Journal of Neural Engineering

PURPOSE-LED PUBLISHING™

ACCEPTED MANUSCRIPT

Towards subject-centered co-adaptive brain-computer interfaces based on backward optimal transport

To cite this article before publication: Victoria Peterson *et al* 2025 *J. Neural Eng.* in press https://doi.org/10.1088/1741-2552/addb7a

Towards subject-centered co-adaptive brain-computer interfaces based on backward optimal transport

Victoria Peterson[1,2*], Valeria Spagnolo[1], Catalina Galván[1,2], Nicolás Nieto[1,3], Rubén D. Spies[1,2], Diego H. Milone[3]

[1]Instituto de Matemática Aplicada del Litoral, IMAL, UNL, CONICET, Santa Fe, Argentina.

[2]Departamento de Matemática, Facultad de Ingeniría Química, UNL, Santa Fe, Argentina.

[3]Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i), UNL, CONICET, FICH, Santa Fe, Argentina.

E-mail: vpeterson@santafe-conicet.gov.ar

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Abstract. *Objective.* Controlling a motor imagery brain-computer interface (MI-BCI) can be challenging, requiring several sessions of practice. Electroencephalography (EEG)-based BCIs are particularly affected by cross-session variability. In this scenario, it is crucial to implement co-adaptive systems, where the machine adapts the decoding algorithm while the user learns how to control the BCI. To support the user learning process, it is essential to measure and provide real-time feedback on self-modulation skills. This study aims to develop a method for online assessment of MI modulation capability to build co-adaptive BCIs that improve both user performance and system accuracy. *Approach.* Backward optimal transport for domain adaptation allows across-session MI-BCI usage without classifier retraining. Using the cued label to guide the adaptation, a supportive backward adaptation (SBA) method is defined. The required model effort to perform a trial adaptation is proposed as an online metric of MI modulation skills. We conducted experiments on both real and simulated data to demonstrate that this metric effectively informs about the the discriminability and stability of the EEG patterns related to the MI task. The proposed metric is validated by means of Riemannian distinctiveness metrics. *Main Results.* Our findings show that the associated effort when applying SBA provides a meaningful way of evaluating EEG patterns discriminability, being significantly correlated with Riemannian distinctiveness metrics. *Significance.* This study introduces a novel framework for co-adaptive BCI learning that performs data adaptation while assessing the MI-BCI skills of the user. The proposed SBA approach can enhance BCI performance by facilitating session-to-session adaptation and empowering users with valuable feedback based on their current MI modulation strategy. This framework represents a significant advancement in developing user-centered, co-adaptive MI-BCIs that effectively support and enhance user capabilities.

Keywords: Domain adaptation, co-adaptive BCI, BCI skills, user-centered BCI, Optimal Transport.
Submitted to: *J. Neural Eng.*

1. Introduction

Brain-computer interfaces (BCIs) provide a digital link between a brain and a device. In particular, BCIs based on the mental imagination of movements, i.e. motor imagery (MI), strongly rely on the subject's ability to self-regulate their brain activity [1]. In this context, machine learning (ML) models built upon brain-recorded data should be able to identify both the optimal feature space and the classifier decision boundary to effectively decode the brain activity related to the intended mental task [2]. While the acquisition of MI-BCI control skills requires practice over several sessions of use [3], the ML decoding algorithm should adapt and support the subject's learning process [4]. This simultaneous subject-machine adaptation defines what is known as co-adaptive learning [5].

From the ML perspective, co-adaptive learning involves adjusting the parameters of the decoding algorithm to address changes in data distribution between training and testing datasets. In BCI, training and testing datasets may be referred to as calibration and application sessions, which typically are registered at different days. Several aspects may influence the stability between sessions of MI-BCI. Psychological and physiological factors, such as level of fatigue, arousal and workload can influence the user capability to control the BCI [2, 6, 7]. Moreover, due to potential electrode-cap misalignment and electrode impedance values drift between sessions [8], non-invasive MI-BCI systems based on surface electroencephalography (EEG) are particularly susceptible to data distribution changes. Thus, from the ML decoding algorithm viewpoint, MI-BCIs based on EEG define a non-stationary learning environment [9].

Domain adaptation provides a suitable framework to address changes in observable non-stationary data [10]. In the context of BCIs, the variability across sessions can be managed by "adapting" the feature distribution from one domain (calibration session) to another domain (application session). Recently, we presented a backward formulation of optimal transport for domain adaptation (BOTDA) [11]. This method transforms data distribution of an application session to match the calibration data distribution, avoiding the traditional need for classifier retraining. Experiments mimicking online adaptation in MI-BCI for motor rehabilitation applications, showed that the use of a supervised BOTDA implementation significantly improved decoding performance of an already trained classifier, as well as state-of-the-art data alignment methods, including complete model retraining [11]. Interestingly, for some subjects, decoding performance resulted in similar fashion to a model without any adaptation. These results open a unique opportunity to study the relationship between adaptation success and the user capability to control MI-BCIs [12].

However, unveiling MI-BCI user skills is still an important challenge in the field. Efforts have been made in order to anticipate users' cognitive skills to self-regulate sensorimotor rhythms by using neurophysiological markers [1, 13], studying visual rotation capabilities [14] as well as other cognitive and personality profiles tests ([14–16]). Nevertheless, these metrics may not reflect the changes related to the user learning

process. Performance metrics from fixed (non-adaptive) ML classifiers can be calculated across sessions; however, they fail to capture the user's MI-BCI control skills. In this direction, the authors in [4] proposed new metrics to measure MI-BCI skills based on covariance matrices in the Riemannian space. These metrics were designed to be independent of the given classifier, and thus only reflect the MI patterns in the EEG data. Among them, class distinctiveness was proposed to measure how distinguishable within a session MI-EEG patterns are from each other, showing to be correlated with the spatial visual abilities of the participants, the mayor predictor of MI-BCI skills [17]. Nevertheless, user's MI-BCI skills should be measured online to provide appropriate feedback that would help them in better self-modulate their brain patterns.

