cdots$, and X_T.

Under this setting, we consider the problem of reaching a target distribution π' such that $\pi' \neq p_Q(X_T)$ from the initial distribution π. In order to reach π', the brain needs to modulate Q. In other words, the brain must follow a controlled transition probability matrix, P, that satisfies $p_P(X_0) = \pi$ and $p_P(X_T) = \pi'$ (p_P is another probability distribution of random variables $X_0, X_1, \cdots$, and X_T).
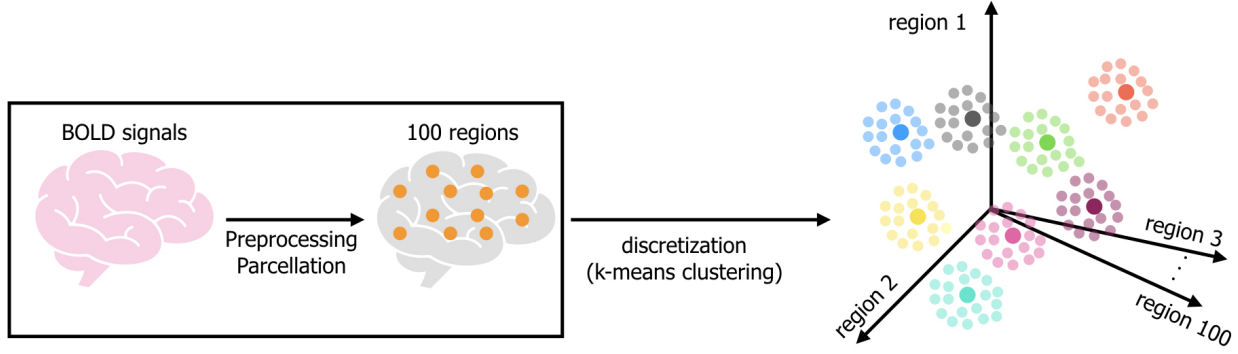
3

Figure 2: Clustering fMRI data. After preprocessing raw fMRI data, we used k-means clustering to group similar brain activity patterns into eight coarse-grained brain states.

While there are many possible Ps that satisfy these conditions, we want to find the optimal P^* that requires the minimum control cost. In Schrödinger's bridge problem, it is known that the optimally controlled transition matrix, P^*, is "closest" to the uncontrolled transition matrix, Q, which can be determined by solving the following minimization problem:

$$P^* = \arg\min_P D(P||Q) \tag{1}$$

where $D(P||Q)$ is the Kullback-Leibler divergence between Markov chains with transition probability matrices P and Q:

$$D(P||Q) = \mathbb{E}_{p_P} \log \frac{p_P(X_0, X_1, ..., X_T)}{p_Q(X_0, X_1, ..., X_T)} \tag{2}$$

Since the sequence $(X_0, X_1, ..., X_T)$ is Markovian and we can set $T = 1$ without the loss of generality[26], $D(P||Q)$ is simplified as follows:

$$D(P||Q) = \sum_{i \in \mathcal{S}} \pi_i \sum_{j \in \mathcal{S}} P_{ij} \log \left(\frac{P_{ij}}{Q_{ij}} \right) \tag{3}$$

The existence and uniqueness of P^* are guaranteed and can be obtained by solving Lagrangian multipliers:

$$P_{ij}^* = Q_{ij} e^{\lambda_i + \eta_j} \tag{4}$$

where $(\lambda_i)_{i \in \mathcal{S}}$, $(\eta_i)_{i \in \mathcal{S}}$ are real-valued constants that satisfy the following recursive relations:

$$\lambda_i = -\log(\sum_{j \in \mathcal{S}} Q_{ij} e^{\eta_j}) \tag{5}$$

$$\eta_i = \log(\pi_i^*) - \log(\sum_{j \in \mathcal{S}} \pi_j Q_{ji} e^{\lambda_j}) \tag{6}$$

