NeuroImage 245 (2021) 118652

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Causal decoding of individual cortical excitability states ☆

J. Metsomaa [a,b], P. Belardinelli [a,b,c], M. Ermolova [a,b], U. Ziemann [a,b,*], C. Zrenner [a,b,d]

[a] Department of Neurology & Stroke, University of Tübingen, Tübingen, Germany
[b] Hertie Institute for Clinical Brain Research, University of Tübingen
[c] CIMeC, Center for Mind-Brain Sciences, University of Trento, Italy
[d] Temerty Centre for Therapeutic Brain Intervention, Centre for Addiction and Mental Health, and Department of Psychiatry, University of Toronto, Toronto, ON, Canada

ARTICLE INFO

Keywords:
TMS
EEG
Excitability
Brain state
Classification
Machine learning

ABSTRACT

Brain responsiveness to stimulation fluctuates with rapidly shifting cortical excitability state, as reflected by oscillations in the electroencephalogram (EEG). For example, the amplitude of motor-evoked potentials (MEPs) elicited by transcranial magnetic stimulation (TMS) of motor cortex changes from trial to trial. To date, individual estimation of the cortical processes leading to this excitability fluctuation has not been possible.

Here, we propose a data-driven method to derive individually optimized EEG classifiers in healthy humans using a supervised learning approach that relates pre-TMS EEG activity dynamics to MEP amplitude. Our approach enables considering multiple brain regions and frequency bands, without defining them a priori, whose compound phase-pattern information determines the excitability.

The individualized classifier leads to an increased classification accuracy of cortical excitability states from 57% to 67% when compared to μ-oscillation phase extracted by standard fixed spatial filters. Results show that, for the used TMS protocol, excitability fluctuates predominantly in the μ-oscillation range, and relevant cortical areas cluster around the stimulated motor cortex, but between subjects there is variability in relevant power spectra, phases, and cortical regions. This novel decoding method allows causal investigation of the cortical excitability state, which is critical also for individualizing therapeutic brain stimulation.

1. Introduction

A ubiquitous property of the brain, apparent in non-invasive brain-stimulation with transcranial magnetic stimulation (TMS) (Hallett 2007; Barker and Jalinous, 1985), is the large variability in the response evoked by the identical repeated stimulus (Kiers et al., 1993). To the degree that such variability is explained by the dynamics of ongoing neural activity (Arieli et al., 1996), the distribution of evoked response amplitudes reflects the distribution of fluctuating cortical excitability states. In this view, variability is not a "bug" but a "feature" in that it allows us to probe different instantaneous brain states. In TMS of primary motor cortex (M1), the amplitude of the motor-evoked potential (MEP) is an indicator of corticospinal excitability at the time that the TMS pulse was applied. By combining TMS of M1 with electroencephalography (EEG), we can group trials post-hoc into high- and low-excitability trials (classes) based on MEP amplitude and ask whether a feature of the EEG signal immediately preceding the respective stimulus can predict between the two classes for an individual subject. This approach is known as "decoding" and it allows a causal investigation of the brain state relevant to the evoked response.

In the primate, the phase of distally recorded local field potentials was shown to predict local spike rates (Canolty et al., 2010), and the phase of invasively recorded sensorimotor μ-rhythm correlates with neuronal spiking (Haegens et al., 2011). EEG oscillations reflect the synchronized activity of different neuronal populations, similar to intracranial invasive extracellular recordings (Buzsaki et al., 2012). Thus, the identification of fluctuating cortical excitability states from ongoing EEG data is a mapping from macroscopic bioelectric phenomena to the depolarization state of specific neuronal cell populations.

It is however not clear how to best extract this information from a window of EEG data for a specific cortical location and subject. Previous attempts to find a relationship between pre-stimulus EEG and MEP amplitude were based on pre-specified assumptions on the locations and/or temporal filters which were expected to yield the relevant predictive EEG features (Schaworonkow et al., 2019; Torrecillos et al. 2020; Maki and Ilmoniemi 2010; Zrenner et al., 2018). These studies have reported

☆ Postal address: Department of Neurology & Stroke, Hertie Institute for Clinical Brain Research, University of Tübingen, Hoppe-Seyler-Straße 3, 72076 Tübingen, Germany.
* Corresponding author.
E-mail address: ulf.ziemann@uni-tuebingen.de (U. Ziemann).

https://doi.org/10.1016/j.neuroimage.2021.118652.
Received 25 June 2021; Received in revised form 30 September 2021; Accepted 11 October 2021
Available online 21 October 2021.

variable results on how MEP amplitude was found to be linked with the phase of the pre-stimulus EEG data on average. Absence of phase relation with the MEP amplitude or very small effect size have also been reported (Mitchell et al., 2007; Madsen et al. 2019; van Elswijk et al. 2010).

It is an open question whether the predictive power of the EEG signal to detect cortical excitability states is inherently limited, or whether individualized modeling with fewer *a priori* assumptions could improve the MEP prediction accuracy. In practice, spatial and temporal filters can serve to extract relevant phase patterns from EEG. Previous attempts to individualize these filters based on the properties of the pre-stimulus EEG signal were moderately successful (Schaworonkow et al., 2018), whereby signal-to-noise ratio (SNR) in the μ-alpha range was optimized in order to increase the accuracy of phase estimation (Zrenner et al., 2020). Here, we utilize the post-stimulus MEP that was "caused" by an interaction between the pre-stimulus EEG signal and the TMS pulse to decode cortical excitability states from pre-stimulus EEG using supervised machine learning approaches: MEP amplitudes are used to label the trials post-hoc as either high or low excitability and the pre-stimulus EEG data is then used to train and to test the classifier.

Specifically, we use logistic regression to compute an optimal linear spatio-temporal pre-stimulus EEG filter extracting the excitability score, separately for each individual subject. Our model assumes that the amplitudes of the underlying cortical activations at specific latencies within a window prior to the TMS pulse are predictive of cortical excitability at the time of the stimulus. This assumption corresponds to instantaneous phase being predictive of instantaneous cortical excitability.

The proposed methodology is highly suitable for studying the excitability state encoding under different measurement settings and conditions since no prior assumptions regarding relevant locations, frequencies, or optimal oscillation phases are required. We tested the classification approach with measured EEG–TMS–MEP datasets to evaluate the classification accuracy of the individually estimated model, and this was compared to a non-individualized model where the same predefined spatial filter was used across all subjects.

Supervised machine learning approaches have previously been used in brain–computer interfaces (BCI), to distinguish between different task-related brain states, such as motor imagery, based on the amplitude of different brain oscillations estimated over several seconds (Blankertz et al., 2011; Blankertz et al., 2007). Our scenario differs in that we seek to estimate rapidly fluctuating instantaneous states, that are encoded by the phase of oscillations, from ongoing EEG preceding a stimulus event. Furthermore, we lack true knowledge of each trial's class membership, as the mapping from cortical excitability state to MEP amplitude is noisy and fluctuates on different time scales. Finally, resting-state is not a single stationary process, but comprises a wide range of spatio-temporal EEG patterns, which can emerge and vanish over the course of the measurement (Betzel et al., 2016).

An important aspect of our decoding approach is that the resulting classifiers are physiologically interpretable: Each decoding filter corresponds to a cortical activity cascade preceding the individual optimal high-excitability state, and the conversion into such a spatio-temporal EEG pattern has been shown to be unique (Haufe et al., 2014). The individual cortical source activity underlying the resulting EEG patterns can be estimated using MRI-guided inverse modeling to localize and characterize the dynamics of cortical activity reflecting instantaneous excitability states. Source modeling also enables verification of neurophysiological plausibility of the machine learning results and allows us to assess the variability of cortical dynamics underlying excitability states across study participants.

To address the experimental uncertainty of labeling individual trials as belonging to either a high-or low-excitability class based on MEP amplitude, we simulated different scenarios with synthetic data (EEG and MEPs) to test how errors in the labeling affect the classification accuracy. The simulation also allowed us to vary EEG signal-to-noise ratio and the sample size to test how sensitive the model is to changes in these parameters.

One motivation for this study is that the accurate identification of cortical excitability states is relevant for improving the efficacy of brain-state-dependent therapeutic brain stimulation. Because the induction of synaptic plasticity depends on excitability of the stimulated circuits, fluctuating excitability states present temporal windows that determine the direction and magnitude of plastic changes induced by stimulation (Zrenner et al., 2018; Baur et al., 2020). To explore the potential of the proposed methodology for online usage, we tested how stable the brain states appear during one experiment and across separate experiments by testing the classifier accuracy over time and subjects.

2. Methods

2.1. Experimental methods

2.1.1. Participants

The study was approved by the local ethics committee at the medical faculty of the University of Tübingen (protocol 716/2014BO2) and conducted in accordance with the Declaration of Helsinki. Eight right-handed adults (5 female, 3 male; age 23.5 ± 3.3 years) with no history of neurological and/or psychiatric pathologies were enrolled and completed the study. All participants provided written informed consent before participation.

2.1.2. Neuroimaging

Anatomical T1-weighted MRI data was acquired on a separate session for the generation of realistic head models (3T Siemens Prisma with a 32-channel head coil, GRE pulse sequence, TE 2.22 ms, TR 2400 ms, FA 8°, FoV 256, Phase FoV 93.8%).

2.1.3. Experimental setup

A concurrent EEG–TMS setup was employed for the experiment. A TMS stimulator (PowerMAG Research 100, MAG & More, Munich, Germany) was configured to deliver biphasic pulses through a passively cooled figure-of-eight coil (PMD70-pCool, 70 mm winding diameter, MAG & More, Munich, Germany). EEG and electromyography (EMG) were recorded simultaneously at a sampling rate of 5 kHz using a 24-bit biosignal amplifier in DC mode (NeurOne, Bittium, Oulu, Finland). MEPs were recorded from the abductor pollicis brevis (APB) and first dorsal interosseous (FDI) muscles of the right hand in a bipolar belly-tendon montage using adhesive hydrogel electrodes (Kendall, Covidien, Dublin, Ireland). EEG was recorded using a TMS-compatible 128-channel cap with Ag/AgCl sintered ring electrodes (EasyCap BC-TMS-128, EasyCap, Herrsching, Germany), placed according to the International 10–5 system.

Head position was maintained using a vacuum pillow (Vacuform, Salzbergen, Germany), the TMS coil was positioned using a mechanical arm (Fisso, Baitella, Zürich, Switzerland). A stereoscopic neuronavigation system (Localite, Bonn, Germany) was used to co-register the participant's head to an individual MR image, to pinpoint the locations of the EEG electrodes on the scalp, and to real-time monitor coil positioning throughout the experiment.

2.1.4. Experimental procedures

The hand representation of left M1 was targeted orienting the coil such that the strongest field was induced in a posterior-lateral to anterior-medial direction. The motor hotspot was defined as a position and orientation of the coil resulting in the largest MEP amplitudes in the right APB or FDI. Resting Motor Threshold (RMT) was defined as the minimum stimulation intensity eliciting MEPs with peak-to-peak amplitudes > 50 μV in 50% of test stimulation pulses (Groppa et al., 2012).

2.1.5. Experimental session

The study consisted of a single experimental session for each participant, lasting about 3 h. After preparation of EEG, EMG, and neuronavigation and pinpointing the EEG electrodes, the hotspot location was determined. Participants were seated comfortably and instructed to fixate a visual target (fixation cross approximately 1 m in front of them). 1000 single TMS pulses were then applied in one run with an interstimulus interval of 2 ± 0.25 s at a stimulation intensity of 110% RMT. An additional 5 min resting-state EEG recording was obtained immediately before and after the stimulation.

2.1.6. Head model generation and computing the EEG forward model

Individual realistic head models were generated to estimate the anatomical locations of the sources of EEG activity recorded at the scalp. Source estimation was not employed for computing EEG filters used in the decoding algorithm itself, but for the localization and interpretation of the resulting spatio-temporal EEG patterns, to address the question which brain areas are involved in which temporal order in the high-excitability brain state.

Head models were used for both forward and inverse solutions. For the cortical mesh representing the source space, the T1-weighted MRI volume was realigned and resliced to "anterior commissure–posterior commissure" (ACPC) space, the skull was stripped and a source model based on a cortical sheet was created with FreeSurfer (Dale et al., 1999), and the HCP-workbench (Van Essen et al. 2013). The resulting cortical mesh was downsampled and registered to a common spherical template, allowing for one-to-one comparisons of activations on the same vertex across subjects. For each subject, a mesh with 15,684 vertices was created.

A 3-shell boundary element model (BEM) was then created to compute the forward model between source mesh and the EEG sensors: The T1-weighted MRI volume was segmented using the SPM toolbox (https://www.fil.ion.ucl.ac.uk/spm/), resulting in surfaces for the scalp, as well as the outer and inner skull. Standard conductivity values (skin, bone, brain: 0.33 S/m, 0.0041 S/m, 0.33 S/m) were used for the lead-field computation. EEG electrode alignment was performed manually based on anatomical fiducials using Fieldtrip (Oostenveld et al., 2011).

2.2. Preprocessing the electrophysiological data

2.2.1. Toolboxes

Preprocessing and data analysis were performed using Matlab (The Mathworks, Inc., Natick, MA, USA), the EEGLAB toolbox (Version 14.1.2b) (Delorme and Makeig, 2004), the FastICA toolbox (Hyvarinen, 1999), as well as custom scripts. The analysis pipeline was designed to run automatized with minimal user input to make the preprocessing systematic and consistent across all datasets.

2.2.2. EMG preprocessing

The goal of the EMG preprocessing was to clean the signal such that the peak-to-peak MEP amplitudes at the APB and FDI right hand muscles could be accurately extracted in each trial also for small responses: EMG data was epoched using EEGLAB in time windows of [−0.5 s, 0.5 s] relative to the TMS pulse. Slow drifts were eliminated by Laplacian trendline fitting and 50 Hz noise was removed in each trial individually by projecting out two 50 Hz sinusoidal fixed-amplitude waveforms with 90° phase difference (see Appendix A.1 for details).

Trials containing EMG activity in the period of [−300 ms, −5 ms] before the TMS pulse indicative of pre-innervation or other EMG artifacts were removed by visually inspecting all trials, with a range exceeding 20 μV. The TMS-related decay artifact was then removed from the EMG signal by fitting an exponential curve to the signal in each EMG channel and trial separately (see Appendix A.2 for details). Finally, the peak-to-peak MEP amplitude was determined in each trial and channel as the range of the EMG signal within the manually defined data interval containing the MEPs. Each MEP was individually inspected.

2.2.3. EEG preprocessing

EEG data was extracted from the pre-stimulus window between [−1000 ms, −4 ms] and downsampled to a sampling frequency of 1 kHz using the EEGLAB toolbox. Slow trends were then removed by applying a Laplacian-based trend detection (see Appendix A.1 for details).

To identify noisy channels and trials, we estimated the channel-wise uncorrelated noise signals using the DDWiener method (Mutanen et al., 2018). The standard deviation of the noise in each channel and trial was computed to estimate the respective noise level. To avoid biased noise-level estimates, we normalized them by dividing by the noise-level estimates obtained using the spherical-head-model lead-field matrix (see Appendix A.3) as input data for DDWiener. The median of these standard deviations was computed excluding the values above the 98th percentile. If the noise level in a channel exceeded 10 times the median in over 10% of the trials, the channel was considered poor and removed from further analysis. On average, 30 channels were excluded per subject, with the standard deviation of 5, the minimum of 17, and maximum of 34. We also computed the range of the noise signal in each channel and trial and took the maximum range across all channels in each trial to detect high-amplitude baseline changes in the signal. Trials were excluded if this measure exceeded two times the median (computed excluding values above the 98th percentile). On average, 176 trials were removed from each data set. After rejecting poor channels and trials, the average reference was applied over all channels using the remaining channels.