While in MI-BCIs the classifier's performance depends on both the user's capability to generate distinct and consistent EEG patterns and the optimized ML model parameters [18], most current closed-loop BCIs offer feedback to the user based solely on the model output. Nonetheless, the classifier's effectiveness will only reflect the user's ability to operate the BCI if the model is reliable, that is, its complexity, number of trainable parameters, training strategy and used data are appropriate for the classification task. For building co-adaptive closed-loop BCIs, not only the method should adapt to the user as they learn how to control the BCI, but also the delivered feedback should be based on their generated EEG patterns. As such, feedback is expected to be informative about users' current MI skill levels, helping them develop better BCI control capabilities [12]. In this regard, in co-adaptive BCIs we argue that the algorithmic support, that is, the effort the model needs to make to perform the adaptation, should be measured and delivered as feedback to the BCI user to inform them in real-time about the quality of their EEG patterns. In this context, let us consider the following example, in which the performance of two users in MI-BCI utilizing a co-adaptive algorithm is compared. For a "high-performing" user, the co-adaptive method achieved a final accuracy of 90%, starting from a baseline accuracy of 80% using a basic (no adaptive) ML model. In contrast, for a "lower-performing" user, the same final accuracy of 90% was reached, but the initial accuracy of the basic ML model was only 60%. In the latter case, the effort in performing the adaptation was higher to that of the high-performing user. Thus, we argue that by measuring the adaptation cost, real-time feedback could be delivered to the users to help them improve their MI modulation strategies effectively.

In the present work, we show that our supervised backward domain alignment method constitutes a reliable online adaptive framework. In particular, we demonstrate that the effort required to transform a given trial so as to match the distribution of the calibration data offers a novel way for online assessment of the discriminability and stability of EEG patterns. We shall refer to this method as supportive backward adaptation (SBA). SBA not only facilitates successful adaptation between sessions in MI-BCIs, but it also enables the online evaluation of the user's ability to produce distinguishable and stable EEG patterns. This opens a new venue in co-adaptive BCI learning, where meaningful feedback based on adaptation-related information can be

provided to the users, guiding them throughout the BCI learning process. The present study represents a step forward in developing user-centered co-adaptive closed-loop MI-BCIs.

2. Co-adaptive learning based on backward optimal transport

This section provides a brief review of backward optimal transport for domain adaptation (BOTDA) between sessions of BCI. Implementation details for supervised transfer learning are also given. The new metric for MI-BCI skills assessment based on the associated effort in performing the data adaptation is then presented. The supervised BOTDA implementation together with the MI-BCI skills assessment metric based on the adaptive cost constitute SBA.

2.1. Backward supervised online adaptation method

Backward adaptation consists of coping with changes in data distribution between two BCI sessions, while avoiding classifier retraining [11]. In this context, it is said that data used to train the classifier belongs to the source domain (Ω_s), while the application or testing data are drawn from the target domain (Ω_t). The goal of BOTDA is to adapt or transform features of the target domain so that the distribution become similar to the source domain feature distribution, i.e., where the classifier was trained. Assuming that the feature distribution drift is given by a certain transformation $\mathcal{B} : \Omega_t \to \Omega_s$, BOTDA consists of estimating $\mathcal{B}$ in a cost-effective manner [19].

BOTDA is applied at the feature space level, that is, after features are extracted. Regardless of the method used for learning the feature representation, BOTDA seeks to match the distribution of two datasets, the source ($\mathcal{S}$) and the target ($\mathcal{T}$) datasets. Mathematically, these two datasets can be described as a set of pair data:

$$\mathcal{S} \doteq \{(\mathbf{x}_i^s, y_i^s), i = 1, \ldots, N_s\} \subset \Omega_s \times \mathcal{K},$$
$$\mathcal{T} \doteq \{(\mathbf{x}_i^t, y_i^t), i = 1, \ldots, N_t\} \subset \Omega_t \times \mathcal{K},$$

where N_s and N_t denote the number of trials in the source and target domain, respectively, $\mathbf{x}_i^s \in \Omega_s, \mathbf{x}_i^t \in \Omega_t$ are the feature vectors, with Ω_s, $\Omega_t \subset \mathbb{R}^d$, $\mathcal{K} \doteq \{k_j\}_{j=1}^K$ and $y_i^s, y_i^t \in \mathcal{K}$ represent the indicated mental tasks belonging to K possible classes.

When working with discrete variables, estimating $\mathcal{B}$ results in finding the optimal transportation plan γ^* to minimize a certain cost function m. Then, defining $M = (m_{i,j})$, $m_{i,j} \geq 0$, we have:

$$\gamma^* \doteq \underset{\gamma}{\operatorname{argmin}} \langle \gamma, M \rangle_F, \tag{1}$$

where $\langle \cdot, \cdot \rangle_F$ denotes the Frobenius inner product.

Note that M is a matrix such that $m_{i,j} = m(\mathbf{x}_i^t, \mathbf{x}_j^s)$ represents the cost of moving a unit probability mass from $\mathbf{x}_i^t \in \Omega_t$ to $\mathbf{x}_j^s \in \Omega_s$. Generally, m is chosen as the square of the Euclidean distance [20].

To relax the optimization procedure and to include label information during the transportation plan estimation, two regularization terms $W_e(\cdot)$ and $W_c(\cdot)$ are added to the functional in (1), as follows:

$$\gamma^* \doteq \underset{\gamma}{\operatorname{argmin}} \langle \gamma, M \rangle_F + \lambda W_e(\gamma) + \nu W_c(\gamma), \tag{2}$$

where $\lambda > 0$ and $\nu > 0$ are regularization parameters that weight the contribution of each penalization term. Here $W_e(\gamma) \doteq \sum_{ij} \gamma_{i,j} \log(\gamma_{i,j})$ accounts for the negentropy, while $W_c(\gamma) \doteq \sum_j \sum_k \|\gamma_{\mathcal{I}_k,j}\|_2$, where $\mathcal{I}_k$ denotes the set of indices that belongs to class $k \in \mathcal{K}$, promotes transporting "together" samples with the same label. The problem defined by (2) is known as the Sinkhorn group-LASSO optimal transport problem.