The above relations can be solved numerically with an iterative method[26]. We define $D(P^*||Q)$ as the minimum control cost for steering a system from the initial distribution π to a target distribution π'.
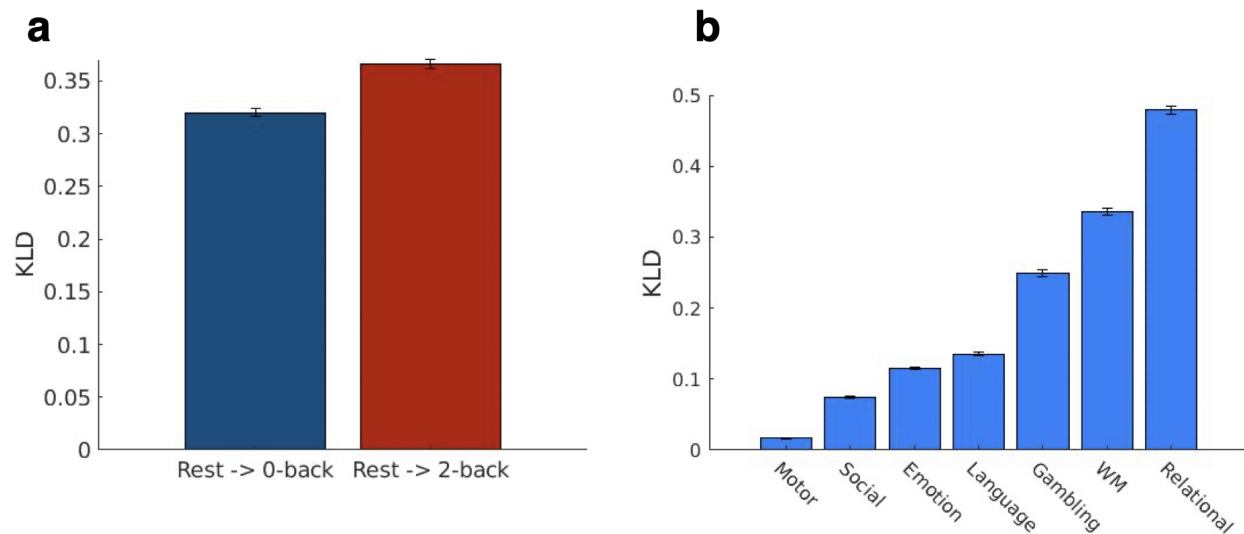
4

Figure 3: Brain state transition cost from the resting state to tasks. We computed the transition cost from the rest to cognitive tasks in the HCP data, as the minimum Kullback-Leibler divergence from the optimally controlled transition probability to the baseline uncontrolled (resting state) transition probability. (A) The transition cost from the 2-back task (more difficult) is larger than the transition cost to 0-back task (easier) (one-sided t-test, $p \ll 0.001$, $t > 60$, $df = 198$). (B) Transition cost from the rest to the seven cognitive tasks in the HCP dataset. Values are averaged over 100 bootstrapping trajectories and error bars indicate one standard deviation estimated with trajectory bootstrapping.

Quantification of brain state transition cost in fMRI data

To test the utility of our proposed method, we applied the Schrödinger's bridge-based framework to real fMRI data. We used resting-state fMRI and task-fMRI (emotion, gambling, language, motor, relational, social, and working memory) from the Human Connectome Project (HCP)[29]. As shown in Fig. 2, we first performed preprocessing of the BOLD signals and parceled them into 100 regions. We concatenated the preprocessed data of all subjects and applied the k-means clustering algorithm to group similar brain activity patterns into $k = 8$ coarse-grained states. We then computed the probability distributions and transition probability matrices of eight coarse-grained states for seven cognitive task states and the resting state. Here, we assume that an uncontrolled path in the brain is the resting state transition probability. Using these probability distributions and transition probability matrices, we calculated brain state transition cost as the minimum KL divergence (see Methods for more details.)