Independent component analysis (ICA) was then applied using the FastICA algorithm in its symmetric mode and with the 'tanh' contrast function to estimate 35 independent components within the subspace spanned by the 35 largest principal vectors (Hyvarinen, 1999). We visually identified the eye movement components and other outlier components and removed them from the data (median 3.5 components, range [2, 6]). Epochs were then low-pass filtered at 35 Hz with zero-phase response using a Hamming window-based FIR filter of order 20.

An alternative parallel run of ICA was performed with strict cleaning to study the generalizability of the classification over time and subjects assuming stationary subject-wise noise. To mimic the stationarity, we removed all noisy (spiky) components whose power was distributed unevenly across the trials. Such components, representing non-stationary noise, would deteriorate the classification even if excitability encoding patterns were unchanged. On average, 20 components (range, 17–23) were rejected. Downsampling to 200 Hz was applied after low-pass filtering at 35 Hz to reduce the computational times of the forthcoming analyses.

For cross-subject analysis, only the channels accepted for all subjects were preserved. This choice was made after discovering that even a greatly reduced number of channels was enough for the classification at individual level. The power level of the data from each subject, estimated as the variance of the EEG over all channels and samples, was normalized to unit size to prevent the estimation from biasing due to different individual EEG power levels.

Lastly, EEG and EMG data were brought into correspondence and trials where either part was affected by artifacts were removed, resulting in a set of clean pre-stimulus EEG data windows with corresponding post-stimulus MEP amplitudes.

2.3. Description of the decoding algorithm

2.3.1. Goal

We aimed at predicting cortical excitability level based on a period of pre-stimulus EEG data, optimized on an individual basis. Our hypothesis was that a linear spatio-temporal filter could capture the essential information of the underlying cortical excitability state from an epoch of resting-state EEG. This filter is able to approximate the brain state defined by the instantaneous phases or shapes of the underlying cortical source waveforms. The entire EEG/EMG analysis pipeline from the

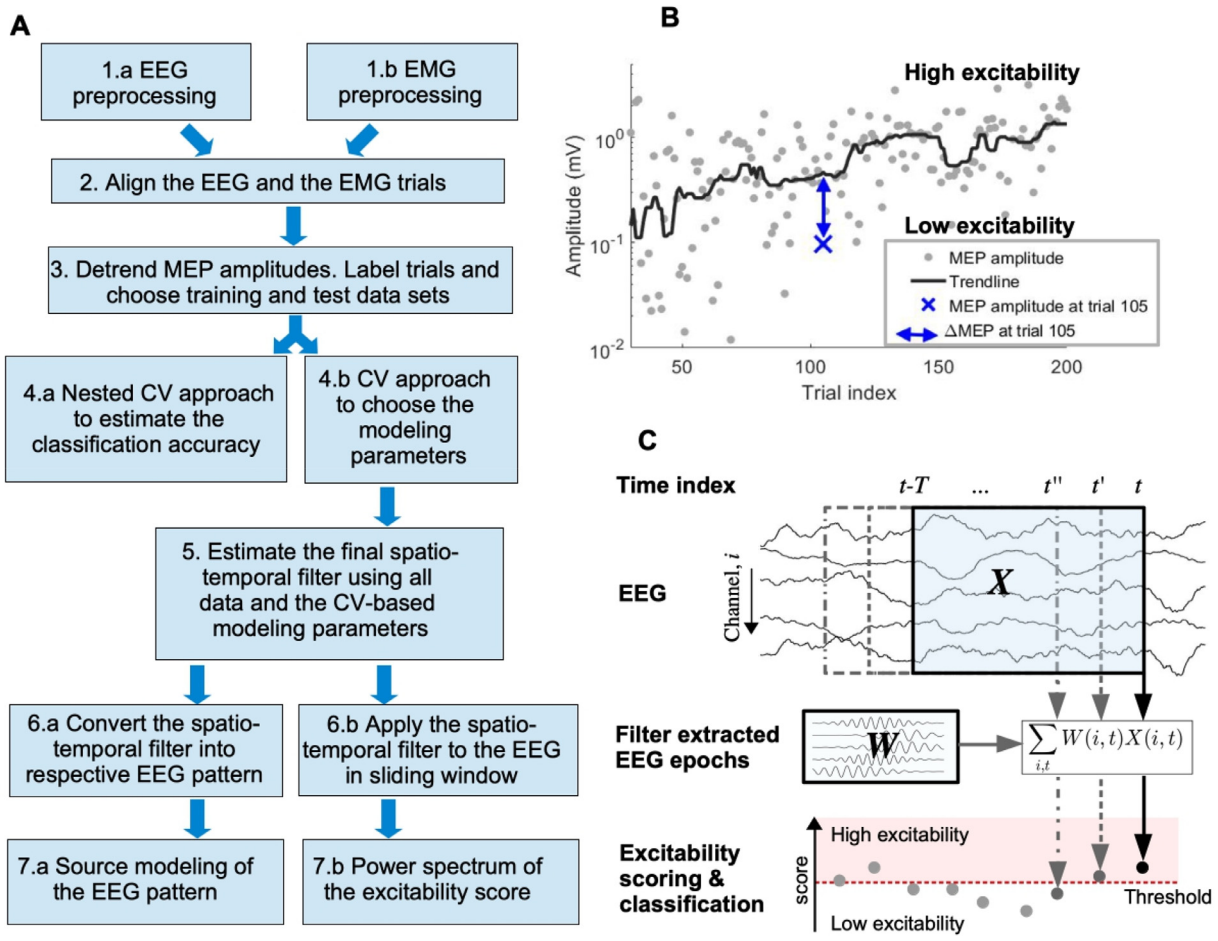

Fig. 1. A. Analysis pipeline. First (1.a and 1.b), EEG and EMG preprocessing protocols are separately run, after which the corresponding EEG and EMG trials are matched, and the leftover data trials are removed in step 2. MEP amplitudes are corrected by detrending over successive trials by step 3. The trials are labeled, and the test data and training data sets are defined according to the detrended MEP amplitudes to high and low excitability states. Logistic regression was used to train a classifier in two parallel analysis paths (4.a and b): To evaluate the classification accuracy, in step 4.a, the nested cross-validation scheme is adapted for choosing the features and testing the data in nested iterations. In step, 4.b, all the data is used in a cross-validation process to define the features. Thereafter, all the data is used to train the model with the chosen features in step 5, which yields a spatiotemporal filter. The filter can be transferred into the respective spatio-temporal pattern to reveal the pattern which most likely represents the high-excitability state for the given data in step 6.a. In step 7.a, source localization is used to estimate the corresponding cortical activation cascade. After applying the spatio-temporal filter to the resting-state EEG (step 6.b), the power spectrum of the fluctuating brain state can also be estimated (step 7.b). B. The trend of the MEPs (black) is estimated by median filtering. Subtracting the trend from the individual MEP amplitudes (gray) yields the variation signal $\boldsymbol{\Delta MEP}$. The labels are assigned to high- or low-excitability states according to positive and negative variations, respectively. C. Matrix filtering in a sliding window yields the excitability score as a function of time.

preprocessing, through the filter estimation, up until the model interpretation is illustrated in Fig. 1A.

With a linear model, it is possible to highlight the spatio-temporal EEG pattern representing the highest excitability in an individual data set. This particular EEG pattern represents the data with the underlying collection of sources at optimal phases with respect to the high-excitability state. Furthermore, the opposite phase information (180° shifted phases of the underlying sources) represents the lowest excitability. The filter measures the similarity of a measured EEG data epoch and the optimal spatio-temporal pattern, while simultaneously suppressing the irrelevant parts of the data.

Here, MEP amplitudes were used as a read-out of the instantaneous corticospinal excitability state, and we used this knowledge to train the predictive filter in a supervised way. Because MEP amplitudes have multiple modulators along the motor tracts up until the EMG recording setting, we did not aim at predicting the exact amplitude itself, but rather at predicting whether a measured amplitude is higher or lower than its expected value based on the preceding and following MEP sizes in the measurement. Therefore, we trained a binary classifier which would allocate pre-stimulus EEG epochs into either high- or low-excitability states (classes).

We first evaluated the benchmarking classification results when using minimal cleaning of the data and including the most reliably labeled trials in the classifier training and testing. These results were compared against those obtained from the fixed-spatial-filter paradigm. We then proceeded to study the generalizability of the classification from three perspectives, which would be relevant for online usage, *i.e.,* to what extent the ML-trained classifier could be generalized 1) across different subjects, 2) over time of individual measurement sessions, and 3) when decreasing the number of channels.

2.3.2. *Defining excitability states based on EMG*

We classified EEG epochs (trials) into two classes: 0 represents low cortical excitability state, and 1 represents high cortical excitability state. To use a supervised learning algorithm to train a classifier and to test its predictive power, we assigned the trials into these states by EEG-independent means. MEP amplitudes were used as the read-out of the instantaneous excitability state. Basically, large and small MEPs were

interpreted to represent high and low corticospinal excitability states, respectively. It is noteworthy that this labeling could not be considered entirely accurate but still informative of the actual excitability states.

During a long measurement, slow fluctuations were observed in the MEP amplitudes in the course of the trials. This phenomenon has been observed previously with low-frequency single-pulse TMS, and might be attributed to cumulative effects of low-frequency TMS on corticospinal excitability (Pellicciari et al., 2016), or infraslow fluctuation of cortical activity (de Goede and van Putten, 2019). The slowly changing MEP trendline was estimated with a median filter of order 20, and this trend was subtracted from the individual amplitudes (Fig. 1B). The resulting positive and negative difference MEPs (ΔMEP) were interpreted to represent high and low corticospinal excitability states, respectively. Before filtering, on both edges of the signal, the signal was mirrored to provide enough samples for the median filter.

Because the labeling could not be considered fully accurate, we picked a subset of the trials, which were considered most accurately labeled, for training and testing the classifier. The detrended MEPs, $\Delta\mathrm{MEP}_i$, were computed for both measured EMG (*i.e.*, APB, FDI) channels $i = 1, 2$, after which the trials, where the signs of the detrended MEPs were opposite for the two channels, were rejected from further modeling. We further calculated the joint amplitude of both EMG channels as $\mathrm{sign}(\Delta\mathrm{MEP}_1)\sqrt{(\Delta\mathrm{MEP}_1)^2 + (\Delta\mathrm{MEP}_2)^2}$, after which the epochs with 150 smallest and 150 largest joint amplitudes were preserved for further analysis.

For the brain-state stability evaluation, we picked 200 largest and smallest joint amplitudes which were picked as evenly as possible throughout the recording. To this end, normalization of the joint amplitudes was performed to keep their variance steady: In a sliding window of 50 consecutive trials, and moving forward one trial at a time, the MEP amplitudes were normalized by dividing them by their standard deviation. To prevent very high outlier amplitudes from biasing the normalization, MEPs, whose amplitudes exceeded three times the standard deviation of all MEPs were excluded from the normalization.

2.3.3. Spatial and temporal filtering of EEG epochs

We aimed to decode the cortical excitability state at a given time instant based on the EEG epoch recorded from M channels and T time points preceding the time instant. The following linear model was assumed to give rise to the (M x T) EEG epoch $\mathbf{X}$:

$$\mathbf{X} = \sum_i \mathbf{a}_i \mathbf{s}_i^{\mathrm{T}}, \tag{1}$$

where (M x 1) $\mathbf{a}_i$ is the topography of the ith underlying source, and (T x 1) $\mathbf{s}_i$ is the respective waveform (time course). Here, the summation in Eq. (1) goes over all neural, artefactual, and noise sources. Furthermore, the physical or the statistical properties of the sources are not restricted, and the topographies/waveforms in Eq. (1) can be either the true or the estimated ones.

When applying a spatial filter $\mathbf{w}_{\mathrm{s}}$ to the measured EEG epoch, we get a T x 1 waveform vector $\mathbf{s}$ representing the time series of the extracted phenomenon:

$$\mathbf{s}^{\mathrm{T}} = \mathbf{w}_{\mathrm{s}}{}^{\mathrm{T}}\mathbf{X}\,. \tag{2}$$

Optimally, the spatial filter has zero inner product with all irrelevant EEG signals as described by their topographies, *i.e.,* $\mathbf{w}_{\mathrm{s}}{}^{\mathrm{T}}\mathbf{a}_i = 0$ for all irrelevant sources i in Eq. (1).

Applying then a (T x 1) temporal filter $\mathbf{w}_{\mathrm{t}}$ to the waveform means computing the inner product $\mathbf{s}^{\mathrm{T}}\mathbf{w}_{\mathrm{t}}$. This filtering can reveal temporal dynamics of the time course, *e.g.,* phase or shape information of the source waveform.

We used the spatial and temporal filters to get a value which is maximal at the highest excitability state, while the minimum is reached at the lowest excitability state. We refer to this value as the excitability index. If a single known source, whose waveform is captured by $\mathbf{w}_{\mathrm{s}}$, is encoding the excitability state based on its instantaneous phase/shape information, highlighted by filter $\mathbf{w}_{\mathrm{t}}$, we can measure the excitability index I_{BS} by

$$I_{\mathrm{BS}} = \mathbf{w}_{\mathrm{s}}{}^{\mathrm{T}}\mathbf{X}\mathbf{w}_{\mathrm{t}}. \tag{3}$$

Normalizing I_{BS} with respect to the power of the waveform $\|\mathbf{s}\| = \|\mathbf{w}_S^T\mathbf{X}\|$ yields $\|I_{\mathrm{BS}}\| = \mathbf{w}_s^T\mathbf{X}\mathbf{w}_{\mathrm{t}}/\|\mathbf{w}_s^T\mathbf{X}\|$, which is needed if the power is varying over the epochs, but it is not predictive of the excitability. To summarize, the spatial filter can reveal the source waveform, while the temporal filter can decode the relevant instantaneous phase information in the signal. In the following, we assume that each data epoch has been normalized with respect to power. How to perform normalization in practice is discussed later.

A spatio-temporal matrix filter is obtained by combining the spatial and temporal filters into a matrix filter $\mathbf{W} = \mathbf{w}_{\mathrm{s}}\mathbf{w}_{\mathrm{t}}{}^{\mathrm{T}}$, after which the brain state index in Eq. (3) is obtained by

$$I_{\mathrm{BS}} = \sum_{i,j} \mathrm{W}_{i,j}\mathrm{X}_{i,j} = \mathbf{W}, \mathbf{X}_{\mathrm{F}} = \mathrm{vec}(\mathbf{W})^{\mathrm{T}}\mathrm{vec}(\mathbf{X}), \tag{4}$$

where $\cdot, \cdot_{\mathrm{F}}$ denotes the Frobenius inner matrix product, *i.e.,* summation over all matrix elements after entry-wise product, and $\mathrm{vec}(\cdot)$ denotes vectorization of the matrix by concatenating all of its columns. For brevity, we use the Frobenius inner matrix product notation when applying such a spatio-temporal filter.

If the sign of the waveform is reversed in Eqs. (2) and (3), I_{BS} gets opposite sign. Changing the sign corresponds to flipping the phase content by 180°, which according to our hypothesis, represents the opposite excitability state. Maximal and minimal excitability indices thus encode maximal and minimal excitability states, respectively.