The optimal transport plan γ^*, estimated by means of (2), is a matrix of the same size as the cost matrix M, which describes how to distribute the mass probability between the domains. In particular, in the backward formulation, each element $\gamma^*_{i,j}$ indicates how much probability mass from the target sample $\mathbf{x}^t_i$ is transferred to the source sample $\mathbf{x}^s_j$. Thus, in the context of online adaptation, and noting that the last two terms in (2) are independent of $\mathbf{x}$, for a given ℓ^{th} trial from the application session, transportation is the result of the following mapping:

$$\hat{\mathbf{x}}^t_\ell = \mathcal{B}_{\gamma^*}(\mathbf{x}^t_\ell) \doteq \underset{\mathbf{x}\in\mathbb{R}^d}{\operatorname{argmin}} \sum_{j=1}^{N_s} \gamma^*_{\ell,j}\, m(\mathbf{x}, \mathbf{x}^s_j). \tag{3}$$

When the ℓ^{th} target sample is used during the optimal transport plan estimation, the transformation defined in (3) results in a simple weighted barycenter mapping. That is, $\hat{\mathbf{x}}^t_\ell = \sum_{j=1}^{N_s} \gamma^*_{\ell,j}\, \mathbf{x}^s_j$.

Considering real-life BCI applications, the changes between the source and the target domains must be estimated by using only a portion of data (as small as possible) of the current application session. In most BCIs setups, a few trials at the beginning of the session are registered without feedback; this is usually referred to as recalibration phase. The resulting recalibration data is used to learn the transportation plan from the target to the source domain. Therefore, in the sequel we shall refer to the recalibration data as the *transportation set*, and will be denoted by $\mathcal{V} \doteq \{(\mathbf{x}^v_i, y^v_i)\}_{i=1}^{N_v} \subset \mathcal{T}$, where $N_v \ll N_t$. Online adaptation based on BOTDA was proposed by adding the current ℓ^{th} trial to the transportation set, that is:

$$\mathcal{V}_\ell \doteq \left\{(\mathbf{x}^t_i, y^t_i)\right\}_{i=1}^{N_v} \cup \left\{(\mathbf{x}^t_\ell, y^t_\ell)\right\} \subset \mathcal{T}, \quad \forall \ell = N_v + 1, \ldots, N_t. \tag{4}$$

From a transfer-learning viewpoint, finding the minimizer in (2) for a transportation set as defined in (4) constitutes a supervised domain adaptation problem [10]. In the context of the present study, we shall refer to this supervised BOTDA implementation, as supportive backward adaptation (SBA). It is timely to note that the class label of the current trial, as reported by the BCI system, does not necessarily reflect the mental task performed by the user. Therefore, accounting for the trial cued label does not necessarily imply that after adaptation such a trial will be correctly classified, as we argue and demonstrate next.

2.2. Supportive backward adaptation to guide the user learning process

As previously explained, SBA performs supervised domain adaptation at the feature space level making use of the trial label to estimate the optimal transport plan. Nevertheless, in synchronous MI-BCIs, the "label" represents the instruction (cue), rather than the actual class label of the user-generated EEG patterns, which may or may not align with the instructed task.

Based on the hypothesis that SBA can successfully perform the adaptation when the user-generated EEG patterns belong to the cued mental task, it is reasonable to think that the associated cost of transporting a trial from the current application session to the calibration one, contain valuable information about the stability and discriminability of the EEG trial *per-se*. In fact, let us closely inspect (2). When the optimal plan $\gamma^* \in \mathbb{R}_+^{N_{v+1} \times N_s}$ is computed for $\mathcal{V}_\ell$ as defined in (4), the fidelity term of the optimal transport functional in (2), known as the Wasserstein distance, reads as follows:

$$W = \sum_{i,j} \gamma^*_{i,j} m_{i,j}, \tag{5}$$

where $\gamma^*_{i,j}$ are the components of γ^*. Thus, every sample i in the target domain is transported as indicated by the optimal plan $\gamma^*_{i,j}$ given a cost $m_{i,j}$ to the j sample in the source domain. In other words, W quantify the amount of "work" done in order to make the sample distribution of the target domain resemble that of the source domain. Considering the way the transportation plan is learned (see (4)), we propose to quantify the algorithmic support in adapting the current ℓ^{th} trial by h, defined by:

$$h \doteq \sum_j \gamma^*_{\ell,j} m_{\ell,j} =< \boldsymbol{\gamma}^*_\ell, \mathbf{m}_\ell >, \tag{6}$$

where $\boldsymbol{\gamma}^*_\ell$ and $\mathbf{m}_\ell$ are the ℓ^{th} rows of the optimal transportation plan γ^* and of the cost matrix M, respectively.

By means of h, we are able to measure the effort exerted by SBA when performing the data distribution matching of a given trial. We shall refer to this measure as *algorithmic support*. It is expected that when the EEG patterns generated by the user are aligned with the indicated mental class, the class label information provided by the system helps the adaptation to be successful. That is, the class of the current trial after adaptation is correctly predicted by the classifier. In this context, it is expected that h be higher for less distinguishable and stable EEG patterns, resulting in more effort made by the model during the online adaptation process. The proposed algorithmic support index works as an online user skills metric which does not require of extra methods or layers of analysis upon the adaptive algorithm itself. Since it is applied at the feature space level, it can be used within any decoding pipeline, including end-to-end deep learning architectures.

3. Leveraging SBA for cross-session MI-BCI

In the present work we focus on a two-class subject-specific MI-BCI system used across sessions. The common spatial pattern (CSP) in conjunction with the linear discriminant analysis (LDA) are used to build the decoding algorithm [21]. As commonly done, the feature extractor as well as the discriminative model are trained following an intra-subject scheme using data from the calibration session (source domain). Data from a new BCI session, where the model is applied, defines the target domain. The subset of the first N_v trials of the target domain defines the recalibration set [11].

The initial step of SBA involves training both the feature extractor and the classifier using the complete calibration session data (training data). Then, both optimized models will always be kept fixed. To maintain the sample proportion between the source and the target domain when learning the optimal transport plan, a subset of the source data equal to N_v is first selected [22]. This subset of source data constitutes the source domain samples from where to learn the transportation plan. We have previously shown that with $N_v = 20$ is sufficient to perform the across sessions adaptation [11].