We began by testing whether transition cost from rest to a more difficult task is larger. For this purpose, we quantified the transition cost from the distribution at rest to those during 0-back (easier) and 2-back (more difficult) tasks in the working memory (WM) task data. We chose the WM task because the WM task data are the only task data in HCP, wherein subjects perform tasks with objectively different levels of task difficulty. As shown in Fig. 3, we found that the transition cost to a 2-back task is larger than that to a 0-back task. This result suggests that our cost metric may capture the level of task difficulty from fMRI data.
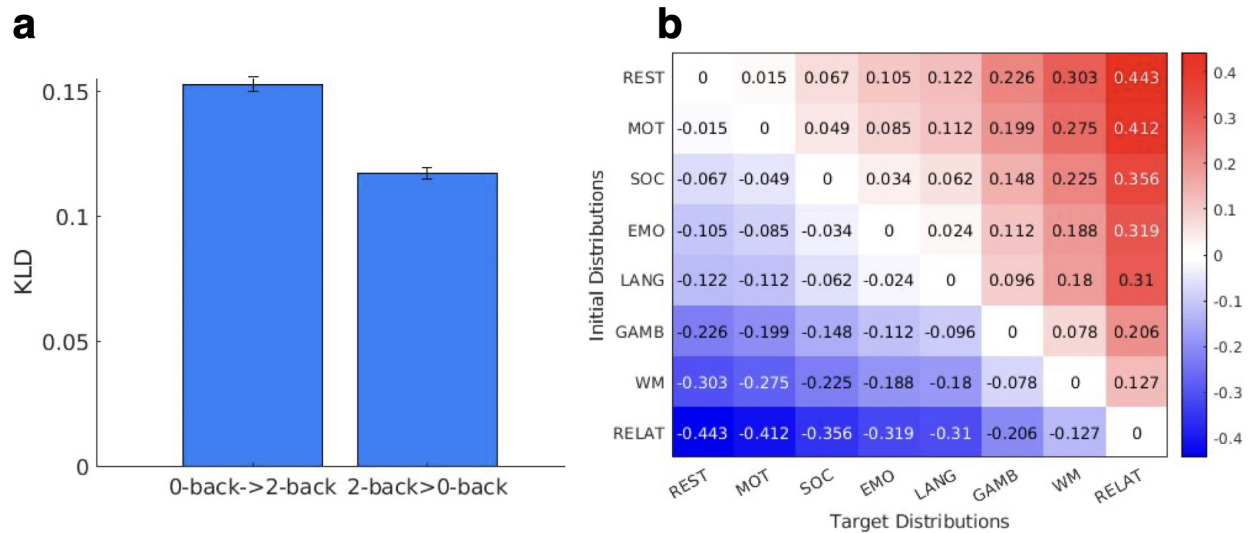
5

Figure 4: Brain state transition cost between task states. (a) Brain state transition cost between the 0-back and 2-back tasks in the working memory task. Values are averaged over 100 bootstrapping trajectories and error bars indicate one standard deviation estimated using trajectory bootstrapping. (b) Asymmetry of brain state transition costs for the rest and seven cognitive tasks. Each element in the matrix represents a difference in transition cost between tasks.

Brain state transition cost to multiple tasks

To further check the behavior of the proposed metric for transition cost, we then computed brain state transition cost to multiple task distributions in the HCP dataset (emotion, gambling, language, motor, relational, social, and working memory). Note that, unlike the working memory tasks, we cannot objectively compare their task difficulties since these tasks are qualitatively different. Thus, the analysis here is exploratory without any prior hypothesis.

We found that the degree of transition cost to the seven cognitive tasks is significantly different. Fig. 3b shows the rank order of transition costs in the seven cognitive tasks. Notably, transition cost to a motor task was smallest among the seven tasks, whereas the transition cost to a relational task was the largest (see Discussion).

Asymmetry of brain state transition cost

We then investigated whether state transition cost between tasks with different task difficulty was asymmetric. We hypothesized that it would require a larger transition cost to switch from an easier task to a more difficult task. To test this hypothesis, we computed the transition cost between 0-back and 2-back tasks in the working memory task. As shown in Fig. 4a, we found that the transition cost from a 0-back task to a 2-back task was larger than the cost accompanying the reverse direction, which agreed with our hypothesis (one-sided t-test, $p \ll 0.001$, $t > 80$, $df = 198$).

Finally, we examined whether the asymmetric property of brain state transition cost would be observed in other tasks whose task difficulties cannot be objectively compared. Here, we checked if the following relationship would hold for all the pairs of tasks (note that here we regarded the rest as a task): If the transition cost from rest to task A is larger than that from rest to task B, then

6

the transition cost from task B to task A is larger than that in the reverse direction,

$$\text{Cost}(\text{task B} \to \text{task A}) > \text{Cost}(\text{task A} \to \text{task B}),$$
$$\text{if } \text{Cost}(\text{rest} \to \text{task A}) > \text{Cost}(\text{rest} \to \text{task B}), \tag{7}$$

where $\text{Cost}(X \to Y)$ represents the brain state transition cost from X to Y, which is quantified by the KL divergence. To evaluate the relationship, we calculated the difference in transition cost between every pair of tasks, which is obtained as follows.