2.3.4. Spatio-temporal filtering of the EEG epochs

We expect that multiple spatio-temporal filters describing a multi-source model are needed for excitability state prediction. Thus, we extend the single-source model in Eqs. (3) and (4) by applying additional filter pairs to measure the excitability index of the EEG epoch as

$$I_{\mathrm{BS}} = \sum_i \left(\mathbf{w}_{\mathrm{s},i}\right)^{\mathrm{T}}\mathbf{X}\mathbf{w}_{\mathrm{t},i}. \tag{5}$$

Since the data epoch $\mathbf{X}$ is fixed in the summation of Eq. (5), we may write the summation of the filtered signals more compactly as

$$I_{\mathrm{BS}} = \sum_i \mathbf{w}_{\mathrm{s},i}\left(\mathbf{w}_{\mathrm{t},i}\right)^{\mathrm{T}},\ \mathbf{X}_{\mathrm{F}} = \mathbf{W}, \mathbf{X}_{\mathrm{F}}, \tag{6}$$

where the filter pairs have been summated into the spatio-temporal filter $\mathbf{W} = \sum_i \mathbf{w}_{\mathrm{s},i}(\mathbf{w}_{\mathrm{t},i})^{\mathrm{T}}$ before applying them to the EEG epoch.

Not knowing how many relevant patterns there are, we establish a real-valued $(M \times T)$ matrix filter $\mathbf{W}$, which can be applied to the EEG epoch as $I_{\mathrm{BS}} = \mathbf{W},\ \mathbf{X}_{\mathrm{F}}$. The goal is to estimate $\mathbf{W}$ such that I_{BS} is predictive of the excitability (MEP amplitude) without constraining the amount of filters a priori.

The spatio-temporal filter can be applied to the resting-state EEG at any given time instant by simply extracting an $(M \times T)$ epoch from the measured data preceding the time of interest. Applying the filter by Eq. (6) to the epoch yields the excitability index at the chosen time. The sliding-window approach, where the extracted epoch is shifted forward one time instant at a time, can be used to follow the instantaneous excitability as a function of time (Fig. 1C).

2.3.5. Logistic regression classifier

Logistic regression estimates the probability of a data sample to belong to class 1 in a binary classification problem, with classes 0 and 1 (Kleinbaum et al., 2002). This probability is assumed to follow a sigmoid curve as a function of the sum of linearly weighted input variables:

$$\mathrm{p}(y = 1|\boldsymbol{\varphi}, \boldsymbol{z}) = \frac{1}{1 + \exp(-\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{z} + b)} = \frac{1}{1 + \exp(-\mathrm{score} + b)}, \tag{7}$$

where $\boldsymbol{z}$ is an $(N \times 1)$ input data vector to be classified, $\boldsymbol{\varphi}$ is the $(N \times 1)$ weight vector describing the model, and b sets the probability of the

model when the score, $\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{z}$, equals zero. When the weights and b are known, the classification of input data $\boldsymbol{z}$ is performed based on the probability given by Eq. (7). We classify the trials into class 1 if the probability exceeds 0.5, *i.e.,* score $> b$, and into class 0 otherwise. The logistic regression model expects a monotonic growth of the probability as the score in Eq. (7) increases.

The weights and b in the logistic regression classifier are estimated with R samples of training data $\mathbf{z}^{(i)}$ and their known classes $\mathrm{y}^{(i)}$, $i = 1, \ldots, R$. Using the probability model of Eq. (7), the likelihood function of the training data can be formulated as

$$\mathrm{p}\left(\mathrm{y}^{(1)}, \ldots, \mathrm{y}^{(R)} | \boldsymbol{\varphi}, \mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(R)}\right) = \prod_{i \in \text{class } 1} \left(\frac{1}{1+\exp(-\boldsymbol{\varphi}^{\mathrm{T}}\mathbf{z}^{(i)}+b)}\right) \prod_{i \in \text{class } 0} \left(\frac{1}{1+\exp(\boldsymbol{\varphi}^{\mathrm{T}}\mathbf{z}^{(i)}+b)}\right). \quad (8)$$

The model parameters are chosen such that they maximize Eq. (8). For this purpose, we used the Matlab inbuilt function 'lassoglm'; the details are described below.

In our application, the input data were EEG data matrix entries. To use EEG epoch as an input to the logistic regression in Eq. (7), we set $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{X})$. Furthermore, writing for the classifier weights $\boldsymbol{\varphi} = \mathrm{vec}(\boldsymbol{W})$ and comparing to Eq. (4), we see that score in Eq. (7) has the same definition as the brain excitability index I_{BS}. Therefore, we used logistic regression to estimate the spatio-temporal filter $\boldsymbol{W}$ for extracting the excitability state.

We used the logistic regression paradigm with the lasso regularization, which minimizes the L^1- norm of the weights simultaneously with optimizing the model (Tibshirani, 1996). To set the regularization constant, which defines the balance between minimizing the norm and fitting the model weights to the training data, 5-fold cross-validation was applied. In practice, the 'lassoglm' function was used with options: 'distribution': 'binomial', 'alpha': 1, 'CV': 5, 'lambda': $[10^{-2}, \ 10^{-1}, \ldots, \ 10^{2}]$, 'standardize': 1. The last parameter means that all the input variables were set to zero mean and unit variance before model estimation to avoid biased modeling.

2.3.6. Dimensionality reduction for classification

Optimizing a weight for every channel and time entry within an EEG epoch requires estimating channels x time points of unknowns, equaling to 30,000 optimizable parameters if 100 channels and 300 time points are used. This number of unknowns greatly exceeds the number of training data trials, which leads to inaccurate estimates. However, the individual entries in an EEG epoch are highly correlated, which means that we can reduce the dimensionality of the data and the optimized parameters.

We used linear feature extraction such that each feature z_i was obtained by $z_i = \boldsymbol{D}_i, \boldsymbol{X}_{\mathrm{F}}$, where $\boldsymbol{D}_i$ is the ith feature extraction matrix. The benefit of such linear feature extraction is that the score obtained from the logistic regression can still be interpreted as the brain excitability index (Eq. (6)): Setting the features as input into Eq. (7), the score is obtained by

$$\text{score} = \sum_i \varphi_i z_i = \sum_i \varphi_i \boldsymbol{D}_i, \boldsymbol{X}_{\mathrm{F}} = \sum_i \varphi_i \boldsymbol{D}_i, \ \boldsymbol{X}_{\mathrm{F}} = \sum_{\mathrm{i}} \varphi_{\mathrm{i}} \mathbf{D}_{\mathrm{i}}, \quad \mathbf{X}_{\mathrm{F}} = \boldsymbol{W}, \boldsymbol{X}_{\mathrm{F}} = I_{\mathrm{BS}} \ . \quad (9)$$

We identify the spatio-temporal filter in Eq. (9) as $\boldsymbol{W} = \sum_i \varphi_i \boldsymbol{D}_i$, which relation can be used to convert the feature weights and the feature extraction matrices into the spatio-temporal filter applicable to the original EEG epochs.

The common spatial pattern (CSP) approach was used to find spatial filters that highlight the relevant activity whose phase-locked pattern is predictive of the brain state. CSP can be used to estimate spatial filters which pick up high-power waveform in data set 1 while the power in data set 2 is low (Blankertz et al., 2007). More specifically, the spatial filters $\boldsymbol{w}$ can be found at the extrema of

$$\text{relative power} = \boldsymbol{w}^{\mathrm{T}}\mathbf{C}_1\boldsymbol{w}/\boldsymbol{w}^{\mathrm{T}}\mathbf{C}_2\boldsymbol{w}, \quad (10)$$

where $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are the covariance matrices of the data sets 1 and 2, respectively. In practice, the solution is based on computing the covariance matrix of the two data sets, and solving the generalized eigenvalue decomposition of these matrices: The spatial filters, chosen among the eigenvectors, extract CSP waveforms corresponding to specific spatial patterns $\boldsymbol{a}$ as in Eqs. (1)–(2).

To reveal the phase-locked signal with respect to the TMS onset, we computed the mean over 25 randomly chosen high-excitability pre-TMS EEG epochs and the mean of 25 random low-excitability epochs. Computing the means suppresses the ongoing EEG signals whose phases are irrelevant to the excitability, which increases the SNR of the relevant phase-locked activity. The means were then subtracted from each other. As we assume that these two states are presented in opposite phases, the high-excitability pattern is highlighted by the subtraction. The subtraction was repeated 400 times, always randomizing new epochs to be averaged within each state among the training data. $\mathbf{C}_1$ was computed as the sample covariance matrix over these 400 surrogate epochs representing high-excitability data.

The other covariance matrix $\mathbf{C}_2$ was estimated using all of the training data, representing various phase patterns and other ongoing brain activity. To prevent overfitting in the CSP solution, ridge regularization was used by adding a stabilizing scaled identity matrix to the second covariance matrix by $\boldsymbol{C}_2 \leftarrow \boldsymbol{C}_2 + \boldsymbol{I}\gamma \ \mathrm{tr}(\boldsymbol{C}_2)/M$, where $\mathrm{tr}(\cdot)$ denotes the trace operation, and M is the number of channels (Lotte and Guan, 2011). Using these covariance matrices as input to Matlab's 'eig' function, CSP was used to find three spatial filters $\{\boldsymbol{w}_{\mathrm{spat},1}, \ \boldsymbol{w}_{\mathrm{spat},2}, \ \boldsymbol{w}_{\mathrm{spat},3}\}$, which most effectively captured the signal in the averaged epochs as compared to non-averaged epochs, *i.e.,* which yielded highest relative power values according to Eq. (10).

The waveforms corresponding to the three filters still had altogether 903 correlated samples in each epoch. To further decrease the dimensionality, we organized the three waveforms in each epoch i into a long vector by $\boldsymbol{s}_3{}^{(i)} = \mathrm{vec}([\boldsymbol{w}_{\mathrm{spat},1}, \boldsymbol{w}_{\mathrm{spat},2}, \boldsymbol{w}_{\mathrm{spat},3}]^{\mathrm{T}}\boldsymbol{X}^{(i)})$ and ran principal component analysis (PCA) using these vectors from R training epochs $[\boldsymbol{s}_3{}^{(1)}, \ldots, \boldsymbol{s}_3{}^{(R)}]$. The first N principal vectors $\{\boldsymbol{u}_1, \ \boldsymbol{u}_2, \ldots, \ \boldsymbol{u}_N\}$, as obtained by Matlab's 'svd' function, were chosen for dimensionality reduction: In each trial, the combined waveforms of the three spatial filters were projected onto the principal vectors, yielding N features capturing the phase information of the trial in a the feature vector $\boldsymbol{z} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N]^{\mathrm{T}}\boldsymbol{s}_3{}^{(i)}$. This feature vector was taken as input data into logistic regression Eq. (7).

To combine the above CSP and PCA steps, the ith spatio-temporal feature extraction matrix in Eq. (9) was obtained by

$$\boldsymbol{D}_i = \boldsymbol{W}_{\mathrm{spat}}\tilde{\boldsymbol{U}}_i = \left[\boldsymbol{w}_{\mathrm{spat},1}, \boldsymbol{w}_{\mathrm{spat},2}, \boldsymbol{w}_{\mathrm{spat},3}\right] \begin{bmatrix} \boldsymbol{u}_{i,1:T} \\ \boldsymbol{u}_{i,T+1:2T} \\ \boldsymbol{u}_{i,2T+1:3T} \end{bmatrix}, \quad (11)$$

where $\boldsymbol{u}_{i,n:m}$ is the subvector of principal vector i extracted from index n until m. Before training the classifier, the feature vectors $\boldsymbol{z}$ were normalized such that their L^2-norms were set to 1. The normalization aims to prevent the classifier from learning based on the power of the training data epochs.

The model had two hyperparameters describing dimensionality reduction, the CSP regularization factor and the number of features, which had to be determined before computing the features for training the logistic regression classification model. We used cross-validation to decide these hyperparameters as explained in detail in Section 'Testing and Cross-validating the Classifier'.

2.3.7. Reducing the number of channels

It was further investigated how the accuracy changes if the number of channels is reduced. The accuracy was evaluated for six sets of electrodes including 20, 30, 40, 50, 60, and 80 channels. The channels were

selected individually in an iterative process of removing channels one at a time by the following procedure: In each iteration, the channel which had least contribution in the CSP step was removed:

$$\text{removed channel} = \underset{i\in\text{channels}}{\arg\min}\left[\max_{k\in\{1,2,3\}} |(w_{\text{spat},\,k})_i \cdot (a_{\text{spat},\,k})_i|\right], \tag{12}$$

where $w_{\text{spat},\,k}$ and $a_{\text{spat},\,k}$ are the kth spatial filter and pattern from CSP (see previous Section), respectively, and subscript i denotes picking the respective entry. The heuristically chosen measure to be minimized in Eq. (12) is small if either the CSP filter weight or the topography amplitude is small (in comparison to the other one) across the spatial filters/patterns for channel i.

By Eq. (12), the channel with the smallest 'importance' was removed, CSP was rerun with the remaining channels, and the next channel was eliminated for the subsequent CSP round until the chosen number of channels was reached. Finally, the features were computed by Eqs. (9), (11), and the classifier training was completed.

2.3.8. Testing and cross-validating the classifier

Nested cross-validation (CV) process was used to train, test, and validate the model in two nested iteration loops. The inner loop was used to test the model with several hyperparameters to choose the optimal ones, while the outer loop was used to validate the model with the chosen hyperparameters. See, *e.g.,* (Varoquaux et al., 2017) for details of the nested CV protocol. For both tuning the hyperparameters and validating the model, we used accuracy to measure the model performance. Accuracy was computed as the number of correctly classified samples divided by the total number of classified samples.

Since labeling the excitability states based on MEP amplitudes was not accurate, we selected the most reliably labeled trials for testing and validating the model. A subset of 50 out of the 300 trials with the 25 smallest and 25 largest MEP difference values were used as the pool of trials for testing and validating. The outer validation loop was performed in a 5-fold manner, meaning that 10 out of the total of 50 trials were kept apart for validating the trained model, while the remaining trials were utilized in the training/testing. The inner test loop was performed in a 4-fold manner, again holding 10 samples at once for testing with various hyperparameter values.

In the test iterations, to select the hyperparameters, we first fixed the number of features N to eight (in Eqs. (9) and (11)), trained the classifier in each round using each of the CSP regularization factors ($\gamma = \{10^{-4},\ 10^{-3},\ \ldots,\ 10^{2}\}$), and checked the accuracy with the test data. After running through the CV process with all the lambda values, we chose that γ which yielded the maximal average accuracy. Then, using this optimized γ, the number of features was set to $N = \{2,\ 4,\ 8,\ 20\}$. Test data were again used to pick N with the maximal average accuracy.

The hyperparameters were determined in the inner test loop for each round of the outer validation loop. Final model accuracy was evaluated in the validation loop by computing the average classification accuracy of the validation data over the five iterations.

We also analyzed what the individually estimated spatio-temporal filters could tell us about the neurophysiological processes encoding the brain excitability. This analysis was performed separately from the accuracy evaluation. For this purpose, we omitted the outer loop in the nested CV process and implemented a simple 5-fold CV process to estimate the hyperparameters using the pool of 50 most reliable trials for accuracy estimation in a similar setting as explained above. The final filter was estimated using all data and the chosen hyperparameters.

2.3.9. Testing and cross-validating the stability of the classification

For individual brain-state stability evaluation, the training set was chosen among the first half of recorded data and the test data among the latter half. Training set size was set at 200 because only half of the trials were available. To have a balanced training set, it comprised the first 100 high-excitability and 100 low-excitability state trials from the set of 400 largest and smallest normalized joint amplitudes, respectively.