In order to find the training trials that best represent the calibration session, we sample multiple subsets of size N_v from the training data. For each subset, several backward transport plans are learned using combinations of regularization parameters, with the recalibration data serving as target samples. For a given subset, the combination of regularization parameters that achieves the highest accuracy on the transported recalibration data is selected as the optimal. Then, the subset that leads to the highest accuracy on the recalibration data is chosen as the source subset. Thus, both the source subset and the regularization parameters are determined in a data-driven manner based on their performance with respect to the recalibration data [11]. For further implementation detail, please visit our GitHub repository (https://github.com/vpeterson/otda-mibci).

Once the optimum source subset and regularization parameters are chosen, online adaptation can be performed. That is, for every ℓ^{th} trial in the application session, the backward optimal transport plan is computed by adding the current ℓ^{th} trial to the transportation set (see (4)), and solving the backward OT problem as defined in (2) with the source subset serving as source samples and the transportation set $\mathcal{V}_\ell$ as target samples.

Figure 1 illustrates the adaptive process performed by SBA in a two-dimensional CSP space. While the full training set was used to learn the CSP features as well as the LDA decision boundary (Figure 1a), only the source subset is used to learn the transportation plan, as illustrated in Figure 1b. Testing trials without adaptation lie on the feature space as defined by the given discriminative patterns (Figure 1c). When SBA is applied, testing trials are transported towards one or the other class center of mass (Figure 1d). The algorithmic support for a set of trials is shown in Figure 1e. It is timely to note that the distance of a given trial from its corresponding class center of mass in the feature space does not necessarily correlate with the effort exerted by the

model when performing the adaptation. Trials with features lying in close proximity to their corresponding class center of mass can get high algorithmic support values (e.g. trial 2).

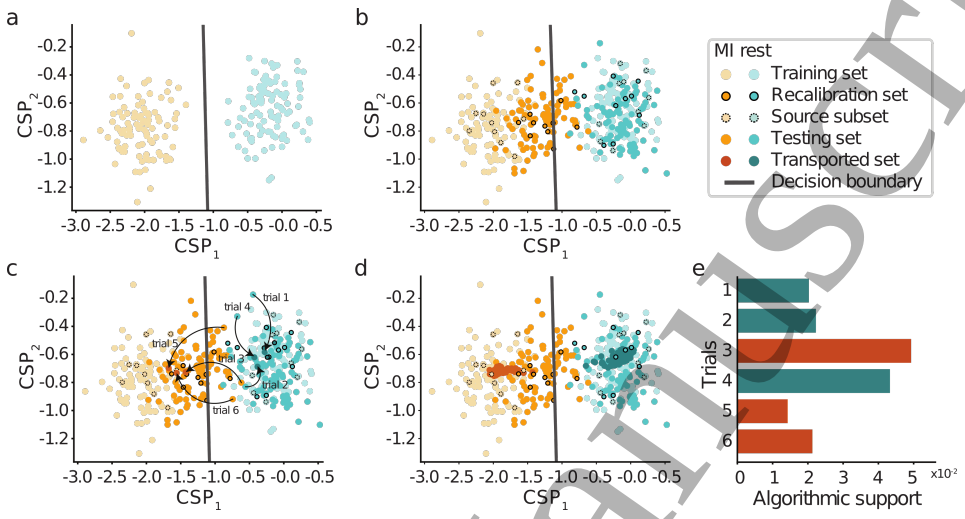

Figure 1: Schematic representation of SBA steps in the CSP feature space. **a.** Training data is used to learn the CSP+LDA decoding algorithm. **b.** By means of the recalibration data the optimal source subset is chosen from where to learn the backward transport plan. **c.** Example of few testing trials transported as indicated by the backward transport plan. Here arrows indicate the direction of the transportation. **d.** Comparison of the data with and without SBA application. **e.** Algorithmic support values for exemplary trials. Bar colors indicate the trial class. The higher the value, the stronger was the adaptive support performed by the algorithm.

4. Riemannian MI-BCI skills metrics

The Riemannian geometry has been widely adopted in the MI-BCI community to find a separable space in the covariance matrices manifold [23–25]. Given a raw EEG trial band-pass filtered in a frequency band of interest (e.g. 8 - 30 Hz), it is widely accepted that its spatial covariance matrix represents the band power of each channel in their first diagonal while the inter-channel covariance values are on the off-diagonal elements [24]. Based on the Riemannian distance between two covariance matrices, the distinctiveness metric D was proposed for measuring how apart patterns belonging to two different states A and B are [4]. That is,

$$D = \frac{\delta_R(\bar{C}_A, \bar{C}_B)}{\frac{1}{2}(\sigma_{C_A} + \sigma_{C_B})}, \tag{7}$$

where δ_R is the Riemannian distance, $\bar{C}$ is the Riemannian mean covariance matrix [24] and σ_C represents the mean absolute deviation of the covariance matrices, calculated as a measure of dispersion of the covariance matrices with respect to their mean. When

comparing two MI classes, D has been shown to be positively correlated with the user's capability to visualize objects and understands their spatial relationship (a.k.a spatial ability). Spatial ability is known to be a major predictor of MI-BCI performance [14, 26], and thus, it is expected that the higher the distinctiveness value is, the more distinguishable the user-generated EEG patterns are, informing about the MI-BCI user skills.

In the present study, D is used as a validated tool to assess users' BCI skills, as done by others [27–29]. Note that when used within a session between classes, it is known as *class distinctiveness* [4]. When the states A and B belong to the same class but across different sessions, D defines a between-session user MI-BCI skills metric, to which we shall refer to as *session distinctiveness*. Similarly, but when computed trial-wise, it can measure how dissimilar a given ℓ^{th} trial for a given k^{th} class in the application session is to the its corresponding training peers of the calibration session. That is, $D_\ell = \delta_R(\bar{C}_k^{tr}, C_{k,\ell}^{te})$, where $\bar{C}_k^{tr}$ denotes the Riemannian mean covariance matrix calculated on the training data for the class k, and $C_{k,\ell}^{te}$ denotes the covariance matrix of the ℓ^{th} trial belonging to class k to be decoded during the application session. We shall refer to D_ℓ computed trial-by-trial as *trial distinctiveness*. Note that, the higher the distinctiveness values are, the more dissimilar the compared states are.

