QUANTIFYING STATE-DEPENDENT CONTROL PROPERTIES OF BRAIN DYNAMICS FROM PERTURBATION RESPONSES

Yumi Shikauchi[1,2], **Mitsuaki Takemi**[3,4], **Leo Tomasevic**[4,5,6], **Jun Kitazono**[1,7], **Hartwig R Siebner**[4,8,9], and **Masafumi Oizumi**[1]

[1] Graduate School of Arts and Sciences, the University of Tokyo, Tokyo, Japan
[2] Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan
[3] Graduate School of Science and Technology, Keio University, Tokyo, Japan
[4] Danish Research Centre for Magnetic Resonance (DRCMR), Department of Radiology and Nuclear Medicine, Copenhagen University Hospital Amager and Hvidovre, Copenhagen, Denmark
[5] Department of Psychiatry and Psychotherapy, University of Regensburg, Regensburg, Germany
[6] Department of Human Sciences, Institute of Psychology, University of the Bundeswehr Munich, Neubiberg, Germany
[7] Graduate School of Data Science, Yokohama City University, Kanagawa, Japan
[8] Department of Neurology, Copenhagen University Hospital Bispebjerg and Frederiksberg, Copenhagen, Denmark
[9] Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Corresponding Authors: Yumi Shikauchi, `yumi.shikauchi@gmail.com`;
Masafumi Oizumic-oizumi@g.ecc.u-tokyo.ac.jp

Abbreviated title: Quantifying Control Properties of Brain by Perturbation

Number of pages: 34 Number of figures = 5, tables = 1, multimedia items = 0, and 3D models = 0

Conflict of interest statement

Hartwig R. Siebner has received honoraria as a speaker and consultant from Lundbeck AS, Denmark, and as an editor (Neuroimage Clinical) from Elsevier Publishers, Amsterdam, The Netherlands. He has received royalties as a book editor from Springer Publishers, Stuttgart, Germany; Oxford University Press, Oxford, UK; and from Gyldendal Publishers, Copenhagen, Denmark.

Acknowledgements

We would like to express our gratitude to Dr. Mikito Ogino for his valuable comments and suggestions on earlier drafts of this manuscript. We extend our heartfelt thanks to Dr. Ben D. Fulcher for his invaluable contributions and insightful discussions that greatly shaped the initial direction of this research. This work was supported by JST Moonshot R&D Grant Number JPMJMS2012 to M. O. and Y. S., JSPS KAKENHI Grant Numbers 18K15341 and 21K19810

to Y. S., the project “Precision Brain-Circuit Therapy - Precision-BCT” from Innovation Fund Denmark (grant nr. 9068-00025B) to H. R. S., the project “ADAptive and Precise Targeting of cortex-basal ganglia circuits in Parkinson’s Disease - ADAPT-PD” from The Lundbeck Fondation (collaborative project grant, grant nr. R336-2020-1035) to H. R. S., and Lundbeck Foundation Experiment grant (grant no. R346-2020-1822) to L. T.

ABSTRACT

The brain can be conceptualized as a control system facilitating transitions between states, such as from rest to motor activity. Applying network control theory to measurements of brain signals enables characterization of brain dynamics through control properties, including controllability. However, most prior studies that have applied network control theory have evaluated brain dynamics under unperturbed conditions, neglecting the critical role of external perturbations in accurate system identification, which is a fundamental principle in control theory. The incorporation of perturbation inputs is therefore essential for precise characterization of brain dynamics. In this study, we combine a perturbation input paradigm with a network control theory framework and propose a novel method for estimating the controllability Gramian matrix in a simple, theoretically grounded manner. This method provides insights into brain dynamics, including overall controllability (quantified by the Gramian’s eigenvalues) and specific controllable directions (represented by its eigenvectors). As a proof of concept, we applied our method to transcranial magnetic stimulation (TMS)-induced electroencephalographic (EEG) responses across four motor-related states and two resting states. We found that states such as open-eye rest, closed-eye rest, and motor-related states were more effectively differentiated using controllable directions than overall controllability. However, certain states, like motor execution and motor imagery, remained indistinguishable using these measures. These findings indicate that some brain states differ in their intrinsic control properties as dynamical systems, while others share similarities that make them less distinguishable. This study underscores the value of control theory-based analyses in quantitatively how intrinsic brain states shape the brain’s responses to stimulation, providing deeper insights into the dynamic properties of these states. This methodology holds promise for diverse applications, including characterizing individual response variability and identifying conditions for optimal stimulation efficacy.

***K*eywords** Data-driven control, network controllability, concurrent TMS-EEG measurement, motor execution, motor imagery, resting state.

Significant statement

The brain can be viewed as a control system transitioning between states, such as from rest to motor activity. Previous studies using network control theory mostly evaluated brain dynamics without external perturbations, neglecting their

role in accurate system identification. This study integrates perturbation inputs with network control theory to propose a method for estimating the controllability Gramian, thereby providing insights into brain dynamics. We applied this approach to TMS-induced EEG responses in motor-related and resting states. Our findings show that controllable directions (eigenvectors) allow better discrimination between states than overall controllability. Our method can quantitatively assess brain state differences, and has potential applications in characterizing individual response variability and optimizing stimulation efficacy.

1 Introduction

Network control theory has been applied to neuronal networks to deepen our understanding of complex cognitive processes in neuroscience [Medaglia et al., 2017, Gu et al., 2015]. The approach models the brain as a control system, highlighting its ability to transition between states, such as from motor activity to rest. It introduces mathematically-grounded and interpretable measures to analyze brain activity as multidimensional time series across different regions. The core concept is controllability, which refers to the brain's capacity to shift its dynamics from one state to another via internal or external inputs. Mathematically, controllability is defined using the controllability Gramian, a matrix that captures how the brain's state changes under the influence of these inputs over time. This framework offers a more holistic understanding of neural processes compared with fragmented approaches focusing on specific activity features.

Pioneering studies have used network control theory to explore how the brain's structural connectivity constrains its dynamics. They have provided mechanistic explanations for transitions between cognitive states, identifying key areas of the brain that can facilitate these transitions [Gu et al., 2015] and providing insight into optimal trajectories between states, such as moving from high activity in the default mode network to high activity in sensorimotor systems [Gu et al., 2017]. These studies have significantly advanced our understanding of how the structural connectivity of the brain governs its control capabilities.

Our study aims to advance the understanding of brain dynamics by shifting the focus from structural connectivity to dynamic neural activity data and examination of state-dependent changes in control properties. Previous research has largely relied on passive data (i.e., data without external inputs) [Li et al., 2023, Kawakita et al., 2022, Kamiya et al., 2023], limiting its ability to fully characterize the brain's dynamic nature. To address this limitation, it is considered essential to apply perturbative inputs is essential for accurate system identification [Jakowluk, 2024], enabling a more precise characterization of brain dynamics.

Although not explicitly grounded in control theory, the importance of perturbative inputs has been demonstrated in studies investigating cognitive states such as working memory [Rose et al., 2016], states of consciousness [Casali et al., 2010, Casali et al., 2013, Lee et al., 2022], and spatial attention [Okazaki et al., 2020]. However, while these studies underscore the value of perturbative approaches, their reliance on one-dimensional, global measures like the Perturbational Complexity Index (PCI) [Casali et al., 2010, Casali et al., 2013, Lee et al., 2022] limits the granularity of

insights because these measures primarily capture the overall complexity or the degree of signal propagation in brain dynamics.

In this study, we integrate perturbation inputs with the control theory framework to introduce a novel method for quantifying multivariate state-dependent changes in network control properties. Our approach provides a simple yet theoretically-grounded method for estimating the controllability Gramian from neural activity data with perturbation inputs. The controllability Gramian offers insights not only into the overall controllability of brain dynamics, as quantified by its eigenvalues, but also into the specific directions in which these dynamics can be effectively influenced, as represented by its eigenvectors. Additionally, under the assumption that the perturbative input acts as an impulse, we show that standard control theory methods can be used to estimate the controllability Gramian directly from the time series responses alone, using standard control theory methods. This approach eliminates the need to estimate the system's connectivity matrix A or input matrix B (see Eq. (1)), thereby simplifying the estimation process.

As a proof of concept, we applied our method to single-trial, single-pulse TMS-EEG data, covering four motor-related task conditions and two resting-state conditions (Table 1). By treating the short TMS pulses (less than 2 ms) as impulse responses, our method allows the controllability Gramian to be directly estimated of the controllability Gramian from the TMS-evoked EEG data. This approach focuses on short-latency responses, minimizing the influence of peripheral effects, as demonstrated in previous studies [Conde et al., 2019, Biabani et al., 2021, Rocchi et al., 2021]. Even with such a short latency, the high temporal resolution of EEG allows for robust state-dependent characterization of neural dynamics. We demonstrate that our proposed method captures the state-dependent change in neural dynamics between motor-related and resting states in terms of both the magnitude and the directions of controllability.

Table 1: Experimental conditions

Condition	Abbreviation	Condition overview (TMS timing)
Motor execution	ME	Right index finger isometric abduction for 3 s (1.5-2.5 s after the task onset)
Motor imagery	MI	Imagery of the motor execution task (1.5-2.5 s after the task onset)
After motor execution	AfterME	After the motor execution task (4-5 s after the task onset)
Resting state with open eyes	OE	(every 4-6 s)
Resting state with closed eyes	CE	(every 4-6 s)
Ipsilateral motor execution	iME	Left index finger isometric abduction for 3 s (1.5-2.5 s after the task onset)

2 Materials and Methods

2.1 Theoretical background to control theory

In this study, we consider a linear auto-regressive (AR) model as a model of multidimensional brain dynamics and quantify the control properties of the system given some external input. Specifically, we consider the following model:

$$\boldsymbol{x}(t) = A\boldsymbol{x}(t-1) + Bu(t-1), \tag{1}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is a n-dimensional vector of the brain activity at time t,

$$\boldsymbol{x}(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T, \tag{2}$$

A is the $n \times n$ connectivity matrixthat describes the interaction between the elements, $u(t-1) \in \mathbb{R}$ is a one-dimensional external input at time $t-1$, and B is an $n \times 1$ input matrix describing how the control input u affects the brain activity x. We consider the dynamics from the time $t = 0$ to $t = m$ (Fig. 1).

We wish to characterize both to what extent and how the external input u can control the system. To this end, the most important quantity to quantify is the so-called controllability Gramian. We will briefly explain the definition and meaning of the controllability Gramian, but for further details please refer to [Brunton and Kutz, 2019].