The test set comprised 50 trials, with the 25 largest and 25 smallest normalized joint-MEP amplitudes among the latter half of the 400 trials and non-overlapping with the training set. A simple two-fold CV-strategy was used, such that the total test set was divided into equal-sized subsets for CV and testing, where the first subset was used to define the hyperparameters yielding the best prediction accuracy, while the latter one was then used for evaluating the prediction accuracy. The CV and test subset roles were then switched to complete the accuracy estimation over the whole test set.

We also tested the applicability of the classifier to a new measurement when training on previously recorded data from other subjects. 300 trials were selected from each subject based on the normalized MEP amplitudes, and nested CV was applied such that the data from one subject at a time was held apart for the accuracy testing. With the remaining data, a 7-fold CV process was run, such that the data from one subject at a time was kept apart for cross-validation, and the data from the rest were used in training the classifier. After selecting the hyperparameters providing the best average accuracy in this inner CV loop, the outer loop iteration was completed by evaluating the accuracy for the test subject data. Having completed the outer loop iterations, the final classification accuracy was obtained as the average test accuracy over all subjects.

2.4. Predicting cortical excitability state using prior information

It has been previously reported that the phase of the μ-rhythm as extracted by the Hjorth-C3 spatial filter is predictive of the MEP amplitude (Schaworonkow et al., 2019; Zrenner et al., 2018). Thus, for comparison, we also set up a classification pipeline using this prior knowledge. The weights of the spatial filter $\mathbf{w}_{\text{s,C3}}$ were 1 for electrode C3, and $-1/4$ for electrodes FC1, FC5, CP1, and CP5 (otherwise 0). We used this spatial filter together with an individualized temporal filter for decoding the excitability state as explained below. This decoding approach is termed here as the *fixed-spatial-filter decoding*.

In case of missing channels, they were interpolated using the intermediate step of source-localization by applying minimum-norm estimation to the remaining channels, after which the source estimates were projected to the missing channels via the forward model as suggested in (Burnes et al., 1998). We used singular-value truncation at the 35 highest values as the regularization method (Lawson and Hanson, 1995).

The power spectrum of the signal retrieved from spatial filtering was estimated using the 300 pre-stimulus EEG epochs of length 1 s. We used Matlab's spectrogram function with a Hamming window, the window length being the epoch width, and the frequencies of {7 Hz, 7.5 Hz,…, 13 Hz}. Then, the power spectrum was estimated as the average of the squared absolute values of the spectrograms over the epochs. After applying log10-conversion to the power spectrum values, pink noise was eliminated by linearly detrending the power spectrum. Robust detrending, described in detail in (de Cheveigne and Arzounian 2018), was used such that the power spectrum samples were iteratively excluded from the fitting process if their detrended values exceeded 1 x the standard deviation of the detrended spectrum.

The frequency which had the maximal power within the spectrum was identified. Two temporal filters, $\mathbf{w}_{\text{t, cos}}$ and $\mathbf{w}_{\text{t, sin}}$, were set as Morlet wavelets that included four cycles of sinusoidal signal at the selected peak frequency and had the phase difference of 90° After applying the spatial and temporal filters to each of the EEG epochs, we normalized the two obtained features by setting their norms to unit size according to the principles of circular-to-linear regression (Zoefel et al., 2019; Kempter et al., 2012), yielding the two-element feature vectors

$$z = \frac{\left[\mathbf{w}_{\text{s,C3}}{}^{\text{T}}\mathbf{X}\mathbf{w}_{\text{t, cos}},\quad \mathbf{w}_{\mathbf{w}_{\text{s,C3}}}{}^{\text{T}}\mathbf{X}\mathbf{w}_{\text{t, sin}}\right]^{\text{T}}}{\left\|\left[\mathbf{w}_{\mathbf{w}_{\text{s,C3}}}{}^{\text{T}}\mathbf{X}\mathbf{w}_{\text{t, cos}},\quad \mathbf{w}_{\mathbf{w}_{\text{s,C3}}}{}^{\text{T}}\mathbf{X}\mathbf{w}_{\text{t, sin}}\right]\right\|}. \tag{13}$$

These vectors were given as an input to the logistic regression model Eq. (7). Matlab's 'lassoglm' function to train the model with the same

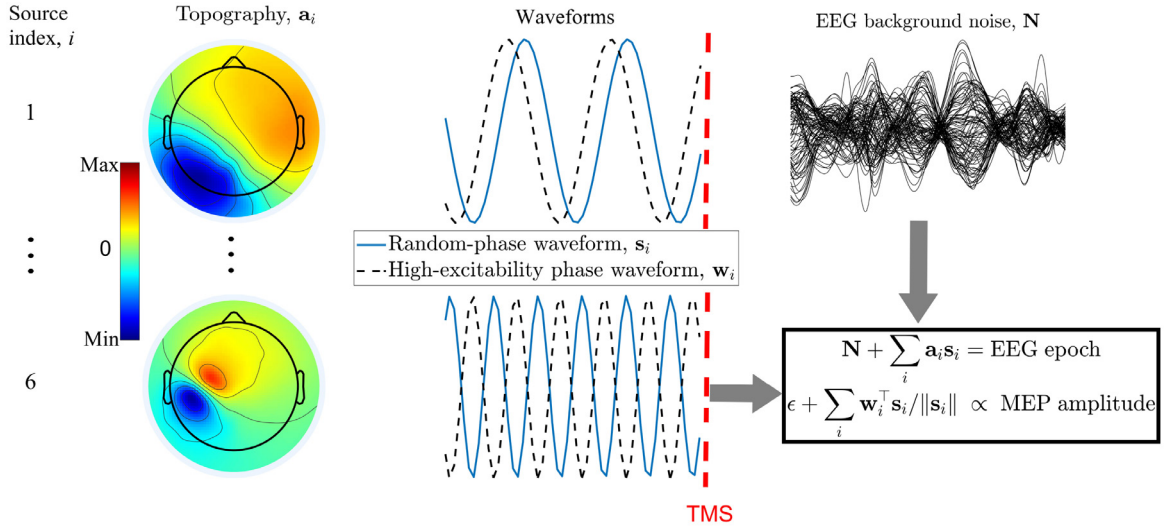

Fig. 2. Simulating data. The simulated data consist of the summation of simulated oscillatory dipolar activity and ongoing noise derived from resting-state EEG measurement. The topographies as well as the frequencies for each source were fixed over the simulated trials, but the phase and amplitude of the waveforms were randomized. Solid line is the simulated waveform **s** and dashed line is the waveform with optimal phase w. MEP amplitude in each trial resulted from the summation of the normalized inner products between the optimum-phase waveforms and the randomly generated waveforms and a noise term ϵ.

parameters as in the individualized decoding. The model was validated against the same set of 50 data epochs as the individualized model and in a 5-fold CV process. The average accuracy over validation set was used to measure the model performance. In practice, the 'lassoglm' function was used with options: 'distribution': 'binomial', 'alpha': 1, 'CV': 5, 'lambda': logspace(−3,−1,15), 'standardize': 1.

2.5. Data simulation and testing the decoding with simulations

To test how well our machine learning approach is able to decode excitability states based on phase-pattern information of EEG data with partly incorrect labeling based on MEPs, we simulated resting-state EEG epochs and respective MEP amplitudes. The data were generated with various levels of added EEG and MEP noise to mimic uncertainty in the data.

The simulation set-up is illustrated in Fig. 2. For each simulation, six cortical dipoles were randomly generated and their topographies $\mathbf{a}_i$ were computed according to a three-layer concentric spherical-head model (see Appendix A.3 for the details of the model). The excitability-encoding source waveforms $\mathbf{s}_i$ were following sinusoidal curves as a function of time. The frequencies of the six waveforms were equidistant in the frequency interval of [7 Hz, 30 Hz]. 200–400 trials with 301 samples and the sampling frequency of 1 kHz were generated. Across trials, the topography and the frequency of each source was fixed, but the phases and amplitudes were randomized. The amplitudes and the phases were distributed evenly in the intervals of [10 nAm, 200 nAm] and $[0,\ 2\pi)$, respectively.

These clean simulated data $\mathbf{X}_{\text{sim}}$ were superimposed with measured data epochs of one subject, which mimicked the background noise $\boldsymbol{N}$ as shown in Fig. 2. To control the SNR, we multiplied the noise with factor k to satisfy

$$\text{SNR} = \sqrt{\frac{\text{tr}\left(\boldsymbol{X}_{\text{sim}}\boldsymbol{X}_{\text{sim}}{}^{\text{T}}\right)}{k^2\ \text{tr}\left(\boldsymbol{N}\boldsymbol{N}^{\text{T}}\right)}} \times 100\%, \tag{14}$$

i.e., SNR = 100% if the powers of the simulated phase-pattern data and the noise were equal.

The MEP amplitudes were simulated such that large MEPs arose when the trial-wise phase of the simulated source was close to optimal, and small MEPs resulted at the opposite phase. Each source had an individual optimal phase, which was randomized in the interval of $[0,\ 2\pi)$ and kept fixed over the simulation trials. In each trial i, the noiseless MEP amplitude deviation from the trendline was generated as

$$\Delta\text{MEP}^{(i)} \propto \sum_n \frac{\boldsymbol{\omega}_n{}^{\text{T}}\boldsymbol{s}_n^{(i)}}{\left\|\boldsymbol{s}_n^{(i)}\right\|} + \epsilon^{(i)}, \tag{15}$$

where $\boldsymbol{\omega}_n$ is the waveform of the optimally phased waveform of source n, $\boldsymbol{s}_n^{(i)}$ is the waveform of source n in trial i, and $\epsilon^{(i)}$ is Gaussian noise with zero mean and the variance of σ^2.

We were specifically interested in whether the MEP amplitude is larger or smaller than the trendline, reflecting high- or low-excitability state, respectively, as explained in Section 2.4. We interpreted ΔMEPs as the detrended MEP amplitudes in Fig. 1B. Thus, a label was set 1 when the respective ΔMEP was positive and 0 otherwise. We kept track of both the *true labels* and *empirical labels* which were set based on noiseless and noisy ΔMEPs, respectively. Increasing the variance of the MEP noise, $\epsilon^{(i)}$, makes the empirical labeling more inaccurate.

The accuracies with respect to the estimated labels were evaluated using the validation set of 25 largest and 25 smallest noisy ΔMEPs, while the rest of the data were used for training the individualized decoding model. For simplicity, without testing, the regularization constant for CSP was chosen to be 10^{-2} and the number of principal components, N in Eqs. (9) and (11), was set to 4. SNR (Eq. (14) was systematically varied over the values of {10%, 30%, 50%, 75%, 100%}, and the standard deviation σ of $\epsilon^{(i)}$ (Eq. (15)) was set to {0, 0.5, 0.75, 1, 1.25, 1.5} × standard deviation of noiseless ΔMEPs. With each parameter combination, the simulation and decoding were run 20 times, each time with randomly generated source locations and optimal phases.

2.5.1. Decoding with perfect prior knowledge

Assuming the ideal case where the source locations, frequencies, and optimal phases for brain state decoding would be known, we built an optimal classifier for the simulated data. We tested how accurate results such a prior knowledge –guided classifier could provide at best. First, beamforming was used to estimate a spatial filter:

$$\mathbf{w}_{\text{s},i} = \mathbf{C}^{-1}\mathbf{a}_i\left(\mathbf{a}_i{}^{\text{T}}\mathbf{C}^{-1}\mathbf{a}_i\right)^{-1}, \tag{16}$$

where $\mathbf{a}_i$ is the topography of source i, and $\mathbf{C}$ is the sample covariance matrix. Here, we deliberately used the exact knowledge of the topography although it is one of the so-called 'inverse crimes'.

After spatial filtering by $\mathbf{s}_i{}^{\text{T}} = \mathbf{w}_{\text{s},i}{}^{\text{T}}\mathbf{X}$, for temporal filtering, we used two sinusoidal waveforms $\boldsymbol{w}_{\text{sin},i}$ and $\boldsymbol{w}_{\text{cos},i}$ with a 90° phase difference

and the frequency exactly at the encoding frequency of each source i. The spatially and temporally filtered EEG epochs were then normalized to unit-norm for each source separately to yield the optimal feature vector for source i as

$$\boldsymbol{z}_i^{\mathrm{OPT}} = \begin{bmatrix} \boldsymbol{s}_i{}^{\mathrm{T}}\boldsymbol{w}_{\mathrm{sin},i} \\ \boldsymbol{s}_i{}^{\mathrm{T}}\boldsymbol{w}_{\mathrm{cos},i} \end{bmatrix} \Big/ \left\| \begin{bmatrix} \boldsymbol{s}_i{}^{\mathrm{T}}\boldsymbol{w}_{\mathrm{sin},i} \\ \boldsymbol{s}_i{}^{\mathrm{T}}\boldsymbol{w}_{\mathrm{cos},i} \end{bmatrix} \right\| \tag{17}$$

according to circular-to-linear modeling (see (Kempter et al., 2012; Zoefel et al., 2019)). Finally, to get the optimal brain-state index, we matched the normalized features with the known optimal phases by calculating the inner product between the filtered EEG signal and the optimally phased sinusoids by

$$I_{\mathrm{BS}}^{\mathrm{OPT}} = \sum_i \left(\boldsymbol{z}_i^{\mathrm{OPT}}\right)^{\mathrm{T}} \begin{bmatrix} \sin\alpha_i \\ \cos\alpha_i \end{bmatrix}, \tag{18}$$

where α_i is the optimal phase of the ith source. If the inner product score exceeded 0, the trial was assigned to class 1 (high-excitability state), and to 0 (low excitability) otherwise.

2.6. Statistical thresholding and statistical comparisons

By classifying randomly several times, the distribution of the random classification results can be simulated. To mimic the classification of a single subject test set, we randomly classified 50 trials into high- and low-excitability states, repeating the process 10^6 times. We determined an accuracy threshold to describe how high accuracy would unlikely be obtained in a randomized setting. We set the threshold at the significance level of 0.05, which means that 95% of the random classification accuracies fell below the threshold accuracy.

We also tested the overall classification results using both the individualized and the fixed-spatial-filter decoding against random classification. To this end, we randomly classified 50 trials into high- and low-excitability states eight times (the number of subjects) and repeated the process 10^6 times. The p-value was computed as the number of times when the average accuracy over the eight subjects exceeded the obtained overall accuracy divided by 10^6.

In addition, using a resampling strategy, we made a statistical comparison of the data-driven decoding against the fixed-spatial-filter decoding. In this comparison, we randomly shuffled between the accuracies of the two decoding methods for each subject. This process was repeated 10^6 times, and each time, the average change of the accuracy of the data-driven decoding relative to that of the fixed-spatial-filter decoding was computed. We then checked how many times the absolute value of the difference resulting from the shuffling exceeded the actual difference and divided this number by the total number of randomizations to get the p-value. The significance level of 0.05 was used in the statistical tests.

2.7. Physiological interpretation of spatio-temporal filters

2.7.1. Converting the spatio-temporal filters into EEG patterns

It is worth noting that, optimally, the spatio-temporal filters do not simply mimic the pattern of EEG that would predict high-excitability state. The filters should rather cancel out the uninteresting parts of the signals (Blankertz et al., 2007). The estimated filter weights can be converted into the EEG epoch that would most likely predict high-excitability state, given the measured data set. We used the relation given in (Haufe et al., 2014) to convert the matrix filter into the respective EEG epoch representing the highest excitability.