5. Data

Simulated and real EEG data were used in this study, comprising two-classes MI-BCI datasets with at least two recordings sessions.

5.1. Simulated MI-EEG data

Different sessions of hand MI vs. rest EEG data were generated with PySimMIBCI [30]. Each session was simulated to define a possible MI-BCI scenario from the user's self-regulation capability perspective. While rest was simulated as idle EEG α-band activity in both hemispheres, for MI the simulation considered the well-known Event-Related Desynchronization (ERD) in the α-band at the contralateral hand motor area. Changes in the percentage of the ERD (ERD%) represents the relative change with respect to the idle state of the α-band during MI. The higher the ERD% is, the better the MI modulation capabilities are.

Each session comprised 200 trials of 4 seconds length with balanced class distribution. The EEG data was simulated with a sampling frequency of 1000 Hz from 41 electrodes localized according to the 10-5 electrode system. Data was then downsampled at 250 Hz and each trial was epoched from 0.5 to 2.5 s.

To study the extent at which the adaptation could be successful, we simulated data as coming from the same participant but in different possible scenarios with respect to both the self-regulation subject capability and the MI-BCI accuracy. While the former was simulated by changing the ERD%, the latter was simulated by including a

percentage of trials where the participant failed to perform the task indicated by the system (fail%). In the following, we shall refer as a simulated session with a given percentage of ERD and failed trials as $S_{ERD\%,fail\%}$. Simulated data sessions can be found in Zenodo (https://zenodo.org/records/13760210).

Simulating different self-regulation capabilities. We simulated sessions with different ERD%, starting from 50%, up to 10% in steps of 5%, leading to 9 MI-BCI sessions ($S_{50,0}$-$S_{10,0}$) from high to low α-band self-regulation capability. For each session, the ERD% was kept fixed and equal to the given value across trials. Figure 2a shows how the α peak decreases as the ERD% increases. Lower ERD% values have peak amplitudes closer to the rest condition, resulting in less distinguishable classes. The distinctiveness metric D was calculated intra-session between-classes and inter-sessions within-class. For the latter, session $S_{50,0}$ was chosen as the comparison reference. As shown in Figure 2b, intra-session distinctiveness decreases when the ERD% decreases, confirming that the lower the ERD% value is the harder the session results for the ML viewpoint. Similarly, conforming the ERD% decreases, the intra-session for MI class increases, showing how different lower ERD% values are from the strongest one (ERD% = 50.) The spatial information also reveals that low ERD% values present a less discriminable topography map, as shown in Figure 2c.

Simulating wrong performed trials. New sessions were created with varying numbers of failed trials (fail%). Specifically, we adjusted the percentage of failed trials from 10% to 50%, in step of 10% by swapping the labels between classes in a simulated session at 45% of ERD. This procedure resulted in five new sessions, named from $S_{45,10}$ to $S_{45,50}$.

Simulating MI related and non-related brain oscillations. We created a new data session of MI vs. rest of 100 trial each by assigning to the MI trials a uniform distribution on the ERD% from 10 to 50. In this way, simulated trials comprised different MI modulations levels within the same BCI session. In addition, following the approach used in [30], we simulated mental fatigue effects by increasing frontal θ and parietal α powers. Considering that mental fatigue typically manifest towards the end of the session, the first half of the simulated trials were fatigue-free, while the second half gradually incorporated increasing fatigue levels, by means of linearly augmenting the amplitude in frontal θ and parietal α power over trials. To show the fatigue effect impact besides MI-BCI skills user capability, the MI trials were simulated as having ERD% equal to 50 with no failed trials.

5.2. Real data

We used two different EEG-based datasets with two MI-BCI sessions. Motivated by the ultimate goal of developing a co-adaptive BCI for motor rehabilitation, the typical two-classes set-up was followed, that is both datasets comprised two MI classes.

Dataset-1. This dataset comprised 10 inexperienced able-bodied BCI users (3 females, 4 left-handed individuals, with a mean age of 25.45 years $\pm$ 2.50) previously

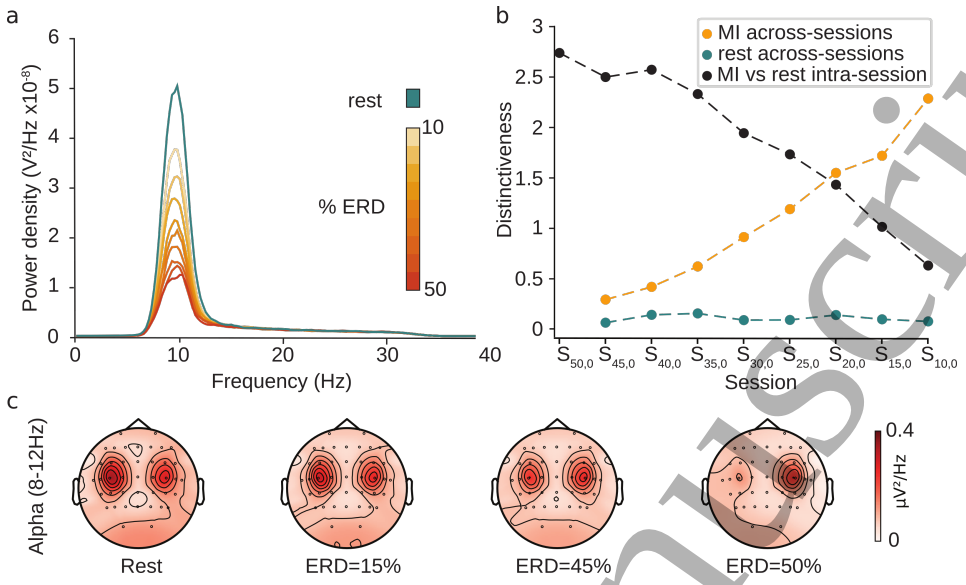

Figure 2: Simulated data main characteristics. **a.** MI self-regulation capabilities were simulated as different amplitude decrements of the alpha peak (10 Hz) with respect to baseline (rest). The higher the ERD% the higher the difference between rest and MI amplitudes. **b.** The distinctiveness metric is used to measure intra-session between classes and inter-session within class similarity. Session $S_{50,0}$ is used as reference for the session distinctiveness. Here dashed lines are used to connect adjacent dots, illustrating their tendency. **c.** Illustration of the alpha power in the channel-space for rest condition as well as for different ERD% values. The higher the ERD% value the more different the topography maps are from the rest condition.