The controllability Gramian W is defined using the controllability matrix $\mathcal{C}$ as follows:

$$W = \mathcal{C}\mathcal{C}^T, \tag{3}$$

where

$$\mathcal{C} = [B, AB, A^2B, \ldots, A^{m-1}B]. \tag{4}$$

The controllability Gramian is important because it gives the solution to an example optimal control problem is: given the initial state $\boldsymbol{x}(0) = \boldsymbol{x}_s$ and the final target state $\boldsymbol{x}(m) = \boldsymbol{x}_f$, find the most efficient control inputs in terms of minimum cost.

Here, to simplify the expression, we introduce a new variable $\tilde{\boldsymbol{x}}(t)$, which is the time series $\boldsymbol{x}(t)$ subtracted from $\boldsymbol{x}(0)$, $\tilde{\boldsymbol{x}}(t) \equiv \boldsymbol{x}(t) - \boldsymbol{x}(0)$, so that $\tilde{\boldsymbol{x}}(0) = 0$ (Fig. 1a). Under this notation, the optimal control problem is formulated by

$$\min_{u(\tau)} \left\{ \sum_{\tau=0}^{m-1} ||u(\tau)||^2 \;\middle|\; \tilde{\boldsymbol{x}}(0) = 0, \tilde{\boldsymbol{x}}(m) = \tilde{\boldsymbol{x}}_f \right\}, \tag{5}$$

where $\tilde{\boldsymbol{x}}_f = \boldsymbol{x}_f - \boldsymbol{x}(0)$. By solving this constrained optimization, we can find the minimal cost and the optimal control input according to:

$$\min_{u(\tau)} \sum_{\tau=0}^{m-1} ||u(\tau)||^2 = \tilde{\boldsymbol{x}}_f W^{-1} \tilde{\boldsymbol{x}}_f, \tag{6}$$

$$u(\tau) = B^T (A^T)^{m-1-\tau} W^{-1} \tilde{\boldsymbol{x}}_f. \tag{7}$$

The equation for the minimum control cost (Eq. (6)) gives the exact mathematical meaning of the controllability Gramian W in terms of the optimal control framework. Assuming that the controllability Gramian has a set of eigenvalues λ_i and corresponding normalized eigenvectors $\boldsymbol{\xi}_i$, then, from Eq. (6), we can easily see that the larger the eigenvalue is, the larger we can move the state of the system in the corresponding direction of the eigenvector. More precisely, given a fixed amount for the sum of the squared control input, i.e., $\sum_{\tau=0}^{m-1} ||u(\tau)||^2 = 1$, the squared dis-

tance we can move in the direction of an eigenvector is exactly given by its eigenvalue. This can be shown by setting the final target state to $\tilde{\boldsymbol{x}}_f = r_i \boldsymbol{\xi}_i$, where r_i corresponds to the distance to move, and substituting this final state into Eq. (6) gives $r_i^2 = \lambda_i$.

In summary, by ordering the eigenvectors of the controllability Gramian according to the highest to lowest eigenvalues from highest to lowest, we can tell which direction we can easily move with the optimal control input: the direction of the largest eigenvalue corresponds to the most easily controllable direction, that of the second largest eigenvalue corresponds to the second most easily controllable direction, and so on (Fig. 1b). Note that the degree of ease of control is defined by the hypothetical optimal control input, not by the actual impulse input. In Fig. 1b, we illustrate this visually by considering the controllability Gramian as a hyperellipse, whose longest axis corresponds to the first eigenvector, whose second longest axis corresponds to the second eigenvector, and so on. In this study, we characterize the control property of the brain dynamics in this way.

2.2 Experimental design and statistical tests

2.2.1 Participants

Seventeen healthy volunteers (11 women, age range: 18–29 years) participated in the experiment. All participants had slept between 6.5 and 11.0 hours the night before the experiment, were right-handed as assessed by the Edinburgh handedness inventory [Oldfield, 1971] (range: 65–100), did not smoke regularly except for one participant, had no history of neurological or psychiatric diseases, and had no contraindications to TMS. After receiving written and oral information about the experiment, the participants gave written informed consent before testing. All TMS-EEG data were collected between December 2016 and March 2017 as a part of a TMS-EEG project approved by the local ethics committee of the capital region of Denmark (H-15008824). The study was carried out under the guidelines of the Helsinki Declaration.

2.2.2 EEG and EMG recording

EEG and electromyographic (EMG) signals were recorded with a NeuroOne Tesla system (Mega Electronics, Kuopio, Finland) with a 63-channel equidistant M10 EEG cap (EASYCAP, Herrsching, Germany) and surface EMG electrodes (Ambu Neuroline 710, Ballerup, Denmark). EMG electrodes were placed on the first dorsal interosseous muscle on the right hand, using a bipolar belly-tendon montage and a ground electrode mounted at the right ulnar styloid process. All EEG electrodes, including a ground electrode and a reference electrode of the cap, were prepared using NuPrep Skin Prep gel and ABRALYT HiCl EEG electrode gel (EASYCAP). EEG electrode impedance was maintained below 5 kΩ throughout the recording periods. EEG and EMG signals were digitized at 5 kHz with a 1250-Hz low-pass filter and DC filter.

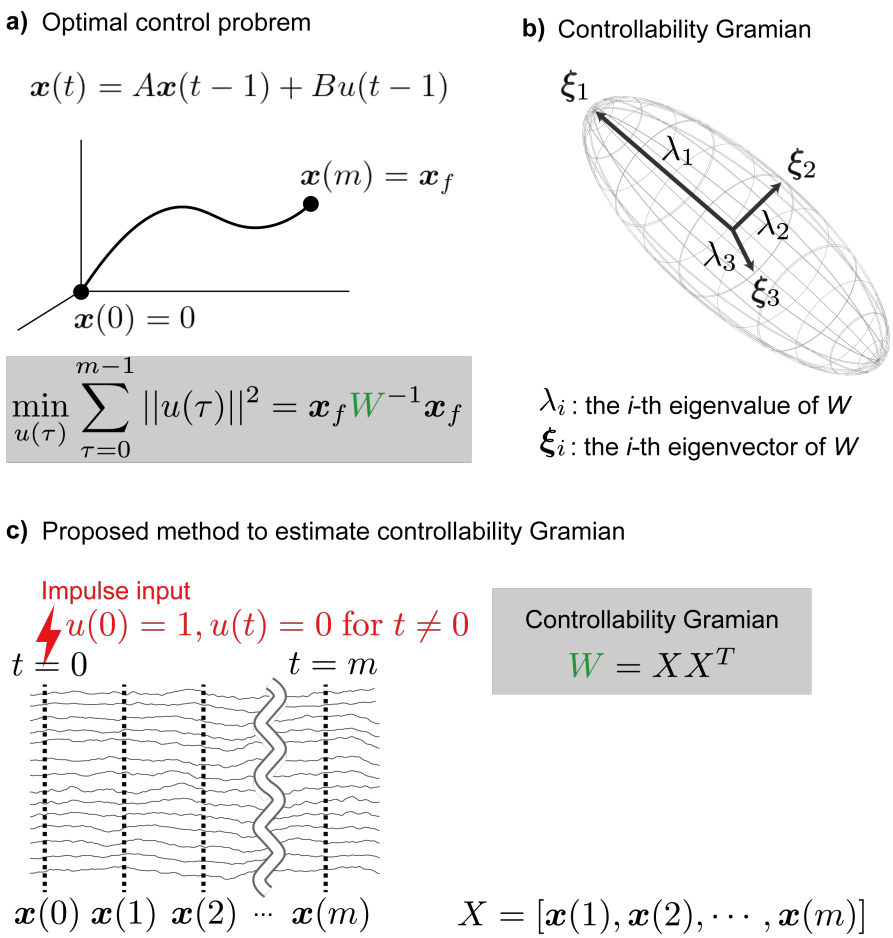

Figure 1: Schematic illustration of the controllability Gramian. (a) In a first-order auto-regressive model, the transition from the initial state $\boldsymbol{x}(0)$ to the final target state $\boldsymbol{x}(m)$ at minimum cost is realized if the inverse of the controllability Gramian is known, as the equation in the gray shaded box shows. (b) The Gramian W is an n-dimensional hyper-ellipse, encompassing all spatial patterns that appear in the controllability matrix. n is the measurement dimension as long as the measurement dimension does not exceed the number of time points. The major axis shows the first eigenvector $\boldsymbol{\xi}_1$. The second and third eigenvectors $\boldsymbol{\xi}_2$ and $\boldsymbol{\xi}_3$ are shown as arrows orthogonal to the first eigenvector. The length of each axis corresponds to the magnitude of change as an eigenvalue $\lambda_\cdot$, and the direction of a specific spatial pattern as an eigenvector $\boldsymbol{\xi}_\cdot$. (c) Assuming that the control input is an impulse, the controllability matrix can be regarded as the measurement signals X, as the equation in the gray shaded box shows. Under this assumption, we obtain the controllability Gramian without estimating the connectivity matrix A and input matrix B. Each line on the left indicates the time series measured by each sensor.

2.2.3 Neuro-navigated TMS-EEG

TMS was applied over the hand area of the primary motor cortex (M1-HAND) of the left hemisphere using a MagPro X100 with an MC-B70 figure-of-eight coil (MagVenture, Farum, Denmark). The coil was placed tangentially on the scalp, with the handle directed posteriorly at a $45°$ angle to the sagittal midline. To reduce TMS artifacts in the EEG signals, a soft silicone rubber sheet (1.5-mm thick) was placed between the coil and the EEG electrodes. Biphasic pulses, in which the second phase induced a posterior-to-anterior current in the cortex, were employed. Continuous monitoring of the TMS coil placement was ensured via a frameless stereotaxic neuro-navigation system (Localite,

Sankt Augustin, Germany). The root mean square of the disparity from the co-registered landmark over the M1-HAND was kept under 2 mm throughout the experiment.

The individual M1-HAND hotspot was identified as the coil position eliciting the most consistent and prominent motor-evoked potentials (MEPs) using suprathreshold stimulation intensity while the subject was at rest. This hotspot was marked in the neuro-navigation system. The resting motor threshold was determined after mounting the EEG cap using maximum-likelihood parameter estimation with the sequential testing approach while participants remained at rest [Awiszus, 2003]. MEPs exceeding 50 µV were classified as 'responding.'

2.2.4 TMS-EEG measurements

Participants underwent six conditions (Table 1), during which a single-pulse of TMS was applied to M1-HAND at a stimulation intensity of 120% of the resting motor threshold for each trial. One hundred trials were carried out per condition. Participants sat in a comfortable armchair, and auditory noise masking was consistently applied throughout the TMS-EEG measurements using in-ear headphones (Insert Earphone 3A 410–3002, 3M systems). Consistent with previous procedures, we employed specific time- and frequency-varying sounds for noise masking [Conde et al., 2019, Herring et al., 2015, Beck et al., 2024]. The sound pressure for noise masking was individually adjusted to a level where participants could not hear the click sound of the TMS pulse with the TMS coil placed on their M1-HAND or until they reached their upper threshold for comfort.