$$\text{Diff}(A, B) = \text{Cost}(\text{task A} \to \text{task B}) - \text{Cost}(\text{task B} \to \text{task A}) \tag{8}$$

The result is summarized in the matrix in Fig. 4b, wherein entries (tasks) are arranged in ascending order by transition cost from rest, i.e., the first row (column) corresponds to the task with the smallest transition cost from rest and the last row (column) corresponds to the task with the largest transition cost from rest. The (i, j) entry of the matrix represents $\text{Diff}(\text{task}_i, \text{task}_j) = \text{Cost}(\text{task i} \to \text{task j}) - \text{Cost}(\text{task j} \to \text{task i})$. We observed that every entry in the upper (lower) triangular parts was positive (negative). This means that the relationship represented in Eq. 7 holds for every pair of tasks in the dataset. That is, the transition cost is asymmetric between tasks with different degrees of transition cost.

Discussion

In this study, we propose a novel framework for quantifying brain state transition cost in stochastic neural systems by framing brain state transition as Schrödinger's bridge problem (SBP). This framework resolves the problem of previous methods that cannot take account of the inherent stochastisity of neural systems[17,31] while still utilizing principled control-theoretic approaches. Under this framework, we assumed that the brain follows the resting state activity as the baseline uncontrolled dynamic, and transitions to other distributions of brain state by modulating the baseline dynamics. Transition cost is measured as the minimum KL divergence from the uncontrolled path to the controlled path with the fixed endpoint probability distributions. We tested the utility of our framework by applying it to fMRI data from the Human Connectome Project release. The results indicated that the transition cost metric proposed in our framework may be useful for elucidating the characteristics of transition cost between tasks with different task difficulties.

Correspondence between transition cost and cognitive demands

In the present study, we aimed to examine the relationship between the degree of brain state transition cost and task difficulty as proof of concept. We refer to task difficulty as objectively quantifiable task difficulty (e.g. 0-back and 2-back) only, not as a subjectively experienced task difficulty, which could vary among subjects. As for the objective task difficulty, we observed that the transition cost to a 2-back task (a more difficult task) is larger than that to a 0-back task (an easier task). Further studies using different types of tasks with various levels of difficulty are needed to determine the generality of this result.

On the other hand, we did not deal with subjectively experienced task difficulty or cognitive demands, as the dataset does not contain subjective reporting on the cognitive demand of each task. Nevertheless, we quantified the transition cost to the seven qualitatively different tasks, whose task difficulty or cognitive demand cannot be objectively quantified. Although it is unclear whether the observed order of transition cost correlates with subjective cognitive demand (Fig. 3b), one may at least consider it reasonable that transition cost to a motor task is substantially smaller than that to a

relational task. This is because performing a motor task only requires subjects to move a part of their body (e.g., right hand or tongue) whereas performing a relational task requires processing multiple visual stimuli and reasoning their relationships, which appears significantly more demanding than performing a motor task. Although we could not further examine if the degree of transition cost correlates with the degree of cognitive demand, investigating this relation in more detail would be an interesting future work. We expect that while there may be a rough correlation between transition cost and cognitive demand, there can never be one-to-one correspondence, as many factors affect subjective evaluation of cognitive demand[32–35]. It would be intriguing to investigate the difference in transition cost and cognitive demand and in what cases these behave similarly or differently. The direction of study proposed in the present study might be an important step toward bridging cognitive demand and brain activity.

Relation to previous theoretical work

In the present study, we considered the brain dynamics as a discrete stochastic process by coarse-graining brain activity patterns, which reduced the computational cost. On the other hand, previous studies using network control theoretic framework[11] employed a linear continuous process. Our framework can also be extended to a linear continuous stochastic process because Schrödinger's bridge problem is not limited to a discrete process but has also been studied in continuous settings as well[22–24]. We therefore expected that we could directly fit the high dimensional neural recording data with a continuous model (e.g. stochastic differential equations) as recently implemented[36] and carry out a similar analysis as the present study. Developing a continuous version of the present framework may allow us to gain more insights into brain state transition cost.