In order to convert an estimated filter into the corresponding pattern, both the spatio-temporal filter $\mathbf{W}$ and the EEG epochs $\mathbf{X}^{(i)}$ $(i = 1, \ldots, R)$ were converted into vector form simply by concatenating the matrix columns vertically into $(M \cdot T \times 1)$ vectors by $\mathbf{x}_{\mathrm{ST}}^{(i)} = \mathrm{vec}(\boldsymbol{X}^{(i)})$, and $\mathbf{w}_{\mathrm{ST}} = \mathrm{vec}(\mathbf{W})$. The corresponding vector-form spatiotemporal pattern, $\mathbf{v}$, was retrieved by

$$\mathbf{v} = \mathrm{cov}\left(\mathbf{x}_{\mathrm{ST}}^{(i)}\right) \mathbf{w}_{\mathrm{ST}} \left(\mathbf{w}_{\mathrm{ST}}{}^{\mathrm{T}}\, \mathrm{cov}\left(\mathbf{x}_{\mathrm{ST}}^{(i)}\right) \mathbf{w}_{\mathrm{ST}}\right), \tag{19}$$

where $\mathrm{cov}(\cdot)$ denotes computing the sample covariance matrix over samples indexed by i. For interpretations and visualization, we converted the vector-form pattern back into an $M \times \mathrm{T}$ EEG pattern as $\mathbf{V} = [\mathbf{v}_{1\,:\,M},\ \mathbf{v}_{M+1\,:\,2M}, \cdots,\ \mathbf{v}_{(T-1)M\,:\,T\cdot M}]$. The spatio-temporal pattern obtained by Eq. (19) would be most predictive of the high-excitability brain state according to the estimated model and for the individual subject. According to the model (Eq. (6)), if the polarity (sign) of this pattern is reversed, the prediction of the brain state would also be reversed.

2.7.2. Source localization of the spatio-temporal patterns

The high-excitability EEG patterns can be used to estimate where the state-encoding source activity arises on the cortex. For each cortical location i, the spatial filter $\mathbf{w}_i$ was obtained by the beamforming formulation of Eq. (16) using each of the lead-field matrix columns as topographies. We used the beamforming ZER localizer (Moiseev et al., 2011) to estimate the power $\mu_{i,t}$ of the source amplitudes at each location i and time index t as

$$\mu_{i,t} = \frac{\mathbf{w}_i{}^{\mathrm{T}}\mathbf{V}_{:,t}\mathbf{V}_{:,t}{}^{\mathrm{T}}\mathbf{w}_i}{\mathbf{w}_i{}^{\mathrm{T}}\mathrm{cov}(\mathbf{N})^{\mathrm{T}}\,\mathbf{w}_i}, \tag{20}$$

where $\mathbf{N}$ is the noise data, and $\mathbf{V}_{:,t}$ is the tth column of the spatiotemporal pattern matrix, *i.e.*, the topography of the EEG pattern at time index t. As noise data, we used all the EEG data over the 300 epochs. The scanning function value is contrasting the power of the interesting signal in cortical location i against the power (mean squared amplitude) of the overall ongoing resting-state activity in the same location.

Beamforming scanning was performed for the high-excitability pattern of each subject separately using their individual lead-field matrices. To define the excitability encoding sources in each individual subject, we identified the sources whose localizer values exceeded a threshold. The threshold was set three times the mean of all the scanning function values over time and source locations.

In order to find common state-encoding sources and time points across subjects, the instantaneous relevancy of the given source at the given time was computed as the proportion of subjects for whom the respective source amplitude exceeded the individual threshold:

$$\mathrm{Relevancy}_{i,t} = \left| \left\{ s \in \mathrm{subjects} \,\middle|\, \mu_{i,t}^{(s)} \right\rangle \mathrm{threshold}^{(s)} \right\} \Big/ |\{ s \in \mathrm{subjects} \}|, \tag{21}$$

where s denotes the subject index, and $|\cdot|$ denotes the number of elements in the set. Furthermore, the overall relevancy at a given source location was computed as the proportion of subjects for whom the amplitude in the source location exceeded the threshold at any time instant, *i.e.*, omitting index t in Eq. (21).

2.7.3. Fluctuating brain states in the resting-state EEG

We may use an estimated spatio-temporal filter to examine how the brain state changes as a function of time by applying the filter to resting-state EEG as depicted in Fig. 1C: If running the filter in a sliding window over the EEG recording moving forward one sample at a time, the resulting score reflects changes in the cortical excitability state. From the time-course of the score, we can evaluate its frequency content.

Using the sliding window scheme, we computed the score signal over all the 300 EEG trials. We then used the 'spectrogram' function of Matlab to compute the single-trial spectrograms using the Hamming window of length 250 and zero overlap. The power spectrum was estimated by averaging the squared absolute values of all the spectrograms. We applied log10 conversion of the power spectra and detrended them by eliminating the linear trend (pink noise) using robust detrending. In the robust detrending, the line was fitted to the spectrum in an iterative fashion always discarding from the fit the samples whose detrended values exceeded their standard deviation. See (de Cheveigne and Arzounian 2018) for details of the robust detrending process.

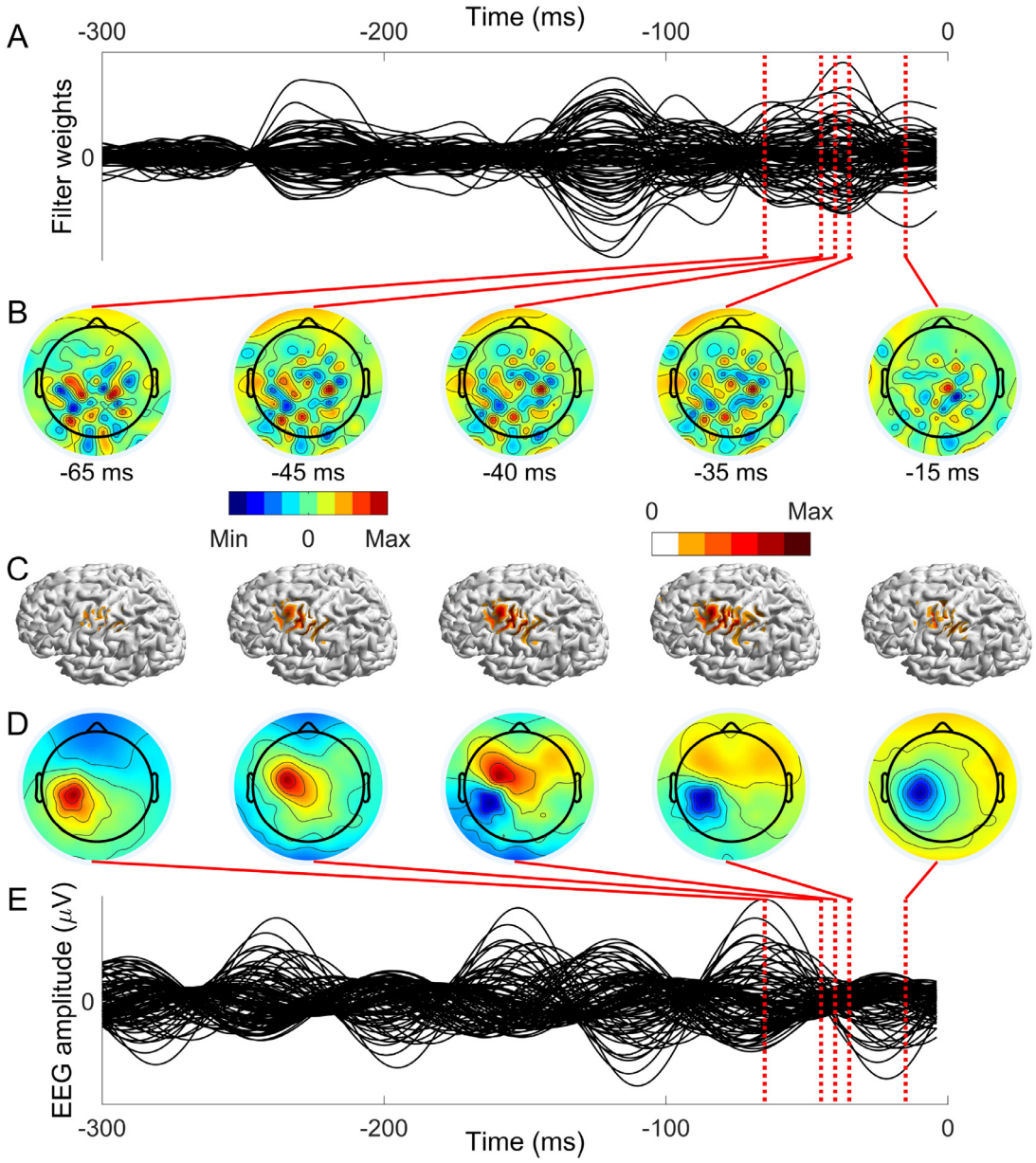

Fig. 3. Final optimized spatio-temporal filter and the respective EEG pattern from one example subject. A. Estimated spatio-temporal filter as a butterfly plot, *i.e.*, each line represents the time-evolving weights for the respective channel. Time is given relative to the TMS pulse at 0 ms. B. The spatial filters at five selected time instant (red dashed vertical lines in A) are shown as topographies. C. The source estimates for the topographies in D in the respective order. The estimates reflect power of the source and are thus always positive. D. The topographies which the spatial filters in B are optimally detecting in the respective order. E. The spatio-temporal pattern (EEG epoch) to which the spatio-temporal filter (in A) is most sensitive to is shown as a butterfly plot. B-D. Color scales are individually set to span the range of values in each subfigure separately.

3. Results

3.1. Results from measured EEG–TMS–MEP data

3.1.1. Example case of estimated spatio-temporal filters, patterns, and source estimates

For each subject, the spatio-temporal filter was individually estimated using the optimized hyperparameters by the CV scheme as explained in Section 2.4. The computational time to run the 5-fold CV process was less than 30 second on a current standard PC hardware. Running the nested CV process for accuracy estimation took about 24 min. An example of the estimated spatio-temporal filter is shown in Fig. 3A. Five time instants from the ST filter near to the onset of the TMS pulse were picked for inspection (Fig. 3B) in this example case. The distribution of the weights over the channels at these time instants can be interpreted as spatial filters. By visual inspection, the topographies of these spatial filters are difficult to interpret.

We used the conversion by Eq. (19) from the filter into the spatio-temporal pattern representing the optimal high-excitability state. The obtained spatio-temporal pattern shows a periodic signal and is clearly different from the filter as seen by the butterfly plot (Fig. 3E). Furthermore, the topographies extracted from the spatio-temporal pattern at the same time instants as the spatial filters show greatest values over the sensorimotor area of the stimulated left hemisphere, where positive and negative voltages are alternating and moving gradually over time (Fig. 3D). The source estimates for each topography are shown in Fig. 3C, and they show a repeated activation pattern moving in a posterior-anterior direction around the central sulcus of the left hemi-

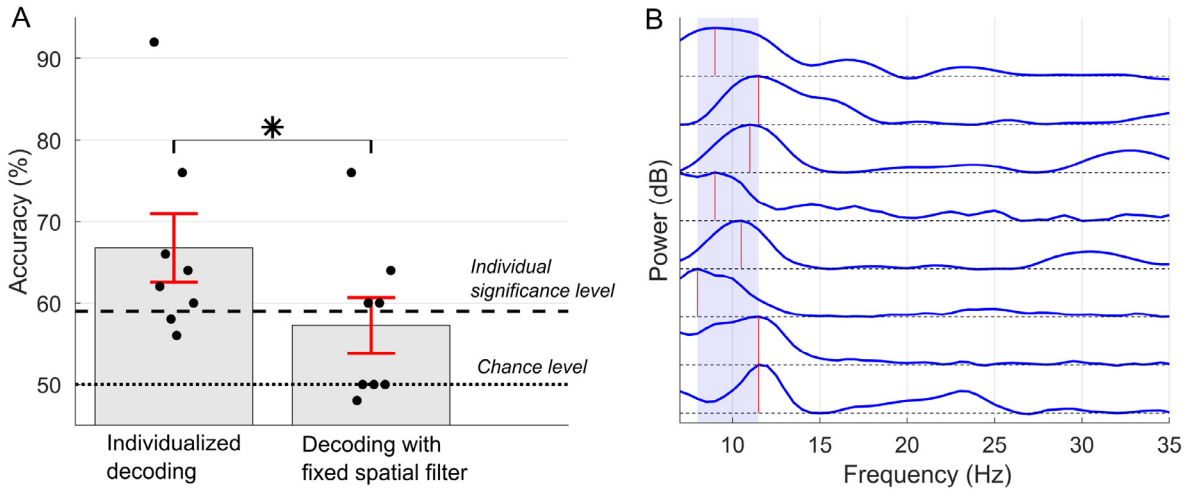

Fig. 4. Decoding results from the measured EEG–TMS–EMG data. A. Accuracies of the two classification methods. The bars indicate the mean accuracies, while the red error boundaries show the standard errors. Individual accuracies are represented by the black dots. The mean accuracy significantly increases when using individualized decoding as compared to fixed-spatial-filter decoding. The individual significance level is the accuracy which is exceeded with a probability of less than 0.05 by a random classifier. B. Power spectra (normalized to the maximum of 1 dB for each subject) of the brain state fluctuations for each subject are shown in blue. The red vertical lines indicate the positions of the maxima. Light blue shadowing shows the frequency range across subjects where the maxima lie.

sphere. Similar illustrations of the estimated filters and patterns for the rest of the subjects can be found in the Supplementary Material. For brevity, in the following, we use the term EEG pattern when referring to the spatio-temporal patterns representing the optimal high-excitability state.

We used the conversion by Eq. (19) from the filter into the spatio-temporal pattern representing the optimal high-excitability state. The obtained spatio-temporal pattern shows a periodic signal and is clearly different from the filter as seen by the butterfly plot (Fig. 3E). Furthermore, the topographies extracted from the spatio-temporal pattern at the same time instants as the spatial filters show greatest values over the sensorimotor area of the stimulated left hemisphere, where positive and negative voltages are alternating and moving gradually over time (Fig. 3D). The source estimates for each topography are shown in Fig. 3C, and they show a repeated activation pattern moving in a posterior-anterior direction around the central sulcus of the left hemisphere. For brevity, in the following, we use the term EEG pattern when referring to the spatio-temporal patterns representing the optimal high-excitability state.

3.1.2. Prediction accuracies

Using the nested CV protocol, the classification accuracy of the individualized decoding was estimated for each subject. In addition, the classification was computed for the fixed-spatial-filter decoding using Hjorth-C3 as the spatial filter (see Section 2.5). The classification accuracies of these filters for the test data are depicted in Fig. 4A. Individualized decoding is superior to the classification by the fixed-spatial-filter decoding. On average, for the individual decoding, the accuracy is 67%, with the minimum of 56% and maximum of 92%. The fixed-spatial-filter classification gives accuracies in the interval of [48%, 76%], with the mean 57%.

Statistical comparison of the average accuracies to the random classification yielded p-values of $<10^{-6}$ and 0.0005 for the individual and fixed-spatial-filter classifications, respectively. Thus, at the group level, both classifiers gave significantly more accurate results compared to random classification. Comparing between the averaged accuracies from the individualized and fixed-spatial-filter decoding retrieves a p-value of 0.012, indicating a significant improvement in excitability estimation when using the individualized model. The random permutations gave a threshold of 59% for the individual accuracy at the significance level of 0.05. Six out of eight subjects had an accuracy over this 59% threshold when using individualized decoding, while this was the case for only four out of eight subjects for the fixed-spatial-filter decoding.