Measurements commenced with either the OE or CE conditions described in Table 1, with the order of these conditions being pseudo-randomized across participants. Participants placed their right hand, palm side down, on the armrest and remained at rest. In the OE condition, participants were instructed to keep their eyes open and fixate on a fixation cross, while the TMS pulse was administered every 4–6 s. In the CE condition, participants were instructed to keep their eyes closed, and the TMS pulse was delivered every 4–6 s.

Following the OE and CE conditions, three conditions involving right-hand movement or motor imagery (ME, MI, and AfterME) were conducted. In these three conditions, each trial began with the presentation of open gray and filled blue circles (ME and AfterME) or an open blue circle (MI) on a computer monitor placed in front of the participants. After the presentation of the two circles, participants performed 3 s of isometric abduction of the right index finger at 10% of maximum voluntary contraction. The size of the filled blue circle increased depending on the force level applied. Participants were instructed to match the size of the two circles as quickly as possible, meaning that the applied force level corresponded to 10% of the maximum voluntary contraction. In the MI condition, after presentation of the open blue circle, participants performed 3 s of kinesthetic motor imagery of the ME content without moving their right hand. After 3 s, a gray cross appeared on the monitor, which prompted participants to cease motor tasks and return to a resting state until the next trial began. TMS was administered 1.5–2.5 s from the start of each trial in the ME and MI conditions and 4–5 s from the start in the AfterME condition. The ME, MI and AfterME conditions were pseudo-randomized per trial and were paused multiple times to allow participants a short break and prevent mental fatigue.

Finally, we conducted the iME condition, in which participants were asked to perform the ME content with their left hand. TMS was administered 1.5-2.5 s from the start of each trial.

2.2.5 EEG data preprocessing

Offline data analysis was performed using with the FieldTrip toolbox for EEG- and MEG analysis [Oostenveld et al., 2011]. Our proposed method compresses the time-series data in the temporal dimension (see Theoretical background to control theory). Since high-amplitude components can have lasting effects beyond their occurrence, the following preprocessing steps were implemented to minimize the impact of cranial muscle artifacts as much as possible. Initially, TMS-induced artifacts were removed from EEG signals using the TESA methods [Rogasch et al., 2017] with MATLAB software (MathWorks Inc., Natick, MA). In brief, the TMS-EEG data underwent the following processing steps: 1) removal of noisy channels; 2) epoching of data from -1 to $+1$ s relative to the TMS pulse for the OE and CE conditions and from -1 to $+6$ s relative to the onset of the motor task for the other conditions; 3) baseline correction with data from -500 to -110 ms relative to the TMS pulse for the OE and CE conditions or relative to the motor task onset for other conditions; 4) removal of periods containing TMS pulse artifacts and TMS-evoked muscle activity (-2 to 15 ms relative to the TMS pulse); 5) interpolation of missing data around the TMS pulse with constant amplitude values; 6) removal of noisy trials; 7) independent component analysis (ICA) and removal of components representing muscle artifacts; 8) replacement of data of the removed window (-2 to 15 ms relative to the TMS pulse) by interpolation using a cubic function; 9) second-order Butterworth filtering (1–80 Hz); 10) notch filtering; 11) replacement of data from the removed window with constant amplitude values; 12) independent component analysis to remove residual artifacts; 13) spatial interpolation of the rejected channels; 14) common average re-referencing; and 15) linear interpolation of the 0–20 ms signal was applied to completely eliminate the residual effects of artifacts. As a result of the above preprocessing, an average of 2.25 (SD 1.51, range $0-8$) channels, an average of 13.99% (SD 12.10, range $0-58$) of trials per session were excluded. Through these preprocessing steps, we obtained TMS-induced artifact-free EEG data.

2.2.6 Extracting the controllability matrix from single-shot TMS-EEG

We then extracted the controllability matrix corresponding to Eq. (13) from the TMS-induced artifact-free data. Focusing on cortical responses, signals with short latencies up to 60 ms were included in the analysis, thereby avoiding peripheral responses, which have a strong influence on long latencies [Biabani et al., 2019, Rocchi et al., 2021]. The analysis was performed according to the following steps: 1) scaling by the standard deviation of signals from -200 to -5 ms before stimulation to correct for differences in signal values per electrode; 2) downsampling to 1070 Hz to align the number of time points with the number of channels; 3) baseline correction with the values immediately previous to satisfy the assumption that $\tilde{\boldsymbol{x}}(0) = 0$; and 4) segmentation from 1 to 60 ms from TMS onset. This resulted in a 63-channel $\times$ 63-time-points data matrix of EEG signals for each trial. There is no need to estimate the connectivity matrix A and the input matrix B in Eq. (1).

When the analysis was performed over longer time windows, i.e., 1–120 ms and 1–300 ms, the downsampling frequency was set to 530 Hz and 220 Hz, respectively, so that the controllability matrix was a square matrix.

2.2.7 Definition of pairwise distance

We treated the single-shot TMS-EEG data as the controllability matrix described in the previous section, from which we computed the controllability Gramians (Eq. (3)). For each trial, the resulting controllability Gramian was further analyzed through its eigenvalues and eigenvectors. To characterize trial-by-trial variations, we quantified pairwise distances between trials on the basis of these properties. This analysis was implemented in MATLAB using in-house code and the publicly available SPDtoolbox on GitHub (`https://github.com/kyoustat/papers`) [You and Park, 2021].

We used the following three distance measures.

Gramian-based distance The controllability Gramians are considered to be on the symmetric positive definite (SPD) space, which is a Riemannian manifold. For any two SPD matrices w and w', the affine-invariant Riemannian metric between them is defined as [Bhatia, 2009]:

$$D^{\text{total}}(w, w') = \|\text{Log}(w^{-1}w')\|_F, \tag{8}$$

where $\text{Log}(a)$ is the matrix logrithm of a square matrix a, and $\|a\|_F$ is the Frobenius norm of a.

When focusing only on up to the mth eigenmode (i.e., excluding the eigenmodes after the $m+1$st eigenmode), the reconstructed controllability Gramian $\hat{w}$ could be obtained accoding to the following. The original controllability Gramian W was decomposed into its eigenvalues and eigenvectors according to:

$$W = U\Lambda U^T$$

where U is the matrix of eigenvectors and Λ is the diagonal matrix containing the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ in descending order. To achieve dimensionality reduction, we reconstructed the Gramian matrix using only the top m eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$ and their corresponding eigenvectors. The reconstructed Gramian $\hat{W}$ is given by:

$$\hat{W} = U_m \Lambda_m U_m^T$$

where U_m consists of the first m columns of U, and Λ_m is the diagonal matrix containing the first m eigenvalues. This allowed us to approximate the original Gramian in a lower-dimensional space.

Eigenvalue-based distance Eigenvalues are considered as points on the complex number plane and Euclidean distance is used. The inter-trial distance of the average contorollability was calculated using the following equation:

$$D^{\text{magnitude}} = \frac{|\sum_{i=1}^m \lambda_i - \sum_{i=1}^m \lambda_i'|}{\sum_{i=1}^m \lambda_i + \sum_{i=1}^m \lambda_i'},$$

where λ_i is the i-th eigenvalue of one trial, and λ'_i corresponds to that of another trial. m represents the number of eigenmodes, which is typically equal to the number of channels ($m = n = 63$) when all dimensions are utilized. However, m can be set to a smaller value than n when focusing only on the upper eigenmodes.

Eigenvector-based distance The angle between the two eigenvectors was used to compare the direction independently of the magnitude. The distance between the i-th eigenvector of one trial x_i and the j-th eigenvector of another trial y_j was defined using the cosine distance as $d^{\text{direction}}_{x_i,y_j} = \frac{x_i \cdot y_j}{\|x_i\|\|y_j\|}$. Using $d^{\text{direction}}_{x_i,y_j}$, the distance of the controllable direction up to the mth eigenmode of a trial pair was determined as follows:

$$D^{\text{direction}}_{x,y} = \frac{\sum_{i=1}^{m} \min(d^{\text{direction}}_{x_i,y_1}, d^{\text{direction}}_{x_i,y_2}, \cdots, d^{\text{direction}}_{x_i,y_m})}{m}, \tag{9}$$

where m represents the number of eigenmodes. $D^{\text{direction}}_{x,y}$ is the order-free cosine distance, which uses the distance of the most similar eigenvectors, regardless of the order of the magnitude of the eigenvalues. Since the resulting distance matrix is asymmetric ($D^{\text{direction}}_{x,y} \neq D^{\text{direction}}_{y,x}$), we symmetrized it by taking the arithmetic mean of each pair of corresponding trials, as follows:

$$D^{\text{direction}} = \frac{D^{\text{direction}}_{x,y} + D^{\text{direction}}_{y,x}}{2}. \tag{10}$$

2.2.8 Classification of conditions according to inter-trial distance matrices

Each of the three distance matrices we obtained used a different metric, and were thuson different scales and were difficult to compare. Therefore, we performed six-condition classification using the k-nearest neighbor method with each distance matrix and obtained a common index of classification performance ($k = 5$). Each trial was classified into one of six conditions, yielding a 6×6 confusion matrix. The confusion matrices were averaged across participants. Setting k to 1 or 3 and focusing on more local neighborhoods did not change the results.

2.2.9 Measure for dimensionality of controllable directions

To verify the dimensionality of the controllability Gramian, the cumulative contribution rate up to the m-th eigenmode was determined as follows:

$$\text{Cumulative contribution rate} = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_i}{\sum_{i=1}^{m} \lambda_i}, \tag{11}$$

where λ_i is the i-th engenvalue. Rank, which is the number of eigenvectors with eigenvalues above a certain value (e.g., $\lambda_. > 0.0001$), is often used as another measure of dimensionality. When rank is used, factors such as measurement conditions (e.g., number of noisy channels) and preprocessing (e.g., how many independent components were excluded in the preprocessing) that are unrelated to the dimensionality of the controllable space and have no significant impact on the overall signal (eigenvalues are expected to be small) are affected. Therefore, in this study, dimensionality was evaluated using the cumulative contribution ratio, which allowed us to focus only on modes with large eigenvalues.

3 Results

Building on the theoretical background of control theory outlined in the Methods section, we first propose a method to estimate the controllability Gramian for evoked neural activity data acquired following an impulse response without explicitly estimating the system's parameters A and B. The method itself is well known in standard control theory, but to our knowledge, it has not yet been used in neuroscience applications. Then, we demonstrate the utility of the proposed method by applying it to our TMS-EEG data. By integrating the proposed theoretical framework with empirical data analysis, we investigate the state dependence of the control properties of the EEG dynamics in response to single-shot TMS.