acquired in [31]. The experiment included two identical sessions, spaced 5 days apart. Approval for the experiment was granted by the local ethics committee (BASEC-Nr. Req-2017-00631, Cantonal Ethics Commission, Zurich, Switzerland). Brain signals were recorded using a portable 64-channel EEG system (eego™rt Ant Neuro, Netherlands) with a sampling frequency of 512 Hz. Electrodes were positioned based on the international 10-20 system, with CPz as the reference and AFz as the ground electrodes. EEG signals were filtered within the frequency range of 0.5 Hz to 40 Hz. Participants performed two mental tasks: i) imagining the movement of their dominant hand (grasping motion, MI) and ii) a rest/relaxation condition (rest). Each session consisted of four runs with short breaks in between. Every run comprised 40 trials (20 for each task), resulting in a total of 160 trials per session. No feedback was provided to the subjects during the sessions. For this study, data was downsampled to 128 Hz and epoched from 0.5 to 2.5 s after the onset of the visual cue. 28 electrodes covering the sensorimotor areas were selected, following [32].

Dataset-2. Known in the literature as *Lee*2019_*MI*, is a large publicly available

dataset‡ obtained from 54 able-bodied participants in two sessions of MI-BCI [33]. The EEG recordings were acquired with the BrainAmp (Brain Products, Germany) using the nasion as reference and the electro AFz as ground. 62 electrode signals at 1000 Hz were collected as positionated by the 10-20 international montage system. In each session, participants performed a calibration and a test phase. Each phase had 100 trials with balanced right and left hand imagery tasks. During the online test phase, real-time classifier outputs were used to delivered feedback. No feedback was provided during the calibration phase. As before, the EEG signals were windowed from 0.5 to 2.5 s after the onset of the visual cue. Signals were band-pass filtered between 0.5 and 40 Hz, and then downsampled to 125 Hz. 18 channels covering the sensorimotor area were selected, as done in [33]. Given that for the online test phase no information about the indicated mental task to be performed was given, in the current work we used only those trials from the calibration phase. That is, we ended up with two MI-BCI sessions of 100 trials each.

6. Experiments and Results

Experiments were designed to give support and address the following main hypothesis:

H. "In supervised backward online adaptation, the associated cost of transporting a testing trial to match the calibration data distribution reflects the MI user self-regulation capability."

Performing a backward domain adaptation undoubtedly convey the need of relying on a discriminative source domain to where perform the adaptation. In addition, given that such backward adaptation is based on the cued label, the extend at which such information bias the adaptation performance must be studied to fully understand the nuances of the proposed algorithmic support metric. In this way, to prove the aforementioned hypothesis, we first needed to understand whether the following assumptions were held:

Supervised backward adaptation based on the cued label information can successfully address across-sessions data distribution changes if:

A1. the user-generated EEG patterns are discriminative and aligned to the indicated mental task, and

A2. the source data, where the classifier is trained, defines a discriminative space.

In the following, we first present experiments and results to prove $A1$ and $A2$ assumptions. Then, experiments and results that support the main hypothesis H are presented. Throughout these experiments we always used as decoding algorithm the standard CSP+LDA pipeline. Given the way simulated data was generated, only two CSP components were extracted, defining a 2-dimensional feature space. For real data, the typical six components of spatial filters were used [21]. For evaluating the decoding

‡ Available at http://gigadb.org/dataset/100542

performance, classification accuracy was used. In both simulated and real data, EEG signals were filtered in between 8 and 30 Hz before feature learning. To quantify the user MI-BCI skills, D was used between classes (class distinctiveness), between sessions per each class (session distinctiveness) and trial-wise (trial distinctiveness). For reference, the standard calibration (SC) framework, where not adaption is applied between sessions, is also implemented. The simulations as well as the corresponding source codes implemented in this work are publicly available at GitHub (https://github.com/NiCALab-IMAL/SupportiveBackwardAdaptation).

6.1. Dependence of adaptation success on the user-generated EEG patterns

To evaluate the impact of the user-generated EEG patterns and their correspondence to the indicated mental task (assumption $A1$), two experiments were conducted here. In both experiments, to focus only on the data discriminability at the target domain level, the calibration session (source domain) was simulated as the most well-suited scenario. Thus, the decoding algorithm was trained on simulated data $S_{50,0}$, a session with clear discriminability between MI and rest conditions and no failed trials (see Figure 2). In the subsequent applications, both trained CSP and LDA models were kept fixed, and the recalibration set was defined by the first 20 trials of the testing session ($N_v = 20$).

For the first experiment, we evaluated the decoding performance of the already trained classifier using as testing data sessions from $S_{45,0}$ to $S_{10,0}$. In this way, all EEG patterns belong to the indicated mental task, but with different simulated MI decoding capability. The regularization parameters λ and ν in (2) were kept fixed to 0.1 and 1, respectively. These values were selected to reinforce with positive bias with respect to the indicated mental task information (the higher ν, stronger is the impact of accounting with the cued label to learn the backward transport plan). Results are shown in Figure 3a, demonstrating that the performance of the already trained classifier when SBA is used achieved perfect or near perfect classification values across all simulated sessions. The classification performance of the standard pipeline (SC) is shown for reference. To illustrate in the CSP space how the transportation works, Figure 3b shows for a given session ($S_{15,0}$) the training data, the original testing data as well as the resulting transported testing data. Notably, for most trials, SBA effectively compensates for data distribution shifts between testing and training sets, resulting in high classification accuracy. In the context of the present study, we shall refer to as *adaptation success* those cases at which testing trials are backwardly transported by means of SBA to the correct side of the decision boundary, and thus are correctly classified.