3.1.3. Neurophysiological interpretations

We computed the power spectrum for the excitability state fluctuations for each subject. The peak frequency of the power spectra has the mean of 10.25 Hz and the range of [8.0 Hz, 11.5 Hz] across subjects. The power spectra are shown in Fig. 4B. By visual inspection, three subjects also had additional smaller local maxima in their spectra at higher frequencies, but these did not show consistent patterns across subjects.

The source localization was individually performed with each subject's EEG pattern using beamforming as the source reconstruction method as explained in Section 2.8. Relevancy (*i.e.,* the proportion of subjects where a source is participating in encoding the excitability state) over all time instants on the cortex is depicted in Fig. 5A. The highest relevancy values were found around the sensorimotor region of the stimulated left hemisphere, the highest ones reaching 0.8. These sources are encoding excitability in most subjects. At the individual level, as shown by all relevancy values exceeding 0, relevant cortical regions spanned also over areas on both ipsi- and contralateral parieto-occipital and also temporal regions with respect to the stimulation target. Notably, contralateral pre- and postcentral gyri were not relevant in encoding the excitability state in any of the subjects.

When computed at each time instant separately, the relevancy was 0.5 at highest. These time-domain relevancies are depicted in Fig. 5B for the subset of 72 sources which had higher than 0.5 overall relevancy. Most sources have low relevancy values in time domain (Fig. 5B–C), which indicates that the temporal activity leading to the high-excitability state is highly heterogeneous over subjects. The highest relevancy values as well as the largest amount of relevant sources tend to appear less than 100 ms before the TMS onset (Fig. 5B–C).

The distribution of hyperparameter values, given by cross-validation, to obtain the patterns were as follows: The median number of features (N in Eqs. (9) and (11)) was 4, with the 1st and 3rd interquartiles of [2, 8], and the minimum and maximum of 2, and 20, respectively. The median CSP regularization coefficient γ was 10–2, with the 1st and 3rd interquartiles of [10–2, 10–1], and the minimum and maximum of 10–4, and 1, respectively.

Due to the long measurements and to the stimulation frequency of around 0.5 Hz, gradual changes in MEP sizes could arise, *e.g.,* due to plastic effects, which were removed by detrending of the MEPs in our

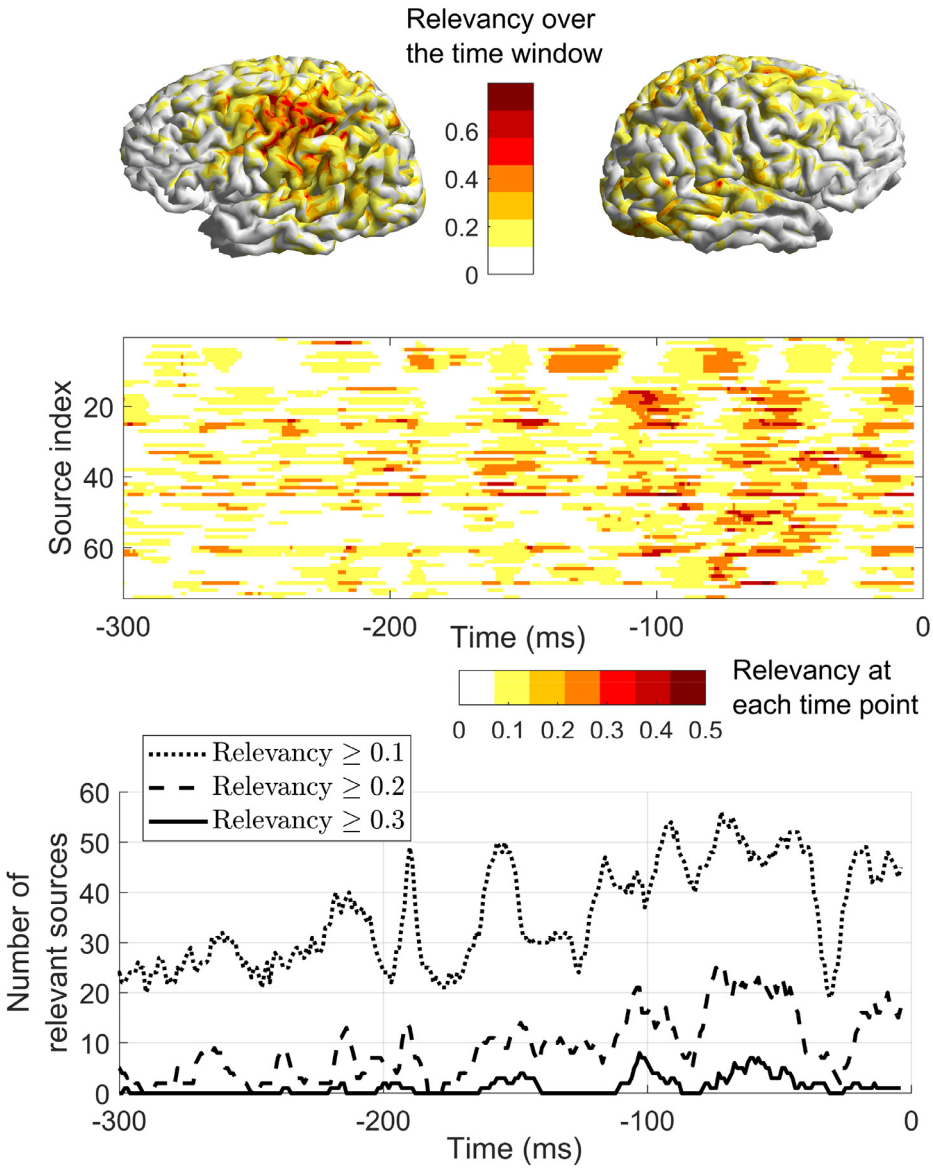

Fig. 5. A. Relevancy (of a source) describes the proportion of the subjects which have high-amplitude activity encoding the high-excitability state in each source location. Here, relevancy was estimated across the eight subjects and over all time points. B. Relevancies as a function of time with respect to TMS. The subset of 72 sources with the overall relevancy (in panel A) of over 0.5 are depicted. Note that the color scale is different from A. for visibility. C. The number of active sources, having higher-than-threshold relevancy in B. as a function of time.

analysis. We tested whether there was a systematic change in the MEP amplitudes across the trials, in case detrending is not applied, by aggregating all MEP amplitudes (two channels from all subjects) across the measurements. The results indicate that there was a rising trend: The mean average difference of the MEP amplitudes between the second and first halves of each session was 71 μV, with the p-value of 0.0017 given by permutation testing. After detrending, no significant change appears. Further EEG estimator stability tests were also performed, as explained in the following section.

3.1.4. Generalizability of the classification

Stability tests were performed both over experiments and within each subject over the course of individual experiments to see how generalizable the estimated filters were across individuals and across the measurement time. For cross-subject stability, leave-one-subject-out nested CV gave an average accuracy of 54% with the standard error of 1.6%, and range [47%, 60%]. Testing the stability over the individual measurements, where the latter half of data were decoded after training the classifier with the first half, gave an average accuracy of 64% with the standard error of 3.2%, and range [46%, 78%]; see Fig. 6.A with 'ALL' electrodes for the accuracy distribution. This average accuracy is close to the benchmarking decoding accuracy in Fig. 4.A, and significantly higher than random classification accuracy (p-value $< 10^{-6}$).

We further continued with stability testing, to see how the accuracy changes when reducing the number of electrodes. Electrodes were eliminated in each subject by Eq. (12), yielding the results in Fig. 6.A where it shows that the decoding accuracy stays rather intact with all tested numbers of electrodes, down to as few as 20 electrodes. In Fig 6.B, the distribution of 30 chosen electrodes is depicted over the subjects. There are no evident most important confined regions of electrode locations; the chosen electrodes are scattered sparsely around the EEG cap, but electrodes ipsilateral to the TMS target are more often chosen to be used in the classification.

3.2. Simulation results

Artificial data were generated with the simulation protocol explained in Section 2.6. The instantaneous phases of the underlying sources were predictive of the MEP size (excitability) (see Fig. 2 for the simulation setup). Using the perfect knowledge of the topographies, the frequencies, and the high-excitability phases, we built the optimal spatial and temporal filters to predict the excitability state according to Eq. (18). At the SNRs of 100% and 10%, the prediction accuracies were 98% and 96%, respectively, see Fig. 7B.

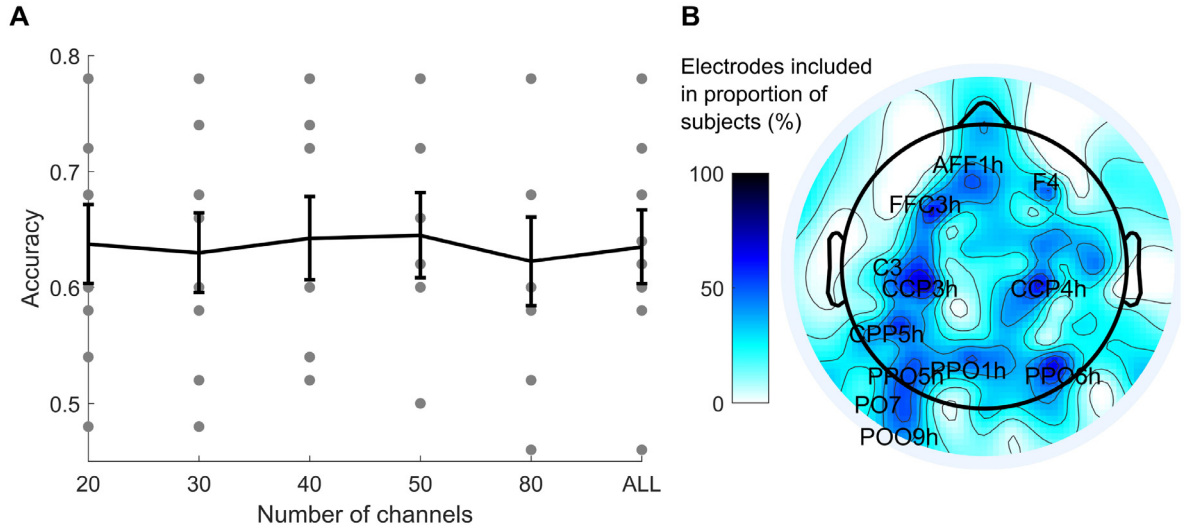

Fig. 6. Effect of the number of EEG channels used and distribution of the chosen electrodes. A. Accuracy as a function of number channels. ALL stands for individual maximum number of channels. gray dots are individual accuracies. The black line shows the mean accuracy, and the error bars indicate the standard error of accuracies for each number of channels separately. B. For each electrode, the%-value indicates the proportion of subjects where the electrode was included among the 30 most important ones. The labels are displayed for electrodes included within at least 50% of the subjects.

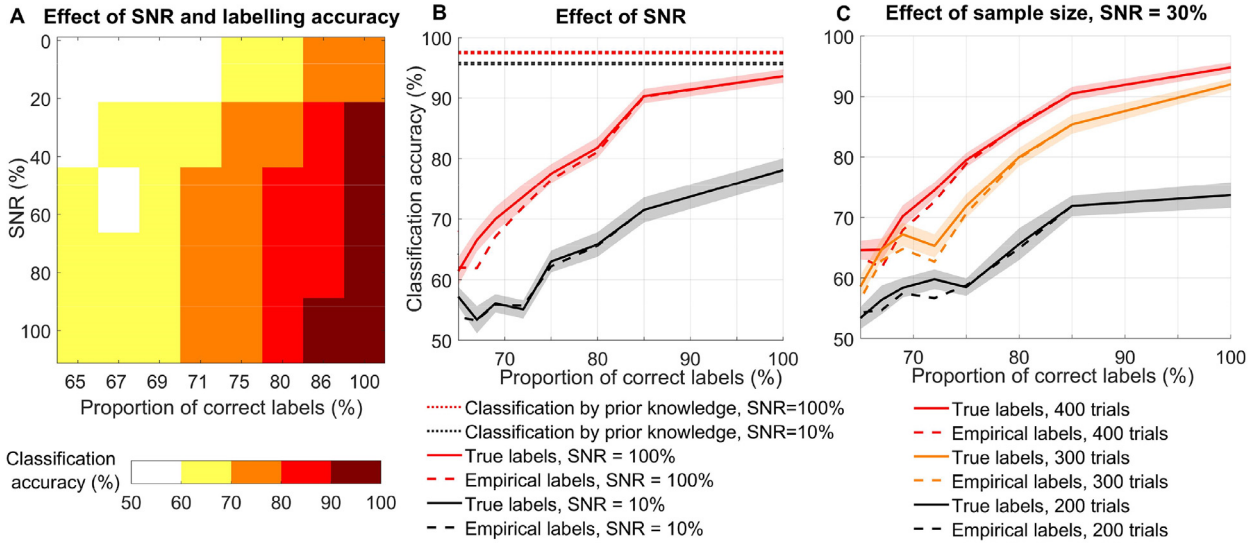

Fig. 7. Results from simulated data. A. Accuracy of the trained spatio-temporal decoding as a function of SNR of the EEG data and proportion of correctly labeled trials based on MEP amplitudes. B. Classification accuracies measured with true and empirical labels. SNRs of 10% and 100% are shown as a function of correctly labeled trials. If the source properties encoding the excitability are known exactly, the accuracies exceed 95% ('Classification by prior knowledge'), and there is no dependency on labeling because training is not required. There is a minor difference when computing the accuracies with respect to true labels *vs*. empirical labels (based on noisy MEPs). C. Classification accuracies with fixed SNR = 30% and different numbers of samples. Both empirical and true accuracies are shown here. In B-C, the shadowed areas span the standard error of the accuracy around the 'true labels' curves.

In reality, perfect prior knowledge is not available, and we rely on the training data to build the predictive model. We tested the proposed individualized spatio-temporal decoding approach using the simulated data. The results are illustrated in Fig. 7. When the SNR was high (*i.e.*, 100% with equal clean signal and noise powers) and the labeling fully correct, we could reach a high prediction accuracy of 94% (Fig. 7*A*–*B*). Training was clearly sensitive to incorrect labels due to MEP noise; the accuracy decreased rapidly when the proportion of correct labels fell below 85%. On the contrary, the approach could better tolerate decreasing SNR in EEG; the accuracy stayed above 90% when SNR was above 30% and the labels were correct (Fig. 7A).

The above-described accuracies were given with respect to the true labels obtained from noiseless MEPs. However, with measured data, we can only check empirical accuracies with respect to the labels estimated from noisy MEPs. Thus, we also checked how large a difference there is between these two accuracy measures. According to the results in Fig. 7B, the true and the empirical accuracies were very similar on average. There was a tendency for the true accuracy of being slightly above to the empirical accuracy when the labeling got more incorrect.

Finally, we also tested the classification accuracy against increasing or decreasing the amount of training data trials by 100 when compared to 300 trials at SNR = 30% (Fig. 7C). The decrease in the number of trials clearly worsened the results: the maximal accuracy reached 74% with 200 trials. In addition, to exceed the 59% accuracy (the individual significance level), at least 80% of the labels needed to be correct. Increasing the number of trials from 300 to 400 had a positive effect on

the accuracy, but the improvement was not as large as the negative effect when decreasing the number of trials. Furthermore, at low labeling accuracies, the accuracy was almost the same irrespective of the number of trials.