3.1 Proposed method to estimate the controllability Gramian

Here, under the assumption that the external stimulus is an impulse input, we propose a novel method to directly estimate the controllability Gramian, bypassing the need to estimate the system's parameters. In general, to estimate the controllability Gramian from brain activity data with general external stimulation not restricted to an impulse input, we need to estimate A and B in Eq. (1). Once we have estimated A and B using some statistical methods, we can compute the controllability Gramian given by Eq. (3). However, in the special case where the external input is an impulse input, i.e., the input $u(t)$ takes a non-zero value only at the time of stimulation ($t = 0$) and then takes zero values for the rest of the time ($[u(0), u(1), \dots u(m)] = [1, 0, \dots, 0]$), we do not need to explicitly estimate A and B, but compute the controllability Gramian directly.

The theoretical validity of this simplification can be demonstrated as follows. If we again introduce $\tilde{\boldsymbol{x}}(t)$ as in the Materials and Methods, and assume that the external stimulus is an impulse input, the time series of neural activity X is given by

$$X = [\tilde{\boldsymbol{x}}(1), \tilde{\boldsymbol{x}}(2), \cdots, \tilde{\boldsymbol{x}}(m)] = [B, AB, \cdots, A^{m-1}B]. \tag{12}$$

Comparing the above equation with the controllability matrix in Eq. (4), we can see that the time series of neural activity, X, is exactly equal to the controllability matrix $\mathcal{C}$,

$$\mathcal{C} = X = [\tilde{\boldsymbol{x}}(1), \tilde{\boldsymbol{x}}(2), \cdots, \tilde{\boldsymbol{x}}(m)]. \tag{13}$$

Finally, using Eq. (3), we obtain the controllability Gramian as

$$\begin{aligned} W &= CC^T && (14)\\ &= XX^T && (15)\\ &= \begin{pmatrix} \tilde{\boldsymbol{x}}_1(1:m)\cdot\tilde{\boldsymbol{x}}_1(1:m) & \dots & \tilde{\boldsymbol{x}}_1(1:m)\cdot\tilde{\boldsymbol{x}}_n(1:m) \\ \vdots & \ddots & \vdots \\ \tilde{\boldsymbol{x}}_n(1:m)\cdot\tilde{\boldsymbol{x}}_1(1:m) & \dots & \tilde{\boldsymbol{x}}_n(1:m)\cdot\tilde{\boldsymbol{x}}_n(1:m) \end{pmatrix} && (16) \end{aligned}$$

where $\tilde{\boldsymbol{x}}_i(1:m)$ is the time series of the i-th element from $t = 1$ to $t = m$,

$$\tilde{\boldsymbol{x}}_i(1:m) = [\tilde{x}_i(1), \tilde{x}_i(2), \cdots, \tilde{x}_i(m)], \tag{17}$$

where $\tilde{x}_i(t) = x_i(t) - x_i(0)$ (Fig. 1c).

3.2 Examples of controllability Gramians from TMS-EEG responses

Using the proposed method illustrated in the previous section to compute the controllability Gramian, we here show several examples of controllability Gramian estimated from our TMS-EEG data. The analysis pipeline is summarized in Fig. 2a (see Methods for details). This study uses a dataset of EEG measurements made during TMS (Table 1). TEPs were recorded with a 63-channel EEG system ($n = 63$, see Materials and Methods). We set the time $t = 0$ to the time of TMS onset, and $\boldsymbol{x}(t)$ is a 63-dimensional vector of TEPs at time t after the TMS onset. The TMS pulse is considered to be the impulse input, i.e., $u(0) = 1$ and $u(t) = 0$ for any $t \neq 0$. An example of TMS-EEG response data acquired during the ME condition is shown in Fig. 2b. Then, using the equation for the controllability Gramian under an impulse input (Eq. (16)), we obtain an $n \times n$ controllability Gramian matrix from the EEG data after each single-pulse TMS. These matrices are illustrated in Fig. 2c, where the top and bottom panels show results from two different trials under the ME condition.

3.3 Inter-trial distances based on the controllability Gramian

To explore the extent to which controllability characteristics vary among conditions, we computed the inter-trial distance based on the three metrics: the controllability Gramian-based, eigenvalue-based, and eigenvector-based distances (Figs. 2c–2e)(see Methods for details). As explained in the previous section, the controllability Gramian contains all the information for control properties associated with the TMS input. The distance in terms of the controllability Gramian represents the overall difference between trials in terms of control properties. By contrast, the distance for the sum of eigenvalues or eigenvectors represents the difference in terms of partial aspects of control properties, i.e., the overall degree of controllability and the direction in which the system can be controlled.

Figure 3a shows the inter-trial distance matrix based on the controllability Gramian. The indices of the matrix correspond to the trials of the TMS experiments, with each condition corresponding to one of six conditions: ME; MI; AfterME; OE; CE; or iME (Table 1). The trials are ordered from top to bottom according to the order of measurement for each session. The right-hand motor-related task conditions were performed in random order within the same sessions, while the other three conditions were performed in different sessions for each condition (see Materials and Methods for more details). Each session was arranged in the following order: the right-hand motor-related task conditions, i.e., ME, MI, and AfterME, and then OE, CE, and iME. The colors in the matrices represent the distance values, with blue colors representing low difference (i.e., the control properties of the trials are similar) and green to yellowish colors representing high difference, i.e., the control properties of the trials are different. We found that for all conditions, trials within the same sessions were close to each other, as indicated by the relatively small values of

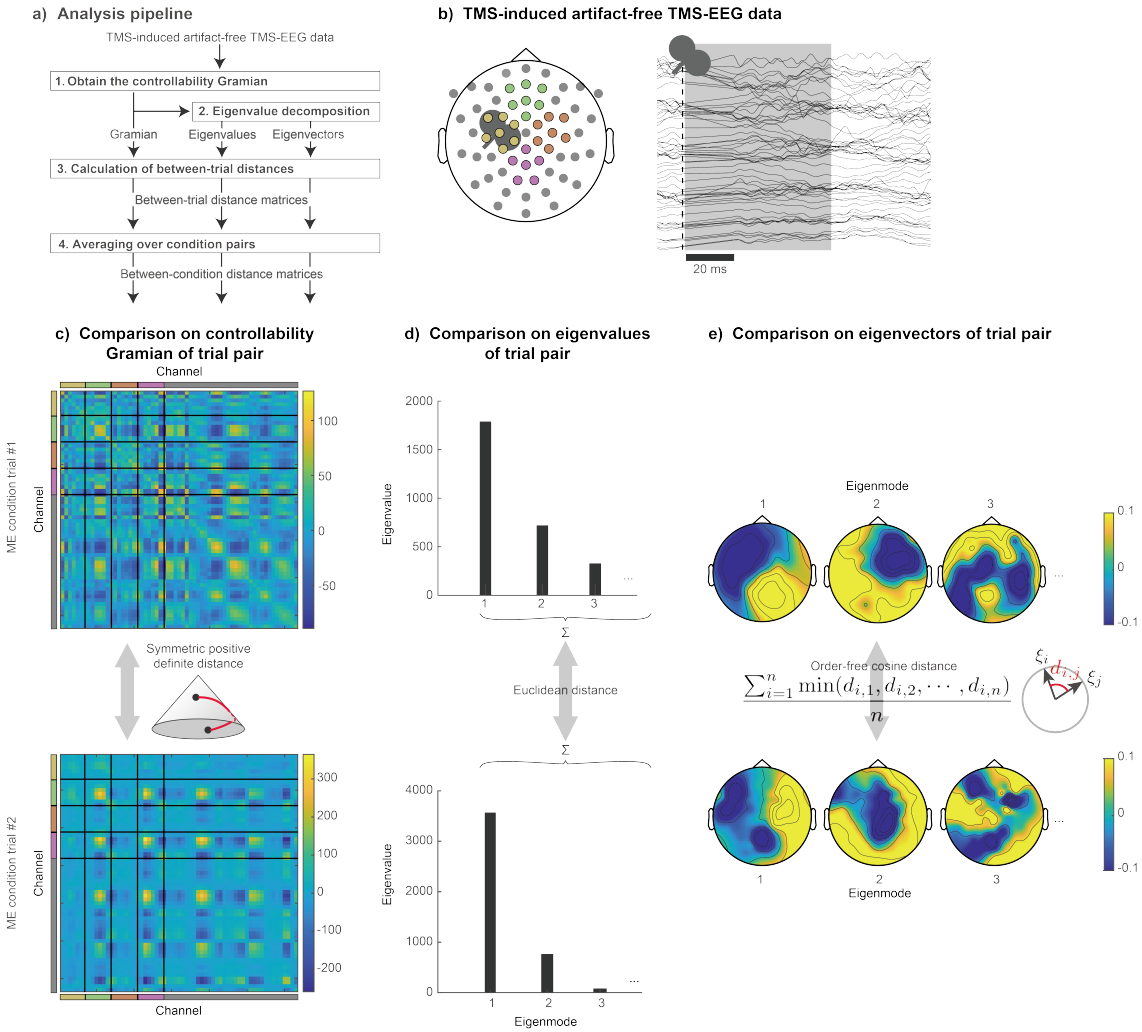

Figure 2: Overview of the proposed methodology and examples single trial analysis results. a) The controllability Gramian is derived from the single-trial TMS-EEG data after removing TMS-induced artifacts. Subsequently, the eigenvalues and eigenvectors of the controllability Gramian are computed. The comparative analysis entails the comparison of controllability Gramians and summation of their eigenvalues, and eigenvectors across each pair of trials to establish the distance matrices (see (c) and Methods). The average distance among trial pairs within the same condition is calculated to generate between-condition distance matrices. b) The left panel displays the electrode arrangement (with colors corresponding to (c)) and a sample waveform from a typical single trial. The temporal segments shaded in gray (1–60 ms from TMS onset) were used for the analysis. Note that the signal from 0 to 20 ms is linearly complemented to remove residual artifacts and the effects of artifact removal. c) The controllablity Gramians obtained from the EEG data of two different trials belonging to the same motor execution (ME) condition, are calculated for the distance in symmetric positive definite space. d) The Euclidean distance of the sums was used as the trial-to-trial distance for the eigenvalues. For the eigenvectors ξ, the effect of order defined by the magnitude of the eigenvalues was ignored, and only the directions were compared. e) Since eigenvectors are represented by relative values scaled between channels, each trial can be represented as n topographies. The distance between the i-th eigenvector of trial #1 ξ_i and the j-th eigenvector of trial #2 ξ_j was weighed by the cosine distance $d_{i,j}$. The distance between trials was the average of the distances of the closest eigenvector pairs (see Methods). Note that the controllability Gramians, eigenvalues, and eigenvectors illustrated in c, d, and e were obtained by the proposed method applied to two typical trials.

the elements of the diagonal blocks corresponding to respective sessions. For a more intuitive understanding of the distances between trials, we have converted the distance matrix shown in Fig. 3a to scatter plots in low-dimensional space with each trial as a point (Fig. 3b and Supp. Fig. S1). Dimension reduction using multidimensional scaling revealed that the two resting conditions (OE and CE) were clearly separated from the other motor-related conditions. In addition, iME was more localized, although it was mixed with the contralateral motor-related condition. It is worth noting that the clusters separated by these conditions span different sessions for each condition (Fig. 3c, numbers indicate sessions); that is, beyond intra-session similarity, clusters were identified by similarity between conditions.