For the second experiment, we evaluated the performance of the perfect trained decoding model in testing sessions with high ERD% but increasing number of failed trials, namely $S_{45,0}$ and from $S_{45,10}$ to $S_{45,50}$. As before, to reinforce positive bias with respect to the system information, the regularization parameter of the group-LASSO penalty was kept fixed and equal to 1 across testing trials. The Sinkhorn regularization parameter was also kept to 0.1. Accuracy results are presented in Figure 3c, where is

it clear that SBA achieved comparable classification performance than SC, i.e, a model without adaptation. Figure 3d illustrates for a given testing session ($S_{45,30}$) that failed trials were transported not based on the indicated mental task class but on the feature information. In other words, a trial with EEG patterns resembling MI but "labeled" as rest, was transported to their MI trials peers, regardless of having a rest label, and vice-versa. This result indicates that accounting for the cued label will effectively aid in backward adaptation only when the user-generated EEG patterns align with the indicated mental task. Understanding that SBA successfully achieves adaptation when the EEG patterns correspond to the indicated mental task, assures us that the predicted class of the model will reflect the user's mental intention.

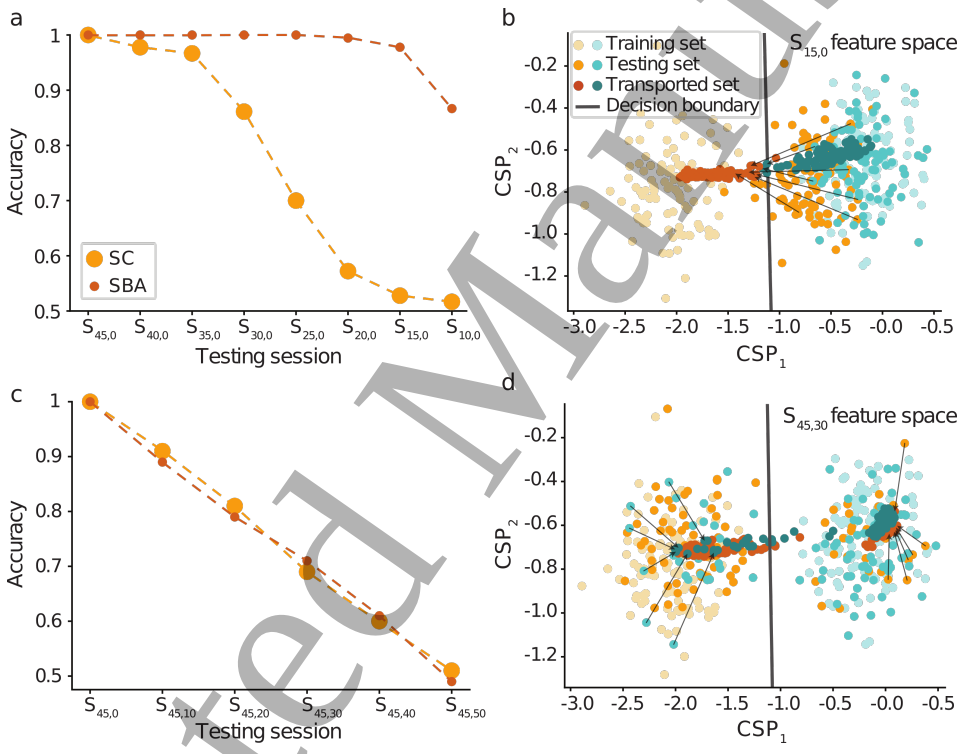

Figure 3: SBA adaptive success dependence on the ERD% and percentage of failed trials. Decoding model trained in a perfect session with high ERD% and no fail trials ($S_{50,0}$). **a.** Accuracy of SC and SBA across different testing sessions with varied ERD% but no failed trials. Dashes lines drawn to connect adjacent points. **b.** Illustration on SBA transportation for a given testing session with no failed trails but low ERD% ($S_{15,0}$). **c.** Accuracy of SC and SBA across different testing sessions with varied fail% but high ERD%. Performance across simulated session are connected with dashes lines to improve results readability. **d.** Illustration of SBA transportation for a given testing session with 30% of failed trails but ERD% equal to 45 ($S_{45,30}$).

.

Altogether, these results evidence the fact that supervised backward adaptation based on optimal transport can address the data distribution changes between the target (testing) and the source (calibration) set when the EEG patterns belongs to

the indicated mental task.

6.2. The effect of the training data discriminability on adaptation

While in the previous experiments we focused on studying the testing sessions (target domain) impact in the backward adaptation to a perfect training session (source domain), here we want to evaluate source domains changes and its impact in the adaptation success. In other words, we want to prove assumption $A2$. The decoding algorithm obtained by different training sessions was tested with session $S_{50,0}$, that is, the perfect session at which all trials are aligned to their corresponding class and have ERD% value equal to 50.

First we evaluated the impact in the self-regulation capability by using as training sessions $S_{45,0}$ to $S_{10,0}$. Class distinctiveness was calculated in each training session to account for a data-driven MI-BCI user's skills metric. As shown in Figure 4a, the class distinctiveness decreased as the ERD% decreased, but always yielding values highly above zero. Interestingly, training accuracy remained fairly high (above 80%). Note that the accuracy of SBA approached near-perfect classification across all trained models.

On the other hand, we used as training data sessions $S_{45,0}$ to $S_{45,50}$. Here, the class distinctiveness rapidly decreased from session to session, yielding values closer to zero when the number of failed trails was equal or above 40% (Figure 4c). Similarly, training accuracy tended to by-chance levels as the fail% increased. It can be observed that SBA failed to transport samples when the source domain class distinctiveness was too low. Therefore, such sessions seemed not to provide a reliable calibration basis for backward adaptation. Figure 4b and d show the feature space with their corresponding LDA decision boundary as well as the CSP topomaps when trained in $S_{10,0}$ and in $S_{45,40}$, respectively. While for $S_{10,0}$ the model was still able to learn discriminable patterns, that was not the case for $S_{45,40}$. Given that both, CSP and class distinctiveness rely on covariance matrices estimation, it is somewhat expected that EEG patterns with near zero distinctiveness result in a low discriminable space. Although this is an extreme scenario, low class distinctiveness can occur when the subject is not able to self-modulate their brain activity, and thus most of the MI trials overlap with the rest ones.