4. Discussion

4.1. Implications

The results indicate that prediction by the individualized classification is more accurate than the Hjorth-C3 fixed-spatial-filter decoding (Fig. 4A). The odds that individually predicted EEG excitability state matches *vs*. mismatches with that given by the MEP amplitude are 2:1 (67% accuracy). This can be regarded reasonably high given the high uncertainties in the EEG and EMG data discussed below. It is noteworthy that the compared spatial and temporal filters, C3-Hjorth and wavelet within μ-rhythm frequency, are actually targeting the cortical areas and frequencies which were shown to be relevant in excitability decoding, but these filters still yield lower accuracies in the classification. This highlights the need for individualized filters to minimize the leakage signals from irrelevant cortical activity to the classification features.

Excitability-state-dependent personalized repetitive TMS (rTMS) has been applied synchronized to the phase of sensorimotor μ-rhythm (Zrenner et al., 2018; Baur et al., 2020) oscillations in the dorsolateral prefrontal cortex (Zrenner et al., 2020), and theta oscillations in the dorsomedial prefrontal cortex, with the goal of reducing the substantial interindividual variability of neuromodulatory effects of rTMS (Lopez-Alonso et al., 2014; Kar 2019; Hordacre et al., 2017; Hamada et al., 2013). However, the efficacy of brain-state-dependent TMS is limited by the accuracy by which brain-states are estimated from EEG, which is limited even in the case of extracting phase from a single oscillation (Zrenner et al., 2020). The present approach goes beyond the phase of a single oscillation and seeks to classify the relevant excitatory state from a dynamic spatio-temporal EEG pattern. Individually optimized brain-state extraction achieved by supervised learning could improve the therapeutic potential of EEG-triggered brain stimulation in the future.

Decoding the measured EEG−TMS data with the proposed algorithm results in neurophysiologically plausible patterns, without *a priori* hypotheses regarding the locations or frequencies of the relevant source activity: Left primary motor cortex (the target location of TMS) and adjacent parts of the sensorimotor network emerge as relevant cortical regions whose activity encodes the high-excitability state (Fig. 5A), and this state fluctuates most prominently at a frequency within the μ-rhythm range (8 – 11.5 Hz) (Fig. 4B). This result, based on the group average, is consistent with previous findings (Haegens et al., 2011; Zrenner et al., 2018; Schaworonkow et al., 2019), substantiating the applicability of the decoding method.

At the individual level, informative sources are found across large areas of cortex (Fig. 5A), indicating that distributed networks take part in encoding the excitability state. However, while there is a clear overlap across participants of individually relevant cortical activity spatially (when aggregating across the whole time window, Fig. 5A), the temporal waveforms that reflect the optimal high-excitability state show a high variability across subjects (there is little overlap in relevant sources at any given time point in Fig. 5B).

The presented methodology can be applied to study excitability encoding patterns under various conditions and TMS paradigms without imposing strict prior hypotheses. For example, it could be useful to apply the same approach to patients after stroke, and to compare the resulting patterns that predict excitability in healthy persons *vs*. stroke patients. Changes in the state-encoding patterns within individual patients, as immediate response to therapy (*e.g.*, by TMS) or in a longer time span, can give insight into cortical changes reflecting individual recovery. The patterns or their changes could serve as biomarkers to track recovery or to tailor more efficient, individualized therapy protocols.

4.2. Accuracy of excitability state labeling

From the simulated data, we can draw the conclusion that individual data-driven modeling is able to detect brain states defined by the dynamic spatio-temporal pattern of the EEG sources. The accuracy of the model seems to be most sensitive to the accuracy of labeling the epochs into high- and low-excitability classes. SNR of the EEG data is less important. When the data are accurately labeled, and the SNR is at least 30%, the model performs at a very high accuracy of over 90%, which is close to the accuracy obtained by using the perfect knowledge of the source properties encoding the excitability (Fig. 7C). Increasing the number of trials improved the decoding accuracy, but this improvement was not remarkable, especially if the labels were noisy.

Given the importance of accurate labeling, we need to take a critical look at how the classes were assigned based on the EMG data. As we labeled the data epochs based on the MEP amplitudes, we made several heuristic decisions (see Methods for details). The labels could have turned out differently if choosing a different strategy. Moreover, the stimulation intensity may play an important role. Here, the TMS intensity of 110% of the RMT was used, for which reason MEP sizes tended to exceed 50 μV irrespective of the instantaneous brain state. More accurate labeling could result when using an intensity closer to the RMT, making it more straight forward to draw a line between 'response' and 'no response' trials (de Goede and van Putten 2019). In fact, it has been previously reported that low TMS intensities lead to improved predictive power of the MEP amplitude based on the μ-rhythm phase (Schaworonkow et al., 2019).

Despite the arbitrary choices related to MEP detrending and labeling, these steps are important as highlighted by the systematic rise of MEPs during the sessions across subjects, but also to compensate for individual non-systematic time-effects. If such slow changes of baseline MEP amplitude are not corrected, the MEP prediction, based on rapidly fluctuating EEG phase, would inevitably be hampered. Low-frequency rTMS protocols, such as the one used in this study, are typically expected to induce a long-term depression (LTD)-like effect. Previously, 0.5-Hz rTMS has been used to induce decrease of cortical excitability in epilepsy (Sun et al., 2012), whereas in healthy participants, also absence of significant changes has been reported (Schrader et al., 2016). Note that the LTD-like effect is quantified by test pulses before and after the rTMS protocol when rTMS is used as a neuromodulatory intervention. Our study, however, is designed to investigate the trial-by-trial MEP fluctuations during the stimulation, and therefore, ISI included a significant jitter. The fact that this protocol resembles a jittered 0.5-Hz rTMS protocol is a consequence of the need to maximize the number of trials and to minimize the slow fluctuations that occur during long or split sessions.

4.3. Limitations

The trial-to-trial variability of the MEP amplitudes within an experimental session is attributed to a multitude of extrinsic and intrinsic factors. Extrinsic factors include all changes affecting the electric field at the target cortical area; most evidently even small changes in the coil targeting can lead to a remarkable change in the MEP (Kallioniemi et al., 2015; Koponen et al., 2018; Miranda et al., 1997). Intrinsically, individual person's neurophysiological excitability state can fluctuate at multiple levels of the motor tracts from the cortical, to the spinal and peripheral circuits (Funase et al., 1999; Kiers et al., 1993; Magistris et al., 1999). Previously, the MEP amplitudes have also been modelled by the properties of pre-stimulus EMG (Keil et al., 2014; Mitchell et al., 2007). At the cortical level, there can be local variations, affecting the excitability of a small cortical region, or more global ones affecting larger areas, *e.g.*, the primary motor cortices bilaterally (Ellaway et al., 1998). In addition to phase, power of the underlying oscillations has been used in determining the cortical excitability state (Sauseng et al., 2009; Thies et al., 2018).

While excitability undoubtedly changes in a graded fashion, here we used binary modeling. The rationale of choosing such binary classification instead of regressing to the exact MEP amplitude is due to challenges of estimating an unknown mapping function among highly noisy samples. As discussed above, MEP amplitudes are affected by several factors, in addition to the cortical excitability state, which, from the perspective of our modeling framework, represent noise in the system. With the classification approach, we bypass the need of specifying a function, *e.g.*, linear regression, or scaling (linear or logarithmic) to be used in mapping the relationship between an EEG feature and the MEP amplitude.

When reducing the dimensionality of the data, with each subject, we used different regularization parameters and degrees of freedom. These parameters were decided using cross-validation, which was automatized but nevertheless time-consuming. It is commonly a trade-off between tuning the modeling parameters for getting an individually optimized model or pre-setting parameters to get a less accurate estimate faster. For the decoding presented in this work, more model parameters could have been added, such as the width of the decoding window.

Our matrix filter was established in time-domain since, based on preliminary experimenting, we found this way the most straightforward to use with least prior assumptions. If one is most interested in retrieving the phases (of the source waveforms or the channel signals) encoding the high excitability in the frequency domain, these would have to be further estimated from the spatio-temporal patterns, which would contain estimation error. Alternatively, one can create a predictive filter in the frequency domain: The model is then set up using chosen frequencies, and the resulting high-excitability pattern would show the optimal phase/amplitude combination for each of the frequencies.

The average obtained accuracy is rather low compared to those obtained from BCI applications (Blankertz et al., 2008). In BCI applications, however, the experiments are designed such that the SNR of the relevant EEG signal is rather high, and the trial labeling is fully accurate, which is not the case here, and we are rather dealing with probabilities than clear binary predictions. As outlined in the Introduction, highly diverse findings starting from a similar research question as here have been published. Thus, high accuracies of the MEP amplitude prediction were not expected, and indeed, a large variability remains in how effectively high-excitability states can be predicted for a given participant: The prediction accuracies were ranging from 56% to 92% (Fig. 4A), reflecting the noisy prediction scheme.

Our primary goal with the measured data was to validate and test the modeling pipeline. From this small sample size, we do not make strong neurophysiological conclusions. The neurophysiological results presented here are also not to be generalized to other cortical regions or different TMS parameters such as different stimulation intensities, or interstimulus intervals. On the other hand, the presented methods are highly suited for studying how the excitability state encoding changes when TMS parameters are changed or the subject is performing a task, *e.g.*, motor imagery.

Here, ISI was chosen to be around 2 s to minimize the measurement duration and to keep the EEG as stationary as possible. Previous TMS pulses in the session might have gradually changed the brain-state encoding via plasticity, but according to our tests, classification gave the accuracy of on average 64% even when predicting the excitability within the latter half of the experimental trials using the first half of them for training, which is close to the benchmarking accuracy, suggesting that the excitability encoding stays rather stable throughout the measurement.

4.4. Outlook for individualized online classification

An important potential application of the predictable brain excitability states is closed-loop TMS to induce therapeutic changes in brain networks. To enhance a therapy protocol, the prediction accuracy may not need to be very high: If, say, 65% of the stimuli are delivered during the time windows of high excitability (and thus high plasticity), this would make the therapy 30% more effective as compared to the random stimulation, provided that the effectiveness is directly proportional to the number of times when high-excitability state is reached.

For successful classification in a closed-loop experiment, two critical methodological steps need to be implemented and tested: 1) online EEG preprocessing, and 2) training filters for upcoming trials within an experiment, or even training filters for future experiments based on data recorded from different individuals. Here, we explored the prerequisites for the latter task.

The test accuracy for predicting the brain-state with the first half of the data to be used within the latter half of the data was slightly less than from the benchmarking test. The difference is rather small, suggesting that the excitability encoding patterns are relatively stable in time. It is known from previous BCI studies that the classification accuracy tends to decrease over time, and the estimated filters should be re-adapted to the changes in EEG (Blankertz et al., 2008; Shenoy et al., 2006; Zhao et al., 2008). In an online setting, the training could be enhanced by estimating an initial filter estimate with few trials and adaptively optimize the classifier to the changing data (Blankertz et al., 2007; Hill et al., 2006; Shenoy et al., 2006; Zhao et al., 2008).

We also discovered that the number of EEG channels could be greatly decreased while still keeping the accuracy unchanged, which would be useful for online, and especially for clinical usage. A small number of electrodes has been reported sufficient for CSP also earlier (Farquhar et al., 2006). The results indicate that, *e.g.*, 30 channels located sparsely around the relevant cortical region could suffice provided that the SNR of the data is good. Depending on the data quality, effective cleaning of the data online may still require more channels.

The applicability of same ML-optimized filters over different subjects did not show promising because the obtained accuracy was just slightly above 50%. This is not surprising because the SNR of the relevant spatio-temporal patterns among EEG are rather low in the resting-state condition, and the estimated filters are individually adapted to cancel out irrelevant heterogeneous and variable ongoing EEG oscillations. Even though this straightforward attempt of training filters across subjects did not seem profitable, there may be more efficient ways of using the prior information of the spatio-temporal patterns at the sensor- or source-level to enhance filter optimization for individual recordings: If consistent population-level patterns exist, they could be used to guide individualized modeling as prior information to stabilize and speed up the training (Cheng et al., 2017; Lu et al., 2010).

4.5. Conclusion

Cortical excitability states can be modelled in an individual way using supervised learning of the spatio-temporal filter applied to the resting-state EEG epoch preceding a TMS pulse. When applied to the data recorded from healthy subjects, the cortical excitability state could be predicted more accurately than using a fixed-spatial-filter model. The estimated spatio-temporal filter could be used to study how the cortical excitability states are encoded by cortical activity at an individual level and under different conditions. Excitability prediction with individualized EEG filters also pertains to personalized TMS therapy with the potential to increase therapeutic effect sizes.

Declaration of Competing Interest

CZ is a shareholder in a spin-off start-up company, sync2brain GmbH, Tübingen, Germany.

Data and code availability statement

The source code of the scripts used for the analysis are available from the lab's GitHub repository: www.github.com/bnplab/causaldecoding

The data are available upon request. There is no consent from the study participants to make the data publicly available.

Author contributions

JM developed the decoding algorithm, analyzed the experimental data, conceptualized and performed the simulated data experiment, created the figures, and wrote the first draft of the manuscript. CZ and UZ conceptualized the research question. CZ, PB, JM and ME performed the TMS-EEG measurements. PB generated the MRI-based head models. All authors contributed to the interpretation of the results and to the final manuscript.

Acknowledgments

Dragana Galevska, Niko Mangold and Sinan Barus supported conducting the experiments. Natalie Schaworonkow contributed to the conceptualization of the research question and developed a proof-of-concept decoding algorithm.

This work was supported by an EXIST Transfer of Research grant by the German Federal Ministry for Economic Affairs and Energy (NEUROSYNC) [grant number 03EFJBW169], by the European Research Council (ERC Synergy) under the European Union's Horizon 2020 research and innovation program (ConnectToBrain) [grant number 810377], by the KAUTE Foundation, and by the Emil Aaltonen Foundation. The funding sources did not participate in conducting the study at any phase.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118652.

Appendices

A. Laplacian trendline fitting and line noise removal

Before detrending, in EMG data, the time window of the MEPs and the artifacts was manually identified and linear interpolation was applied in this interval. The remaining stationary part of the data are contained in the set of indices $\tau_{\text{stat}} = \{1, \ldots, t_1, t_2, \ldots, T\}$, where the time indices t_1 and t_2 define the start and the end points of the interpolation, respectively. The trendline was estimated for each channel in each data epoch of EEG and EMG by:

$$\mathbf{x}_{\text{trend}} = (\boldsymbol{I} + \boldsymbol{C}\delta)^{-1}\mathbf{x}_{\text{interp}}, \tag{A1}$$

where $(T \times 1)$ $\mathbf{x}_{\text{trend}}$ is the estimated trend, $(T \times 1)$ $\mathbf{x}_{\text{interp}}$, is the interpolated signal, δ is a regularization factor, and $(T \times T)$ $\boldsymbol{C}$ is the (auto-)covariance matrix of the detrended signal. We set $\boldsymbol{C} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$, where $\boldsymbol{P}$ the Laplacian matrix operator:

$$\boldsymbol{P} = \begin{bmatrix} 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ & \vdots & & \ddots & & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix}. \tag{A2}$$

Eq. (A.1) corresponds to trendline computation by minimum mean-squared error estimator assuming that $\boldsymbol{I}$ and $\mathbf{C}$ are the (auto-)covariance matrices of the trendline and the detrended signals, respectively, and that they are uncorrelated with each other. The regularization factor was set to $\delta = 10^5$ and 10^4 for the sampling frequencies of 1 kHz (EEG) and 5 kHz (EMG), respectively. Detrended signal $\tilde{\mathbf{x}}$ was obtained by subtracting the trendline from the original data (before interpolation).