In a second step, with the sum of eigenvalues representing the controllable magnitude, we found that a distinct diagonal block structure in the distance matrix was absent, showing that trials within the same session were no closer to each other than trials from different sessions (Fig. 3d and Supp. Fig. S2), unlike situation for the controllability Gramian. While the sum of eigenvalues tended to be higher on average for the two resting conditions, there was considerable trial-to-trial variation (Fig. 3e). Thus, compared with the controllability Gramians, the sum of eigenvalues alone less well captures the differences between both conditions and sessions.

In a third step, using the eigenvectors we obtained a distance matrix more-or-less similar to that of the controllability Gramian (which is in contrast to the results for the sum of the eigenvalues), i.e., with the characteristic of small values in the diagonal block components of the distance matrix (Fig. 3f). In the reduced-dimensional space achieved by multidimensional scaling, the ability to distinguish between motor-related and rest conditions in the first dimension aligned with that of the controllability Gramians (Fig. 3g and Supp. Fig. S3). However, in the second dimension there was a wide spread of points representing motor-related conditions (cyan, green, purple, and yellow symbols), with no discernible difference from the resting-state conditions (orange and pink symbols). Even within the condition, the trials appeared to be split into multiple subclusters, similar to those found in the controllability Gramian. Both the average distance within a session (ME/MI/AfterME $0.59\pm$ SD 0.01, OE 0.61 ± 0.02, CE 0.62 ± 0.01, iME 0.59 ± 0.01) and between successive trials (ME/MI/AfterME 0.60 ± 0.01, OE 0.61 ± 0.01, CE 0.63 ± 0.01, iME 0.61 ± 0.01) tended to be longer under the CE condition and a Friedman test revealed a significant effect of condition (within session $\chi^2(37) = 3, p < 0.001$, between successive trials $\chi^2(28) = 3, p < 0.001$). These suggest that one of the reasons for the wider distribution or more clusters of motor-related conditions than the resting state conditions is that data were collected in more sessions (ME/MI/AfterME 4.94 ± 0.90, OE 1.28 ± 0.39, CE 1.18 ± 0.39, iME 1.94 ± 0.56) rather than having large within-session or trial-to-trial variability. These subclusters were to some extent cohesive per session, but were not completely bounded. While they do not mix with each other in the OE, CE, and motor-related conditions, they do mix within the same condition across sessions, indicating a clear difference in controllable direction between conditions (Fig. 3h).

The following results were observed. First, there were differences between the resting and motor-related conditions in the Gramians, sum of eigenvalues, and eigenvectors. The sum of eigenvalues was larger in the resting state. Second, Gramians and eigenvectors formed multiple clusters in the lower dimensions. These clusters reflected differences between conditions, which were not solely due to slight variations in the measurements (e.g., electrode impedance

and environmental noise), but also due to variations between the conditions. Third, the sum of eigenvectors showed that conditions were closer to each other across sessions. Subclusters were observed within the same condition, which were not observed in the Gramians. Fourth, the three motor-related conditions were more similar to each other than to the resting state conditions in terms of controllability, as indicated by their non-adjacency to the resting state condition trials in the low-dimensional space of Gramians and eigenvectors. While the influence of measurement states could not be completely eliminated, this finding suggests that the motor-related tasks shared common properties between the measurment sessions.

3.4 Inter-condition distances based on the controllability Gramian

To observe the average trend of the control properties across participants, we next quantified the averaged discriminability between conditions across participants using the distance matrices of the trials. We first averaged the inter-trial distance matrices for each participant across condition pairs and averaged the condition-based distance matrices across participants, obtaining Figs. 4a–4c. The distance matrices are difficult to compare with each other because different distance measures are used for the controllability Gramians, eigenvalues, and eigenvectors. Therefore, we performed six-condition classification using the k nearest neighbor method with each distance measure and obtained classification performance as a common measure (Figs. 4d, e, and f). Additionally, to better illustrate the differences and similarities between conditions, we constructed dendrograms from the distance matrices (Figs. 4g, h, and i).

First, with the controllability Gramians, we found that the three contralateral motor-related conditions (i.e., ME, MI, and AfterME) were indisputably close to each other in the inter-condition distance matrices (Figs. 4a and d). The three conditions related to the right-hand motion were the most similar, with the left-hand motion condition being located between the resting conditions and the other motor-related conditions (Fig. 4g).

Second, using the eigenvalues, we found that the two resting conditions were distant from the other conditions, as was the case with the controllability Gramians (Fig. 4b). However, as seen in Fig. 3e, the variability between trials made it difficult to classify the conditions in each trial (Fig. 4e). The dendrogram based on the distance matrix between conditions showed that iME was most similar to ME, unlike the situation when using the controllability Gramians (Fig. 4h).

Third, using the eigenvectors, we found that in addition to the resting condition, which was clearly distinct from the other conditions in the controllability Gramians, iME also showed a trend toward a greater distance from the other conditions (Fig. 4c). In terms of classification performance, the three right-hand motor-related conditions were indistinguishable from each other, while the other conditions were not misclassified (as with the controllability Gramians) (Fig. 4f). The dendrogram showed that iME, like that of the controllability Gramians, was located between the three right-handed motor-related conditions and the two resting conditions, as was the situation for the controllability Gramians (Fig. 4i).

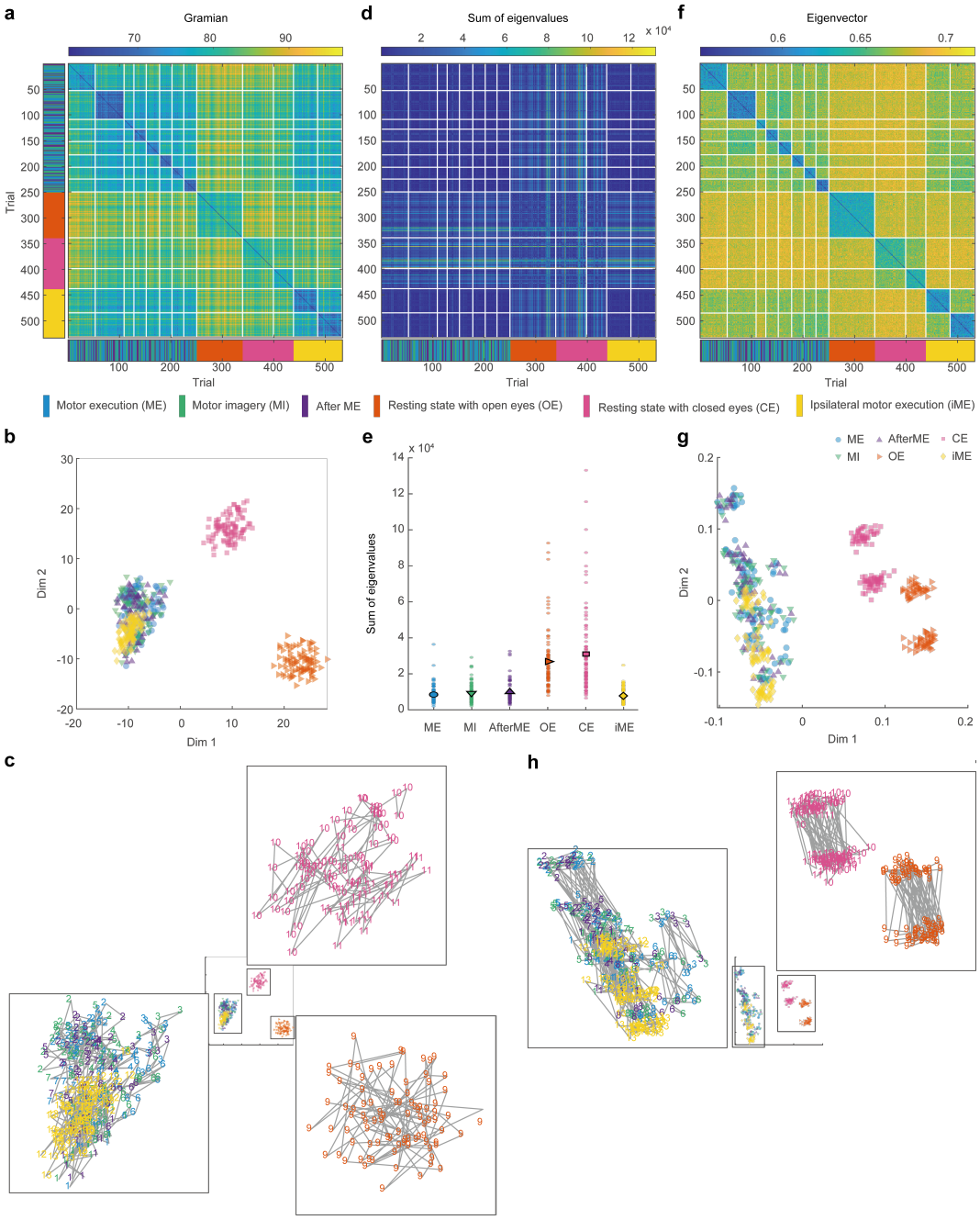

Figure 3: Distances between trials for controllability Gramians, sums of eigenvalues, and eigenvectors for a sample participant. a) A distance matrix of trials for Gramians. Each condition was measured in 100 trials, divided into 1 to 8 sessions. Note that the number of trials used in the analysis differs between conditions because of the noisy trial removal in the preprocessing. White lines indicate session boundaries and the color bar below indicates the condition to which each trial belongs in (b) and (c). b) The distance matrix illustrated in (a) reduced to a lower dimension using multidimensional scaling. Each point represents one trial, with colors corresponding to different conditions in (a). c) To clarify the distances within and between sessions in the same condition, the three clusters in (b) are enlarged. Numbers indicate session indices and colors indicate conditions. Consecutive trials are connected by gray lines. d) A distance matrix of trials for sums of eigenvalues. e) The sum of eigenvalues for each trial is represented as a dot, and the trial average is depicted as a colored symbol. f) A distance map of trials for eigenvalues. g) The distance matrix illustrated in (d) dimensionally reduced. h) The clusters in (g) are enlarged.