Similarly to our framework, some recent works have utilized information theoretic measures to quantify cognitive costs[37] and connectivity changes between brain states[38]. While these studies compute only distance or divergence between the two distributions of brain state, our framework takes account of the underlying baseline activity of the brain. We employ this approach because including this baseline spontaneous activity provides a more accurate transition cost measure from the viewpoint of dynamical system theory.

Brain state transition and reconfiguration of functional connectivity

The brain state transition cost computed in our framework may be related to the reconfiguration of functional connectivity between tasks. Numerous functional neuroimaging studies have reported the alterations in functional connectivity from the resting state connectivity during task performance[39–43]. In our present study, we showed that transitioning to more difficult tasks carries a larger transition cost. This seems to be consistent with Kitzbichler's work, which demonstrated that larger cognitive demand induces a more global alteration in brain activity[28]. It may be the case that our framework captures the cost associated with the degree of reorganization of functional connectivity between different tasks.

Implications to task switching

The Schrödinger's bridge-based framework we proposed in this study may provide a new perspective for studying task switching from brain activity data. One of the most important and replicated findings in the task switching paradigm is the observation of switch cost[27]: switching to a new task takes a higher reaction time and error rate than repeating the same task. Various hypotheses have been proposed to explain the source of switch cost[44,45], including reconfiguration of the mental set for performing tasks. However, few studies have quantified the switch cost from brain activity. A recent

work suggests that task switching involves the reconfiguration of brain-wide functional connectivity[46]. Our framework may be used as a quantitative method for measuring switch cost. Subsequent investigations should study the relationship between switch cost and brain state transition cost by measuring brain activity while a subject is performing a task-switching experiment.

Comparison between the optimally controlled path and the empirical path

We investigated only the computed optimally controlled path, not an empirical path due to the limitation of fMRI recording. To quantify the efficiency of brain state transition, it would be interesting to compare empirical and optimal paths. This may provide insight into individual differences in the performance of task switching. However, the fMRI data from Human Connectome Project does not include recordings in which subjects perform and switch between multiple tasks, and we were therefore unable to compute an empirical transition path between initial and target distributions of brain states. Even if the dataset contained such data, fMRI would not capture rapid transitions between tasks because the time resolution of the fMRI data is not sufficiently high (TR = 0.72 second in the HCP data set). Computing empirical transition paths will require the use of recording data with better temporal resolution, such as EEG, MEG, ECoG, etc. In this regard, our theoretical framework is applicable to other types of recording data beside fMRI.

Acknowledgments

Methods

fMRI data acquisition and preprocessing

The 3T functional magnetic resonance imaging (fMRI) data of 937 subjects were obtained from the Washington University-Minnesota Consortium Human Connectome Project (HCP)[29]. We used minimally preprocessed fMRI data at resting state and seven cognitive task states (emotion, gambling, language, motor, relational, social, and working memory). We selected these 937 subjects as they contain complete data for all the tasks. We then performed denoising by estimating nuisance regressors and subtracting them from the signal at every vertex[47]. For this, we used 36 nuisance regressors and spike regressors introduced in a previous study, consisting of (1-6) six motion parameters, (7) a white matter time series, (8) a cerebrospinal fluid time series, (9) a global signal time series, (10-18) temporal derivatives of (1-9), and (19-36) quadratic terms for (1-18). Following a previous study[47], the spike regressors were computed with 1.5 mm movement as a spike identification threshold. After regressing these nuisance time courses, we also applied a band-pass filter (0.01-0.69 Hz) to the data, in which the upper bound of the filter corresponds to the Nyquist frequency of the time-series. We then applied a parcellation proposed in[48] to divide the cortex into 100 brain regions, which reduced the complexity of the following analysis.