After trendline rejection, 50 Hz noise was identified and eliminated. First, we defined the time-domain subspace of the line noise spanned by the columns of

$$\boldsymbol{U} = \begin{bmatrix} \sin\left(50 \cdot 2\pi \cdot t_1/f_{\text{s}}\right) & \cos\left(50 \cdot 2\pi \cdot t_1/f_{\text{s}}\right) \\ \vdots & \vdots \\ \sin\left(50 \cdot 2\pi \cdot t_N/f_{\text{s}}\right) & \cos\left(50 \cdot 2\pi \cdot t_N/f_{\text{s}}\right) \end{bmatrix}, \tag{A3}$$

where t_i runs over N selected time indices, and f_{s} is the sampling frequency. When the time indices run over the stationary interval τ_{stat} only, we denote $\boldsymbol{U} \leftarrow \boldsymbol{U}_{\text{stat}}$, and when the indices run over the whole epoch, we denote $\boldsymbol{U} \leftarrow \boldsymbol{U}_{\text{all}}$.

The coefficients in (2×1) vector $\mathbf{m}$, describing the 50 Hz –noise content in the signal, were fitted using the stationary data:

$$\mathbf{m} = \boldsymbol{U}_{\text{stat}}^{\mathrm{T}}\tilde{\mathbf{x}}_{\text{stat}}, \tag{A4}$$

after which line noise was removed from the original signal assuming stationarity of the line noise:

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}} \; - -\boldsymbol{U}_{\text{all}}\mathbf{m}. \tag{A5}$$

B. Exponential-curve fitting

In each EMG trial and channel, the fitting of the exponential decay is performed separately to the signal samples $\mathrm{x}(t)$, where t indices time over the fitting interval. These intervals were manually chosen representative time windows before and after the MEP response, where the exponential artifact dominated.

Matlab's 'fminunc' function was used to fit the exponential function $\mathrm{A}\, e^{-\mathrm{k}\,\cdot(t-t_0)}$, where t is the running time index, t_0 the earliest time of the fitting window, and A and k are parameters to be fitted. Fitting was performed such that the mean squared error between the original and the signal and the fitted exponential signal is minimized, *i.e.,*

$$\{\mathrm{A}, \mathrm{k}\} = \arg\min_{\mathrm{A},\,\mathrm{k}} \left\langle \left\| (t) - \mathrm{A}\, e^{-\mathrm{k}\,\cdot(t-t_0)} \right\| \right\rangle_t, \tag{B1}$$

where $\cdot_t$ denotes mean operation over the fitting intervals. The initial guesses for the parameters were set as follows: $\mathrm{A} = \mathrm{x}\,(t = t_0)$, *i.e.,* the signal value at the start of the fitting window, and $\mathrm{k} = {}^{1}/_{t'} - t_0$, where t' is set the smallest time index for which $\mathrm{x}(t) < \mathrm{A}e^{-1}$.

C. Spherical-head model

The spherical head model consisted of three concentric layers: the brain, the skull, and the scalp with the radii of 81, 85, and 88 mm, and the relative conductivities of 1, 0.8, and 1, respectively. The standard electrode positions were set as obtained from EEGLAB. Current dipoles inside the brain 76 mm away from the origin were used as the source space. The equations needed to compute the spherical-model lead fields can be found in [31].

References

Arieli, A., Sterkin, A., Grinvald, A., Aertsen, A., 1996. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273, 1868–1871.
Barker, A.T., Jalinous, R., 1985. Non-Invasive Magnetic Stimulation of Human Motor Cortex. Lancet 1, 1106–1107.
Baur, D., Galevska, D., Hussain, S., Cohen, L.G., Ziemann, U., Zrenner, C., 2020. Induction of LTD-like corticospinal plasticity by low-frequency rTMS depends on pre-stimulus phase of sensorimotor mu-rhythm. Brain Stimul. 13, 1580–1587.
Betzel, R.F., Fukushima, M., He, Y., Zuo, X.N., Sporns, O., 2016. Dynamic fluctuations coincide with periods of high and low modularity in resting-state functional brain networks. Neuroimage 127, 287–297.
Blankertz, B., Lemm, S., Treder, M., Haufe, S., Muller, K.R., 2011. Single-trial analysis and classification of ERP components - a tutorial. Neuroimage 56, 814–825.
Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Muller, K.R., 2008. Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process. Mag. 25, 41–56.
Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F.U., Nikulin, V.V., Müller, K.R., 2007a. Invariant Common Spatial Patterns: alleviating Nonstationarities in Brain–Computer Interfacing. In: NIPS, pp. 113–120.
Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Muller, K.R., 2007b. Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process. Mag. 25, 41–56.
Burnes, J.E., Kaelber, D.C., Taccardi, B., Lux, R.L., Ershler, P.R., Rudy, Y., 1998. A field–compatible method for interpolating biopotentials. Ann. Biomed. Eng. 26, 37–47.
Buzsaki, G., Anastassiou, C.A., Koch, C., 2012. The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. Nat. Rev. Neurosci. 13, 407–420.
Canolty, R.T., Ganguly, K., Kennerley, S.W., Cadieu, C.F., Koepsell, K., Wallis, J.D., Carmena, J.M., 2010. Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. Proc. Natl. Acad. Sci. U.S.A. 107, 17356–17361.
Cheng, M.M., Lu, Z.H., Wang, H.X., 2017. Regularized common spatial patterns with subject-to-subject transfer of EEG signals. Cogn. Neurodyn. 11, 173–181.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.
de Cheveigne, A., Arzounian, D., 2018. Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. Neuroimage 172, 903–912.
de Goede, Annika A.A., van Putten, Michel M.J.A.M., 2019. Infraslow activity as a potential modulator of corticomotor excitability. J. Neurophysiol. 122, 325–335.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.
Ellaway, P.H., Davey, N.J., Maskill, D.W., Rawlinson, S.R., Lewis, H.S., Anissimova, N.P., 1998. Variability in the amplitude of skeletal muscle responses to magnetic stimulation of the motor cortex in man. Electromyogr.Motor Control-Electroencephalogr. Clin. Neurophysiol. 109, 104–113.
Farquhar, J., Hill, N.J., Lal, T.N., Schölkopf, B., 2006. Regularised CSP for Sensor Selection in BCI. In: 3rd International Brain-Computer Interface Workshop and Training Course 2006. Verlag der Technischen Universität Graz, pp. 14–15.
Funase, K., Miles, T.S., Gooden, B.R., 1999. Trial-to-trial fluctuations in H-reflexes and motor evoked potentials in human wrist flexor. Neurosci. Lett. 271, 25–28.
Groppa, S., Oliviero, A., Eisen, A., Quartarone, A., Cohen, L.G., Mall, V., Kaelin-Lang, A., Mima, T., Rossi, S., Thickbroom, G.W., Rossini, P.M., Ziemann, U., Valls-Sole, J., Siebner, H.R., 2012. A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. Clin. Neurophysiol. 123, 858–882.
Haegens, S., Nacher, V., Luna, R., Romo, R., Jensen, O., 2011. alpha-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. Proc. Natl. Acad. Sci. U.S.A. 108, 19377–19382.
Hallett, M., 2007. Transcranial magnetic stimulation: a primer. Neuron 55, 187–199.
Hamada, M., Murase, N., Hasan, A., Balaratnam, M., Rothwelll, J.C., 2013. The role of interneuron networks in driving human motor cortical plasticity. Cerebral Cortex 23, 1593–1605.
Haufe, S., Meinecke, F., Gorgen, K., Dahne, S., Haynes, J.D., Blankertz, B., Biessgmann, F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage 87, 96–110.
Hill, N.J., Farquhar, J., Lal, T.N., Schölkopf, B., 2006. Time-dependent demixing of task-relevant EEG signals. In: Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course, pp. 20–21.
Hordacre, B., Goldsworthy, M.R., Vallence, A.M., Darvishi, S., Moezzi, B., Hamada, M., Rothwell, J.C., Ridding, M.C., 2017. Variability in neural excitability and plasticity induction in the human cortex: a brain stimulation study. Brain Stimul 10, 588–595.
Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. Ieee Trans. Neural Netw. 10, 626–634.
Kallioniemi, E., Kononen, M., Saisanen, L., Grohn, H., Julkunen, P., 2015. Functional neuronal anisotropy assessed with neuronavigated transcranial magnetic stimulation. J. Neurosci. Methods 256, 82–90.
Kar, S.K., 2019. Predictors of Response to Repetitive Transcranial Magnetic Stimulation in Depression: a Review of Recent Updates. Clin. Psychopharmacol. Neurosci. 17, 25–33.
Keil, J., Timm, J., SanMiguel, I., Schulz, H., Obleser, J., Schonwiesner, M., 2014. Cortical brain states and corticospinal synchronization influence TMS-evoked motor potentials. J. Neurophysiol. 111, 513–519.
Kempter, R., Leibold, C., Buzsaki, G., Diba, K., Schmidt, R., 2012. Quantifying circular-linear associations: hippocampal phase precession. J. Neurosci. Methods 207, 113–124.
Kiers, L., Cros, D., Chiappa, K.H., Fang, J., 1993. Variability of Motor Potentials-Evoked by Transcranial Magnetic Stimulation. Electroencephalogr. Clin. Neurophysiol. 89, 415–423.
Kleinbaum, D.G., Dietz, K., Gail, M., Klein, M., Klein, M., 2002. Logistic Regression. Springer.
Koponen, L.M., Nieminen, J.O., Ilmoniemi, R.J., 2018. Multi-locus transcranial magnetic stimulation-theory and implementation. Brain Stimul. 11, 849–855.
Lawson, C.L., Hanson, R.J., 1995. Solving Least Squares Problems. SIAM.
Lopez-Alonso, V., Cheeran, B., Rio-Rodriguez, D., Fernandez-Del-Olmo, M., 2014. Inter-individual variability in response to non-invasive brain stimulation paradigms. Brain Stimul. 7, 372–380.
Lotte, F., Guan, C., 2011. Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. IEEE Trans. Biomed. Eng. 58, 355–362.
Lu, H.P., Eng, H.L., Guan, C.T., Plataniotis, K.N., Venetsanopoulos, A.N., 2010. Regularized Common Spatial Pattern With Aggregation for EEG Classification in Small-Sample Setting. Ieee Trans. Biomed. Eng. 57, 2936–2946.
Magistris, M.R., Rosler, K.M., Truffert, A., Landis, T., Hess, C.W., 1999. A clinical study of motor evoked potentials using a triple stimulation technique. Brain 122, 265–279.
Miranda, P.C., deCarvalho, M., Conceicao, I., Luis, M.L.S., DuclaSoares, E., 1997. A new method for reproducible coil positioning in transcranial magnetic stimulation mapping. Electromyogr. Motor Control-Electroencephalogr. Clin. Neurophysiol. 105, 116–123.
Mitchell, W.K., Baker, M.R., Baker, S.N., 2007. Muscle responses to transcranial stimulation in man depend on background oscillatory activity. J. Physiol.-Lond. 583, 567–579.
Moiseev, A., Gaspar, J.M., Schneider, J.A., Herdman, A.T., 2011. Application of multi–source minimum variance beamformers for reconstruction of correlated neural activity'. Neuroimage 58, 481–496.
Mutanen, T.P., Metsomaa, J., Liljander, S., Ilmoniemi, R.J., 2018. Automatic and robust noise suppression in EEG and MEG: the SOUND algorithm. Neuroimage 166, 135–151.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869.
Pellicciari, M.C., Miniussi, C., Ferrari, C., Koch, G., Bortoletto, M., 2016. Ongoing cumulative effects of single TMS pulses on corticospinal excitability: an intra- and inter-block investigation. Clin. Neurophysiol. 127, 621–628.
Sauseng, P., Klimesch, W., Gerloff, C., Hummel, F.C., 2009. Spontaneous locally restricted EEG alpha activity determines cortical excitability in the motor cortex. Neuropsychologia 47, 284–288.
Schaworonkow, N., Gordon, P.C., Belardinelli, P., Ziemann, U., Bergmann, T.O., Zrenner, C., 2018. mu-Rhythm Extracted With Personalized EEG Filters Correlates With Corticospinal Excitability in Real-Time Phase-Triggered EEG-TMS'. Front. Neurosci. 12.
Schaworonkow, N., Triesch, J., Ziemann, U., Zrenner, C., 2019. EEG-triggered TMS reveals stronger brain state-dependent modulation of motor evoked potentials at weaker stimulation intensities. Brain Stimul. 12, 110–118.
Schrader, L., Sadeghinejad, S., Sadeghinejad, J., Kazanchyan, M., Koski, L., Stern, J., Wu, A., Iacoboni, M., Nuwer, M., 2016. Comparison of low frequency repetitive transcranial magnetic stimulation parameters on motor cortex excitability in normal subjects. Int. J. Epilepsy 3, 2–6.
Shenoy, P., Krauledat, M., Blankertz, B., Rao, R.P.N., Muller, K.R., 2006. Towards adaptive classification for BCI. J. Neural Eng. 3, R13–R23.
Sun, WeiW., Mao, Wei W., Meng, Xianghong X., Wang, Dequan D., Qiao, Liang L., Tao, Wei W., Li, Liping L., Jia, Xiaoying X., Han, Chunyu C., Fu, Mengmeng, Tong, Xiaoyan X., Wu, Xun X., Wang, Yuping Y., 2012. Low-frequency repetitive transcranial magnetic stimulation for the treatment of refractory partial epilepsy: a controlled clinical study. Epilepsia 53, 1782–1789.
Thies, M., Zrenner, C., Ziemann, U., Bergmann, T.O., 2018. Sensorimotor mu-alpha power is positively related to corticospinal excitability. Brain Stimul 11, 1119–1122.
Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. J. R. Stat. Soc. Ser. B-Methodol. 58, 267–288.
Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., U-Minn, W., Consortium, H.C.P., 2013. The WU-Minn Human Connectome Project: an overview. Neuroimage 80, 62–79.
Varoquaux, G., Raamana, P.R., Engemann, D.A., Hoyos-Idrobo, A., Schwartz, Y., Thirion, B., 2017. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. Neuroimage 145, 166–179.
Zhao, Q.B., Zhang, L.Q., Cichocki, A., Li, J., 2008. Incremental Common Spatial Pattern Algorithm for BCI. 2008 Ieee Int. Joint Conf. Neural Netw. 1-8 Vols2656-+.
Zoefel, B., Davis, M.H., Valente, G., Riecke, L., 2019. How to test for phasic modulation of neural and behavioural responses. Neuroimage 202.
Zrenner, B., Zrenner, C., Gordon, P.C., Belardinelli, P., McDermott, E.J., Soekadar, S.R., Fallgatter, A.J., Ziemann, U., Muller-Dahlhaus, F., 2020a. Brain oscillation-synchronized stimulation of the left dorsolateral prefrontal cortex in depression using real–time EEG-triggered TMS. Brain Stimul. 13, 197–205.
Zrenner, C., Desideri, D., Belardinelli, P., Ziemann, U., 2018. Real-time EEG-defined excitability states determine efficacy of TMS-induced plasticity in human motor cortex. Brain Stimul. 11, 374–389.
Zrenner, C., Galevska, D., Nieminen, J.O., Baur, D., Stefanou, M.I., Ziemann, U., 2020b. The shaky ground truth of real-time phase estimation. Neuroimage 214, 116761.