Taken together, the results suggest that controllable direction shows greater preservation of the discriminative ability of the controllability Gramians across conditions. Controllable magnitude represents a distinct aspect of variations in brain states to that represented by controllable direction. Ipsilateral and contralateral motion can be distinguished in terms of controllable direction, but not in terms of controllable magnitude. The three distances had one thing in common: four motor-related conditions including iME were closer to each other than 70% of the maximum distance (red lines in Fig. 4g, h, and i). Additional analyses were conducted and are presented in the supplementary material to address potential concerns regarding trial pairs within the same session and the impact of the time window for analysis. These analyses confirm that the observed results are robust, with the concerns identified concerns do not significantly influencing the outcomes.

3.5 Dimensionality of controllable directions

Finally, we examined the dimensionality of the space controllable by the TMS input using the cumulative contribution of the eigenvalues of the controllability Gramian. In short, when a small number of eigenmodes account for a high contribution rate, the dimensionality is considered low, indicating that a few dominant modes can effectively capture the data. By contrast, when a small number of eigenmodes show a low contribution rate, the dimensionality is high, suggesting that the data require a larger number of modes to be accurately represented. As illustrated in the controllability Gramians concept (Fig. 1b), high-dimensional controllability Gramians approximate hyperspheres, while low-dimensional ones exhibit a more flattened hyperellipsoidal form. For example, when the dimension is effectively 2 or 3, Gramians correspond to ellipses or ellipsoids.

We found that in all participants and in all conditions, 90% of the control properties elicited by TMS could be explained by the first to third eigenmode (Fig. 5a), indicating the low dimensionality of controllable space. Although there was no common trend among the participants as to which condition gave the highest contribution to the first eigenmode, we found that across a majority of participants, common trends emerged in contribution rates up to eigenmode 3: the contribution rate of CE was closest to 1, while that of OE was closer to 1 than the motor-related conditions, but there were no clear differences between the contribution rates of the motor-related task conditions (Fig. 5b). These results suggest that the controllability Gramian can be expressed in about 3 out of 63 dimensions in both resting and motor-related conditions, that resting states are lower dimensional than motor-related states, and that closing one's eyes lowers dimensionality even more.

We additionally examined how these few large controllable directions relate to state dependence. Using only the top three eigenvectors, each trial was classified into six conditions, and we found that the correct rate was significantly lower than when all modes were used (Fig. 6). By increasing the number of eigenmodes used to 5, 10, and 30, the correct rate approached that when all modes were used. This suggests that even the lower eigenmodes, which do not provide a high contribution to amplitude restoration, retain information about the differences between conditions.

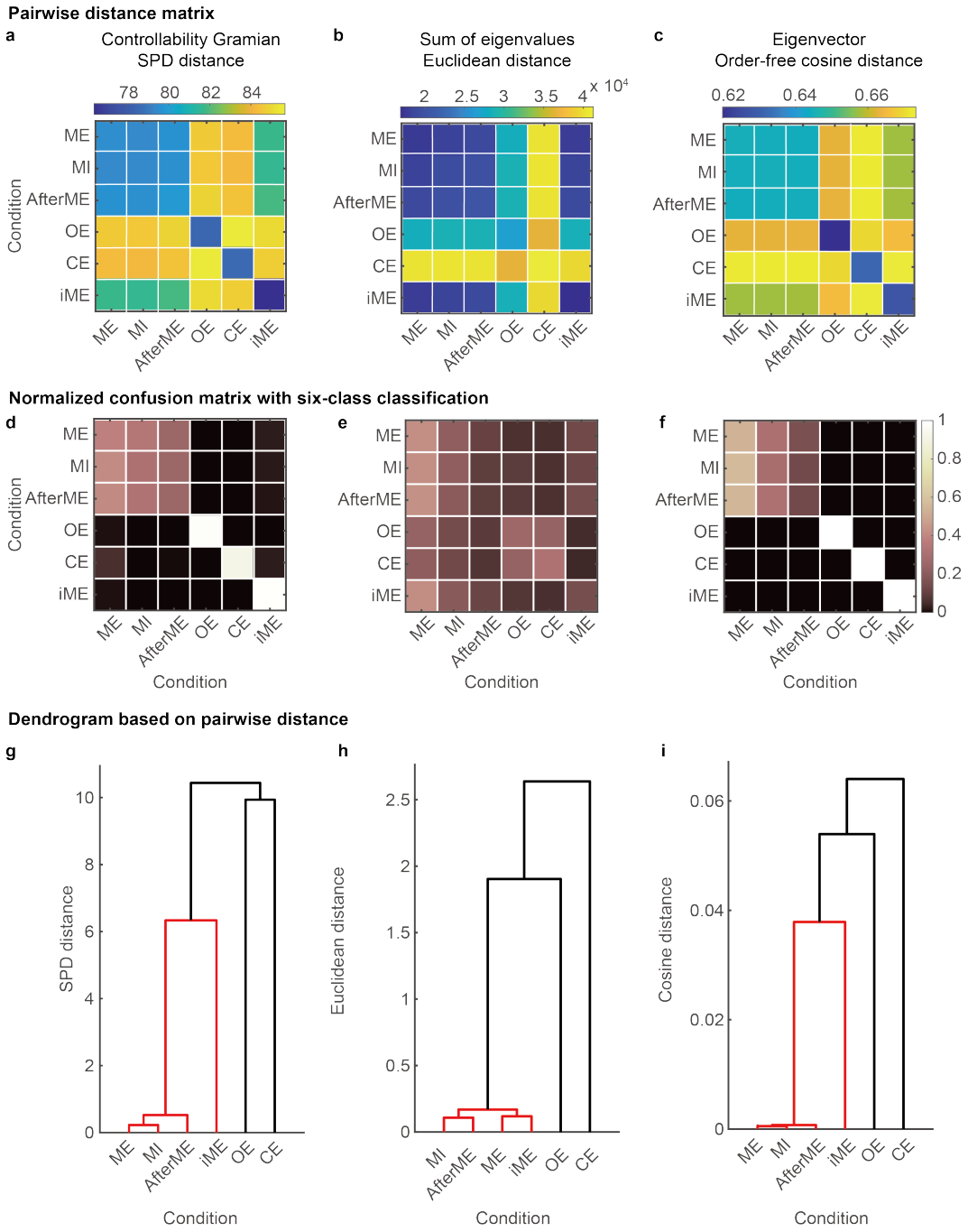

Figure 4: Distance between conditions for the controllability Gramians, sum of their eigenvalues and eigenvectors. a) Distance matrix between-condition pairs for controllability Gramians was averaged over 17 participants. b) Distance matrix between conditions for pairs sums of eigenvalues. c) Distance matrix between conditions for pairs of eigenvectors. d) Normalized confusion matrix for six-class classification using the k-nearest neighbor method (k = 5) in the distance space defined by (a). Rows represent the correct class and the columns the classified class. e) The normalized confusion matrix was obtained from (b). f) The normalized confusion matrix was obtained from (c). g) Dendrograms of the distance space defined by (a). Edges between node pairs that are less than 70% of the maximum linkage distance are indicated by red lines. h) The dendrogram obtained from (b). i) The dendrogram obtained from (c).

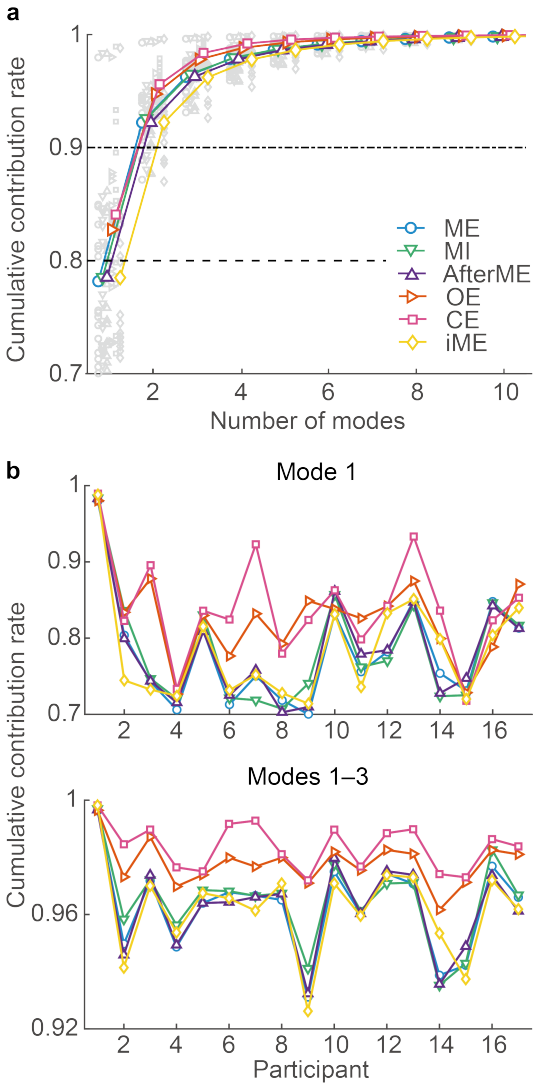

Figure 5: The control properties of M1 under TMS show high cumulative contribution rates for a small number of eigenmodes. a) Colored lines indicate across-participant averages of cumulative contribution rates. The values for each participant are indicated by gray symbols. b) Cumulative contribution rates for the first, and first to third modes. Color and symbols are the same as in a.