Clustering BOLD signals

In order to model brain dynamics as a discrete stochastic process, we coarse-grained brain activity patterns using the k-means clustering algorithm. While there are numerous unsupervised clustering algorithms, we chose the k-means clustering due to its effective fit with the dynamics of neural
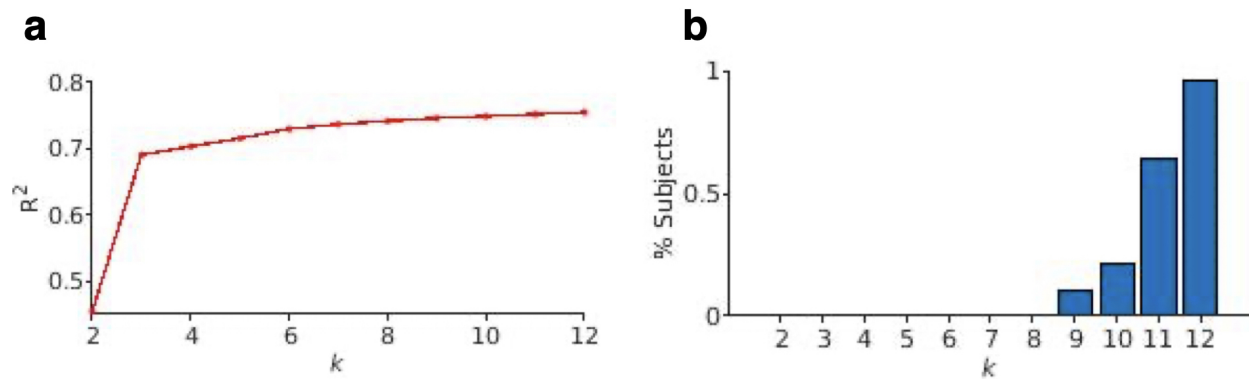
a **b**

Figure 5: Choosing the number of clusters. (a) Percentage of variance explained with increasing k. (b) The percentage of subjects missing at least one cluster in one of the fMRI scans.

activities[14]. We used cosine similarity as a distance measure between centroids of clusters, which is commonly used in high dimensional data. As described in the previous studies[14,30], we concatenated preprocessed BOLD signals of all the subjects during the resting state and seven cognitive tasks. We obtained a $M \times N$ matrix, X, where M is the number of parcels (100), and N is the number of task types times the number of time frames times the number of subjects. In order to prevent the variability in data size across tasks from affecting the clustering results, we used the same number of time frames for each task data. We used a different number of time frames depending on whether we divided the working memory task data into 0-back and 2-back tasks or not. When we divided the working memory task data (Figs. 3a and 4a), we obtain 148 time frames that included either 0-back or 2-back task blocks. Accordingly, we only used the first 148 time frames in the other tasks. When we did not divide the working memory task data (Figs. 3b and 4b), we used only the first 176 discrete measurements in each task since the emotion task - the task with the shortest measurement - was recorded for 176 time frames.

We determined the number of clusters using a similar procedure to that in previous studies[14,30]. We first computed the percent variance explained by the number of clusters varying from $k = 2$ to $k = 12$. We observed that the explained variance plateaued around 75% after $k = 5$ (Figure 4a). We then examined whether all the coarse-grained states would appear in every subject during each task session (Fig. 5b). We found that when we set the number of clusters to be greater than $k = 8$, some coarse-grained states did not appear in the data of some subjects. For these two reasons, we selected the number of clusters to be 8. While we chose $k = 8$, we were able to reproduce the major results with $k > 4$ (data not shown), which indicates the robustness of our results regardless of the number of clusters.

Estimating probability distributions and transition probabilities of coarse-grained states using trajectory bootstrapping

Here, we are limited by the finite length of the time-series data for estimating brain state transition cost, which are calculated using probability distributions and transition probabilities. To ensure the accuracy of estimated quantities, we applied trajectory bootstrapping[30,49] to calculate error bars on the estimated quantities. After we classified brain activity data with k-means clustering, we estimated for each task the $k \times k$ joint probability distribution matrix, M, of coarse-grained states at time $t = t$ and $t = t + 1$, where k is the number of coarse-grained states. To obtain the matrix,

we first created a list of transitions in concatenated time-series data of each task, in accordance with previous work[30]

$$
I = \begin{pmatrix} i_1 & i_2 \\ i_2 & i_3 \\ . & . \\ . & . \\ . & . \\ i_{L-1} & i_L \end{pmatrix}
\tag{9}
$$

where i_l is the coarse-grained state at l-th frame of the time-series, and L is the length of the concatenated time-series (L = the number of time frames × the number of subjects). We sampled a pair of transitions from the list for L times to fill in the matrix, M, and normalized it to be a joint probability matrix. Although the transition list is a concatenated time-series across subjects, we excluded pairs of transitions that took place across subjects; we only sampled pairs within the same subject. By normalizing each row of the matrix, M, to be 1, we constructed a transition probability matrix, Q, where Q_{ij} represents the probability of transitioning from the i-th coarse-grained state at $t = t$ to the j-th coarse-grained state at $t = t+1$. Finally, we computed the probability distribution of a task, π, as the stationary probability distribution of a transition probability matrix, Q, i.e., $\pi = Q\pi$. We calculated a probability distribution and a transition probability matrix for each task including the rest. We followed this process 100 times and computed the error bars on the estimated quantities in this study using the 100 bootstrap trajectories.

When we consider the transition from an emotion task to the motor task, for example, we first need to compute the initial probability distribution, π, from the emotion task data and the target probability distribution, π' from the motor task data. We can compute π as the stationary distribution such that $\pi = Q_e\pi$, where Q_e is the transition probability matrix estimated from the emotion task data. Similarly, we can compute π' as the stationary distribution such that $\pi = Q_m\pi$, where Q_m is the transition probability matrix estimated from the motor task data. We also need to calculate the baseline uncontrolled transition probability, Q, which is the transition probability matrix estimated from the data at rest.

References

[1] McKenna, T. M., McMullen, T. A. & Shlesinger, M. F. The brain as a dynamic physical system. *Neuroscience* **60**, 587–605 (1994).

[2] Breakspear, M. Dynamic models of large-scale brain activity. *Nat. Neurosci.* **20**, 340–352 (2017).

[3] Vyas, S., Golub, M. D., Sussillo, D. & Shenoy, K. V. Computation through neural population dynamics. *Annu. Rev. Neurosci.* **43**, 249–275 (2020).

[4] Kringelbach, M. L. & Deco, G. Brain states and transitions: Insights from computational neuroscience. *Cell Rep.* **32**, 108128 (2020).

[5] Taghia, J. *et al.* Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. *Nat. Commun.* **9**, 2505 (2018).

[6] Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* **36**, 337–359 (2013).

11

[7] Simmering, V. R. & Perone, S. Working memory capacity as a dynamic process. *Front. Psychol.* **3**, 567 (2012).

[8] Deco, G. & Kringelbach, M. L. Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron* **84**, 892–905 (2014).

[9] Aerts, H. *et al.* Modeling brain dynamics after tumor resection using the virtual brain. *Neuroimage* **213**, 116738 (2020).

[10] Adhikari, M. H. *et al.* Decreased integration and information capacity in stroke measured by whole brain models of resting state activity. *Brain* **140**, 1068–1085 (2017).

[11] Gu, S. *et al.* Controllability of structural brain networks. *Nat. Commun.* **6**, 8414 (2015).

[12] Bassett, D. S. & Sporns, O. Network neuroscience. *Nat. Neurosci.* **20**, 353–364 (2017).

[13] Gu, S. *et al.* Optimal trajectories of brain state transitions. *Neuroimage* **148**, 305–317 (2017).

[14] Cornblath, E. J. *et al.* Temporal sequences of brain activity at rest are constrained by white matter structure and modulated by cognitive demands. *Commun Biol* **3**, 261 (2020).

[15] Rieke, F. *Spikes: exploring the neural code* (MIT press, 1999).

[16] Shadlen, M. N. & Newsome, W. T. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896 (1998).

[17] Deco, G., Rolls, E. T. & Romo, R. Stochastic dynamics as a principle of brain function. *Prog. Neurobiol.* **88**, 1–16 (2009).

[18] Deng, S. & Gu, S. Controllability analysis of functional brain networks (2020). `2003.08278`.

[19] Szymula, K. P., Pasqualetti, F., Graybiel, A. M., Desrochers, T. M. & Bassett, D. S. Habit learning supported by efficiently controlled network dynamics in naive macaque monkeys (2020). `2006.14565`.

[20] Schrodinger, E. Uber die umkehrung der naturgesetze. sitz. ber. der preuss. *Akad. Wissen. , Berlin Phys. Math* **144** (1931).

[21] Beghi, A. On the relative entropy of discrete-time markov processes with given end-point densities. *IEEE Trans. Inf. Theory* **42**, 1529–1535 (1996).

[22] Dai Pra, P. A stochastic control approach to reciprocal diffusion processes. *Appl. Math. Optim.* **23**, 313–329 (1991).