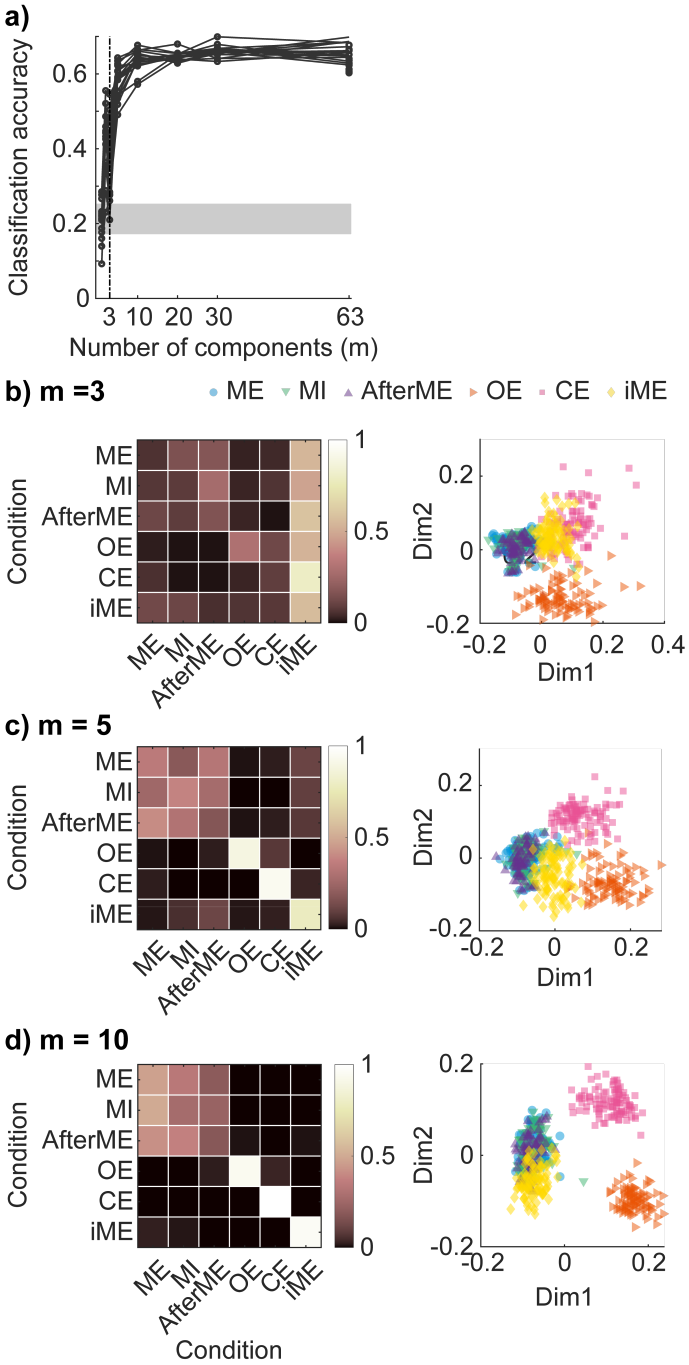

Figure 6: The performance of the six-value classification improved when a fourth or lower mode was added. a) The number of eigenmodes m and the percentage that were classified into the correct condition by the k-nearest neighbor method ($k = 5$). The black line represents the experimental participants, and gray shading represents the correct response rate obtained when predicting the most sampled condition (i.e., the chance level, the lower line represents the minimum between participants and the upper line represents the maximum). b) Confusion matrices (left) when using the first through to third eigenmodes, which account for 90% of the contribution. Classification performance was calculated for each participant and the participant average is shown. Distribution of trials in the classification space for one typical experimental participant (right). This three-dimensional classification space is represented in two-dimensions through the use of multidimensional scaling. Each point represents a trial. Confusion matrices and typical scatter plots when adding the mode used for condition classification from the first to the fifth (c) and first to the tenth (d) eigenmodes are shown.

4 Discussion

In this study, we propose a simple yet theoretically-grounded method based on control theory for analyzing the control properties of brain dynamics under impulse-like brain stimulation. On the basis of the standard results of the optimal control problem, we can regard the time series of neural activity as a controllability matrix, which enables us to simply compute the controllability Gramian without directly inferring the model parameters of linear dynamical systems. We applied our method to single-shot single-trial TMS-EEG data and used the controllability Gramian and its eigenvalues and eigenvectors to examine the differences and similarities in control properties across conditions. These control properties give a reasonable interpretation of responsiveness to local stimuli: the controllability Gramian represents the controllable space with unit input, its eigenvectors represent the controllable direction, and the corresponding eigenvalues represent the controllable magnitude. We found that controllability Gramians estimated from measured signals contains information that allowed distinguishing between open-eye rest, closed-eye rest, contralateral motor-related states, and ipsilateral motor execution state. This distinction is brought about by controllable direction rather than controllable magnitude. The cumulative contribution of eigenvalues indicated that the resting state had less response complexity, regardless of eye opening or eye closing. These results suggest that the state dependence of the EEG response immediately after TMS is characterized by spread rather than amplitude.

4.1 Novel perspectives of the proposed control theory-based measure on perturbation approaches

Although the proposed control theory-based measure is related to some measures quantified in previous studies, its main novel contribution lies in its ability to unify two aspects: the directions in which brain dynamics can be controlled and the extent to which they can be controlled with a fixed amount of input. In terms of relationships with previous studies, excitability and connectivity [Siebner et al., 2009, Bergmann et al., 2016] have previously been quantified as the magnitude of the response activity to a stimulus and the extent to which stimulus-evoked activity propagates [Casali et al., 2010, Casali et al., 2013, Casarotto et al., 2016, Comolatti et al., 2019]. The excitability and connectivity quantified in these previous studies are related to the sum of eigenvalues of the controllability Gramian in our framework, which quantifies the overall degree of propagation of the stimulus-evoked response. However, the eigenvector we focused on indicates the specific direction of this propagation. The novelty of our proposed method is that it quantifies the degree and direction of propagation represented by the eigenvalues and eigenvectors of the controllability Gramian, considering them as being completely separable. In previous explorations of control properties from measurement signals, only the eigenvalues were utilized and the corresponding eigenvectors were neglected [Gu et al., 2015], or they were used as eigenmodes without being separated [Comolatti et al., 2019]. In this study, we showed that differences s between conditions clearly appear in controllable directions rather than in controllable magnitudes.

4.2 Limitations of our controllability Gramian approach

The limitations of our network control theory-based approach stem from assuming a first-order linear AR model without noise. While this simplifies the controllability analysis and avoids the complexities of estimating connectivity

(A in Eq. (1)) and input matrices (B), these assumptions are not fully realistic. First, in terms of model order, EEG responses often require higher-order AR models. Studies suggest an optimal order between 8.6 and 30, depending on the context [Tseng et al., 1995, Chang et al., 2012, Shakeel et al., 2021]. Extending the analysis to a higher-order model would more closely represent the real measurement data, but this would require estimating A and B, which are not needed in the proposed method. Second, regarding the linearity assumption, neural dynamics are in general nonlinear and thus there may be cases where nonlinear models are more appropriate. Note, however, that a previous study demonstrated that linear AR models remain reasonable approximations of macroscopic brain dynamics because of observational and technical limitations [Nozari et al., 2024]. Third, neural dynamics are in general stochastic, and thus accounting for stochasticity could change the prediction of control property models . In our previous work, we explored controllability in stochastic neural systems [Kamiya et al., 2023], and extending the current framework to incorporate stochastic systems is a potential avenue for future research, but this would substantially complicate the estimation of controllability. In addition, the proposed method is currently limited to impulse inputs; addressing more diverse input types would require explicit estimation of the connectivity and input matrices.

Despite these limitations, we found that the Gramian estimated using the simplified model effectively captures state-dependent changes in controllability. A key advantage of this approach is that it simplifies the process by eliminating the need to estimate the connectivity matrix A and input matrix B (Eq. (1)), thereby enabling direct estimation of controllability Gramians. This simplicity makes the method especially suitable for researchers unfamiliar with the framework of control theory, providing a way to easily evaluate the control properties of complex brain dynamics that are typically difficult to assess. Future research should prioritize refining the estimation process by accounting for more realistic factors, such as nonlinearity and noise, to enhance the applicability and accuracy of our proposed method.

4.3 Validity of the results in comparison with previous studies

We found that the controllable magnitude fluctuates more during rest conditions than during motor-related conditions. One possible explanation is the reduced variability in brain activity induced by the visual stimuli used for task instructions, which stabilizes neural activity by constraining it to a specific trajectory [He and Zempel, 2013, He, 2011]. In addition, both motor execution and motor imagery enhance inter-trial coherence within the consistent activity of task-positive networks such as the somatomotor network [Churchland et al., 2010]. By contrast, during resting states, neural networks are less constrained by external demands, allowing greater flexibility and spontaneous variation between interconnected networks. This network-level flexibility may account for the higher trial-to-trial variability observed in resting-state EEG.

Our analysis further reveals that controllable directions retain more information for identifying states than do controllable magnitudes. Controllable direction is a relevant indicator of EEG spatial topography because it captures the relative relationships between EEG channels, rather than the absolute amplitude of each channel. Of note, the right-hand motor-related conditions and the two resting-state conditions formed distinct clusters in analysis based on controllable

directions. Although the specific contributions of motor execution and imagery to these separations remain unclear, the distinction aligns with established findings, such as characteristic EEG topographies for visual evoked potentials and differences between open-eye and closed-eye resting states [Allison et al., 1977, Barry et al., 2007]. These findings confirm that motor-related states with visual instructions, open-eye resting states, and closed-eye resting states can be distinguished by their controllable directions.

We also found differences in motor execution between right and left hands. This is not surprising considering the stream of previous brain-computer interface studies taking advantage of topographic differences caused by differences in effectors [Xu et al., 2011, Huang et al., 2009], but it should be noted that our method does not take into account the sign of the eigenvector. In other words, the left and right-hand movements are not due to a simple lateral reversal of topography, suggesting that different neural activation patterns occur. The influence of afferent sensation on controllability could not be clearly concluded in this study; ME and iME could be distinguished, but not ME, MI, or AfterME. This may be related to the fact that motor imagery causes M1 excitability similar to that caused by motor execution [Takemi et al., 2013].

Taken together, our approach reveals that some brain states, such as motor-related task versus resting states, can be well distinguished by the properties of the controllability Gramian, whereas other states, such as motor execution and motor imagery, cannot be so differentiated. Future work is needed to verify which factors of brain states are reflected in the controllability associated with this TMS-EEG framework and which factors are not, with this involving performing TMS-EEG under more diverse conditions.

4.4 Considerations for Evaluating Control properties with TMS-EEG

In this study, the preprocessing steps used in TMS-EEG analysis could have influenced the results, particularly in terms of distinguishing between conditions. One such preprocessing method is ICA, which can affect the analysis by removing certain components. EEG signals consist of three types of components: (1) components that are common across states, (2) components that differ between states, and (3) noise. If components from category (2) are inadvertently removed, the ability to detect differences between states may be compromised. Therefore, it is crucial to carefully evaluate preprocessing methods to selectively extract components from category (2). Further research is needed to refine these approaches.

Filtering is widely used in EEG analysis and can significantly impact the results. While it helps reduce the influence of extracerebral sources, it may also eliminate signals that are critical for identifying specific states. Therefore, further refinement of preprocessing parameters is necessary to achieve a balance between selectively removing unwanted signals and preserving relevant state-specific information.

A key consideration for future research is the relationship between control properties and physiological indices. It is possible that the controllable direction is more closely associated with measures such as MEP, heart rate, and respiration rather than overall brain state, potentially leading to the emergence of subclusters within conditions. In

this study, higher MEP values were observed in the ME, which is inherently linked to the other two motor-related conditions. This suggests that MEPs do not directly contribute to distinguishing between conditions but warrant further investigation.

Similarly, the relationship between controllability and traditional EEG features is also of interest. Given that TEP amplitude is a measure similar to controllable magnitude, it may have limited utility in differentiating between states. However, examining finer-grained information, such as latency, could provide deeper insights and should be explored in future studies.

References

[Allison et al., 1977] Allison, T., Matsumiya, Y., Goff, G. D., and Goff, W. R. (1977). The scalp topogrphy of human visual evoked potentials. *Electroencephalogr. Clin. Neurophysiol.*, 42(2):185–197.

[Awiszus, 2003] Awiszus, F. (2003). TMS and threshold hunting. *Suppl. Clin. Neurophysiol.*, 56:13–23.

[Barry et al., 2007] Barry, R. J., Clarke, A. R., Johnstone, S. J., Magee, C. A., and Rushby, J. A. (2007). EEG differences between eyes-closed and eyes-open resting conditions. *Clin. Neurophysiol.*, 118(12):2765–2773.

[Beck et al., 2024] Beck, M. M., Heyl, M., Mejer, L., Vinding, M. C., Christiansen, L., Tomasevic, L., and Siebner, H. R. (2024). Methodological choices matter: A systematic comparison of TMS-EEG studies targeting the primary motor cortex. *Hum. Brain Mapp.*, 45(15):e70048.

[Bergmann et al., 2016] Bergmann, T. O., Karabanov, A., Hartwigsen, G., Thielscher, A., and Siebner, H. R. (2016). Combining non-invasive transcranial brain stimulation with neuroimaging and electrophysiology: Current approaches and future perspectives. *Neuroimage*, 140:4–19.

[Bhatia, 2009] Bhatia, R. (2009). *Positive definite matrices*. degruyter.com.

[Biabani et al., 2021] Biabani, M., Fornito, A., Coxon, J. P., Fulcher, B. D., and Rogasch, N. C. (2021). The correspondence between EMG and EEG measures of changes in cortical excitability following transcranial magnetic stimulation. *J. Physiol.*, 599(11):2907–2932.

[Biabani et al., 2019] Biabani, M., Fornito, A., Mutanen, T. P., Morrow, J., and Rogasch, N. C. (2019). Characterizing and minimizing the contribution of sensory inputs to TMS-evoked potentials. *Brain Stimul.*, 12(6):1537–1552.

[Brunton and Kutz, 2019] Brunton, S. L. and Kutz, J. N. (2019). *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press.

[Casali et al., 2010] Casali, A. G., Casarotto, S., Rosanova, M., Mariotti, M., and Massimini, M. (2010). General indices to characterize the electrical response of the cerebral cortex to tms. *Neuroimage*, 49(2):1459–1468.

[Casali et al., 2013] Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., Casarotto, S., Bruno, M.-A., Laureys, S., Tononi, G., and Massimini, M. (2013). A theoretically aased index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.*, 5(198):198ra105.

[Casarotto et al., 2016] Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., Pigorini, A., G Casali, A., Trimarchi, P. D., Boly, M., Gosseries, O., Bodart, O., Curto, F., Landi, C., Mariotti, M., Devalle, G., Laureys, S., Tononi, G., and Massimini, M. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. *Ann. Neurol.*, 80(5):718–729.

[Chang et al., 2012] Chang, J.-Y., Pigorini, A., Massimini, M., Tononi, G., Nobili, L., and Van Veen, B. D. (2012). Multivariate autoregressive models with exogenous inputs for intracerebral responses to direct electrical stimulation of the human brain. *Front. Hum. Neurosci.*, 6:317.

[Churchland et al., 2010] Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., Priebe, N. J., Finn, I. M., Ferster, D., Ryu, S. I., Santhanam, G., Sahani, M., and Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.*, 13(3):369–378.

[Comolatti et al., 2019] Comolatti, R., Pigorini, A., Casarotto, S., Fecchio, M., Faria, G., Sarasso, S., Rosanova, M., Gosseries, O., Boly, M., Bodart, O., Ledoux, D., Brichant, J.-F., Nobili, L., Laureys, S., Tononi, G., Massimini, M., and Casali, A. G. (2019). A fast and general method to empirically estimate the complexity of brain responses to transcranial and intracranial stimulations. *Brain Stimul.*, 12(5):1280–1289.

[Conde et al., 2019] Conde, V., Tomasevic, L., Akopian, I., Stanek, K., Saturnino, G. B., Thielscher, A., Bergmann, T. O., and Siebner, H. R. (2019). The non-transcranial TMS-evoked potential is an inherent source of ambiguity in TMS-EEG studies. *Neuroimage*, 185:300–312.

[Gu et al., 2017] Gu, S., Betzel, R. F., Mattar, M. G., Cieslak, M., Delio, P. R., Grafton, S. T., Pasqualetti, F., and Bassett, D. S. (2017). Optimal trajectories of brain state transitions. *Neuroimage*, 148:305–317.

[Gu et al., 2015] Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Yu, A. B., Kahn, A. E., Medaglia, J. D., Vettel, J. M., Miller, M. B., Grafton, S. T., and Bassett, D. S. (2015). Controllability of structural brain networks. *Nat. Commun.*, 6:8414.

[He, 2011] He, B. J. (2011). Scale-free properties of the functional magnetic resonance imaging signal during rest and task. *J. Neurosci.*, 31(39):13786–13795.

[He and Zempel, 2013] He, B. J. and Zempel, J. M. (2013). Average is optimal: an inverted-U relationship between trial-to-trial brain activity and behavioral performance. *PLoS Comput. Biol.*, 9(11):e1003348.

[Herring et al., 2015] Herring, J. D., Thut, G., Jensen, O., and Bergmann, T. O. (2015). Attention modulates TMS-locked alpha oscillations in the visual cortex. *J. Neurosci.*, 35(43):14435–14447.

[Huang et al., 2009] Huang, D., Lin, P., Fei, D.-Y., Chen, X., and Bai, O. (2009). Decoding human motor activity from EEG single trials for a discrete two-dimensional cursor control. *J. Neural Eng.*, 6(4):046005.

[Jakowluk, 2024] Jakowluk, W. (2024). *Optimal input signal design in control systems identification*. BUT Publishing House.

[Kamiya et al., 2023] Kamiya, S., Kawakita, G., Sasai, S., Kitazono, J., and Oizumi, M. (2023). Optimal control costs of brain state transitions in linear stochastic systems. *J. Neurosci.*, 43(2):270–281.

[Kawakita et al., 2022] Kawakita, G., Kamiya, S., Sasai, S., Kitazono, J., and Oizumi, M. (2022). Quantifying brain state transition cost via schrödinger bridge. *Netw Neurosci*, 6(1):118–134.

[Lee et al., 2022] Lee, M., Sanz, L. R. D., Barra, A., Wolff, A., Nieminen, J. O., Boly, M., Rosanova, M., Casarotto, S., Bodart, O., Annen, J., Thibaut, A., Panda, R., Bonhomme, V., Massimini, M., Tononi, G., Laureys, S., Gosseries, O., and Lee, S.-W. (2022). Quantifying arousal and awareness in altered states of consciousness using interpretable deep learning. *Nat. Commun.*, 13(1):1064.

[Li et al., 2023] Li, Q., Yao, L., You, W., Liu, J., Deng, S., Li, B., Luo, L., Zhao, Y., Wang, Y., Wang, Y., Zhang, Q., Long, F., Sweeney, J. A., Gu, S., Li, F., and Gong, Q. (2023). Controllability of functional brain networks and its clinical significance in First-Episode schizophrenia. *Schizophr. Bull.*, 49(3):659–668.

[Medaglia et al., 2017] Medaglia, J. D., Pasqualetti, F., Hamilton, R. H., Thompson-Schill, S. L., and Bassett, D. S. (2017). Brain and cognitive reserve: Translation via network control theory. *Neurosci. Biobehav. Rev.*, 75:53–64.

[Nozari et al., 2024] Nozari, E., Bertolero, M. A., Stiso, J., Caciagli, L., Cornblath, E. J., He, X., Mahadevan, A. S., Pappas, G. J., and Bassett, D. S. (2024). Macroscopic resting-state brain dynamics are best described by linear models. *Nat Biomed Eng*, 8(1):68–84.

[Okazaki et al., 2020] Okazaki, Y. O., Mizuno, Y., and Kitajo, K. (2020). Probing dynamical cortical gating of attention with concurrent TMS-EEG. *Sci. Rep.*, 10(1):1–10.

[Oldfield, 1971] Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, 9(1):97–113.

[Oostenveld et al., 2011] Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.*, 2011: 156869.

[Rocchi et al., 2021] Rocchi, L., Di Santo, A., Brown, K., Ibáñez, J., Casula, E., Rawji, V., Di Lazzaro, V., Koch, G., and Rothwell, J. (2021). Disentangling EEG responses to TMS due to cortical and peripheral activations. *Brain Stimul.*, 14(1):4–18.

[Rogasch et al., 2017] Rogasch, N. C., Sullivan, C., Thomson, R. H., Rose, N. S., Bailey, N. W., Fitzgerald, P. B., Farzan, F., and Hernandez-Pavon, J. C. (2017). Analysing concurrent transcranial magnetic stimulation and electroencephalographic data: A review and introduction to the open-source TESA software. *Neuroimage*, 147:934–951.

[Rose et al., 2016] Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., and Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354(6316).

[Shakeel et al., 2021] Shakeel, A., Onojima, T., Tanaka, T., and Kitajo, K. (2021). Real-Time implementation of EEG oscillatory Phase-Informed visual stimulation using a least mean Square-Based AR model. *J Pers Med*, 11(1).

[Siebner et al., 2009] Siebner, H. R., Bergmann, T. O., Bestmann, S., Massimini, M., Johansen-Berg, H., Mochizuki, H., Bohning, D. E., Boorman, E. D., Groppa, S., Miniussi, C., Pascual-Leone, A., Huber, R., Taylor, P. C. J., Ilmoniemi, R. J., De Gennaro, L., Strafella, A. P., Kähkönen, S., Klöppel, S., Frisoni, G. B., George, M. S., Hallett, M., Brandt, S. A., Rushworth, M. F., Ziemann, U., Rothwell, J. C., Ward, N., Cohen, L. G., Baudewig, J., Paus, T., Ugawa, Y., and Rossini, P. M. (2009). Consensus paper: combining transcranial stimulation with neuroimaging. *Brain Stimul.*, 2(2):58–80.

[Takemi et al., 2013] Takemi, M., Masakado, Y., Liu, M., and Ushiba, J. (2013). Event-related desynchronization reflects downregulation of intracortical inhibition in human primary motor cortex. *J. Neurophysiol.*, 110(5):1158–1166.

[Tseng et al., 1995] Tseng, S. Y., Chen, R. C., Chong, F. C., and Kuo, T. S. (1995). Evaluation of parametric methods in EEG signal analysis. *Med. Eng. Phys.*, 17(1):71–78.

[Xu et al., 2011] Xu, P., Yang, P., Lei, X., and Yao, D. (2011). An enhanced probabilistic LDA for multi-class brain computer interface. *PLoS One*, 6(1):e14634.

[You and Park, 2021] You, K. and Park, H.-J. (2021). Re-visiting riemannian geometry of symmetric positive definite matrices for the analysis of functional connectivity. *Neuroimage*, 225:117464.