[23] Chen, Y., Georgiou, T. T. & Pavon, M. Optimal steering of a linear stochastic system to a final probability distribution, part I. *IEEE Trans. Automat. Contr.* **61**, 1158–1169 (2016).

[24] Léonard, C. A survey of the schrödinger problem and some of its connections with optimal transport (2013). `1308.0215`.

[25] Chen, Y., Georgiou, T. T. & Pavon, M. On the relation between optimal transport and schrödinger bridges: A stochastic control viewpoint. *J. Optim. Theory Appl.* **169**, 671–691 (2016).

[26] Pavlichin, D. S., Quek, Y. & Weissman, T. Minimum power to maintain a nonequilibrium distribution of a markov chain (2019). 1907.01582.

[27] Monsell, S. Task switching. *Trends Cogn. Sci.* **7**, 134–140 (2003).

[28] Kitzbichler, M. G., Henson, R. N. A., Smith, M. L., Nathan, P. J. & Bullmore, E. T. Cognitive effort drives workspace configuration of human brain functional networks. *J. Neurosci.* **31**, 8259–8270 (2011).

[29] Van Essen, D. C. *et al.* The WU-Minn human connectome project: an overview. *Neuroimage* **80**, 62–79 (2013).

[30] Lynn, C. W., Cornblath, E. J., Papadopoulos, L., Bertolero, M. A. & Bassett, D. S. Nonequilibrium dynamics and entropy production in the human brain (2020). 2005.02526.

[31] Daunizeau, J., Stephan, K. E. & Friston, K. J. Stochastic dynamic causal modelling of fMRI data: should we care about neural noise? *Neuroimage* **62**, 464–481 (2012).

[32] McGuire, J. T. & Botvinick, M. M. Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7922–7926 (2010).

[33] Frömer, R., Lin, H., Dean Wolf, C. K., Inzlicht, M. & Shenhav, A. Expectations of reward and efficacy guide cognitive control allocation. *Nat. Commun.* **12**, 1030 (2021).

[34] Kool, W., McGuire, J. T., Rosen, Z. B. & Botvinick, M. M. Decision making and the avoidance of cognitive demand. *J. Exp. Psychol. Gen.* **139**, 665–682 (2010).

[35] Rosenbaum, D. A. & Bui, B. V. Does task sustainability provide a unified measure of subjective task difficulty? *Psychon. Bull. Rev.* **26**, 1980–1987 (2019).

[36] Nozari, E. *et al.* Is the brain macroscopically linear? a system identification of resting state dynamics (2020). 2012.12351.

[37] Zénon, A., Solopchuk, O. & Pezzulo, G. An information-theoretic perspective on the costs of cognition. *Neuropsychologia* **123**, 5–18 (2019).

[38] Amico, E., Arenas, A. & Goñi, J. Centralized and distributed cognitive task processing in the human connectome. *Netw Neurosci* **3**, 455–474 (2019).

[39] Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S. & Petersen, S. E. Intrinsic and task-evoked network architectures of the human brain. *Neuron* **83**, 238–251 (2014).

[40] Davison, E. N. *et al.* Brain network adaptability across task states. *PLoS Comput. Biol.* **11**, e1004029 (2015).

[41] Spadone, S. *et al.* Dynamic reorganization of human resting-state networks during visuospatial attention. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8112–8117 (2015).

[42] Stitt, I. *et al.* Dynamic reconfiguration of cortical functional connectivity across brain states. *Sci. Rep.* **7**, 8797 (2017).

[43] Cole, M. W., Ito, T., Cocuzza, C. & Sanchez-Romero, R. The functional relevance of task-state functional connectivity. *J. Neurosci.* (2021).

[44] Jersild, A. T. Mental set and shift. *Archives of Psychology* **14**, 89 (1927).

[45] Koch, I., Gade, M., Schuch, S. & Philipp, A. M. The role of inhibition in task switching: A review. *Psychon. Bull. Rev.* **17**, 1–14 (2010).

[46] Daws, R. E. *et al.* Optimisation of brain states and behavioural strategies when learning complex tasks (2020).

[47] Satterthwaite, T. D. *et al.* An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* **64**, 240–256 (2013).

[48] Schaefer, A. *et al.* Local-Global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).

[49] Battle, C. *et al.* Broken detailed balance at mesoscopic scales in active biological systems. *Science* **352**, 604–607 (2016).