These results demonstrate that backward adaptation should be conducted using a training session with reasonable class discriminability. While this is obvious for any decoding algorithm, extra care must be taken when working in backward adaptation, as it strongly relies on train data distribution. These results indicate that despite of having perfectly performed trials during the application session, adaptation will fail when the training session lacks sufficient class discriminability. In real-life scenarios, this could lead to ambiguity and misleading feedback. To avoid user frustration and to properly support their learning process, the supervised backward adaptation must only be implemented after calibration data defines a discriminable feature space.

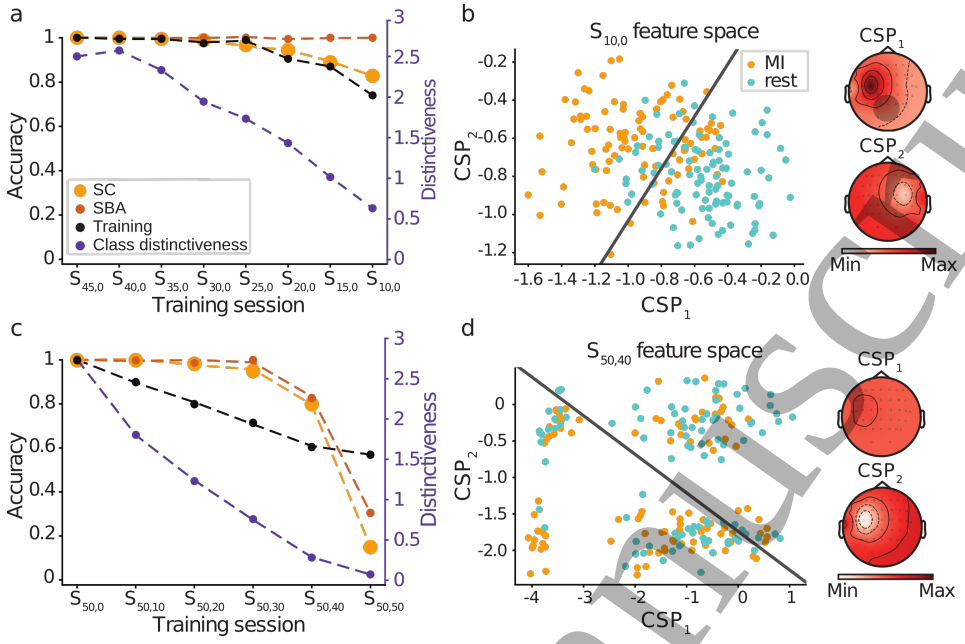

Figure 4: SBA adaptive success dependence on the training data discriminability. Testing data kept fixed to a perfect session with high ERD% and no failed trials ($S_{50,0}$). **a.** Accuracy reached by SC and SBA across different training sessions with varied ERD% but no failed trial. Training model accuracy as well as class distinctiveness is shown. Connecting lines are drawn to improve results readability. **b.** Illustration on the training feature space with the learned LDA decision boundary together with the CSP topomap for a session with low ERD% and no failed trials ($S_{10,0}$). **c.** Accuracy reached by SC and SBA across different training sessions with varied fail% but high ERD%. Training model accuracy as well as class distinctiveness is shown. Here again, connecting lines are drawn to improve results readability. **d.** Illustration on the training feature space with the learned LDA decision boundary together with the CSP topomaps for a session with high ERD% but 40% of failed trials ($S_{45,40}$).

6.3. Relationship between algorithmic support and data discriminability

In the previous experiments we proved that if the calibration data is discriminative enough and the EEG patterns belong to the indicated mental task, SBA succeeds in performing the adaptation. But, while the adaptation seems to be possible regardless the ERD%, what is the *cost* of doing so? In other words, how much is the algorithmic support (h) made by SBA to facilitate the proper data alignment? In this section we investigate the relationship between the algorithmic support and the MI-BCI user skills with simulated data.

As before, CSP and LDA were used for EEG decoding. SBA regularization parameters were kept fixed in 0.1 and 1, for λ and ν, respectively. To ensure a perfect calibration session, session $S_{50,0}$ was used as training data. The algorithmic support h was measured at every testing trial as defined in (6). To understand whether a

correlation exists between the algorithmic support and the simulated self-regulation capability, we used as testing data the session with ERD% uniformly sampled from 10 to 50, as described in Section 5.1. As shown in Figure 5a, there is a strong negative correlation in between the algorithmic support and the ERD% (Pearson correlation $r = -0.8$, p-value < 0.0001), showing that while adaptation was successfully possible (SC accuracy = 0.86, SBA accuracy = 0.98), the lower the ERD%, the higher the cost h. Interestingly, the algorithmic support also increased as long as the simulated fatigue effect increased, as shown in Figure 5b. Here it is observed that for the first half of the trials, where no fatigue was present, the algorithmic support was zero or near zero, while for the second half, the algorithmic support increased linearly, following the fatigue level increments as explained in Section 5.1.

While these two experiments were run based on a fixed but known ν parameter to accomplish with strong group-LASSO penalty in (2), in real-life scenarios the regularization parameters might be selected based on data-driven procedures. As explained in Section 3, the regularization parameters selection relies on the recalibration data and the source subset from where to learn the transportation plan [11]. As before, recalibration data conforms the first 20 trials of the testing session ($N_v = 20$). Thus, in the following experiment we let the model decide the best regularization parameters from a list of 20 values, ranging from 0.1 to 10. Similar to before, training data was $S_{50,0}$, but testing data varied from $S_{45,0}$ to $S_{10,0}$. h was calculated for each testing trial in each testing session. The results, depicted in Figure 5c show the mean algorithmic support value together with their standard deviation across all testing trial in each testing session. In line with the previous results, the algorithmic support increased as the ERD% in testing data decreased. Interestingly, the selected group-LASSO regularization parameter (ν) started to increase from $S_{30,0}$, that is, when the ERD% was lower than 30%. Note that the Sinkhorn regularization parameter (λ) was always chosen to be 0.1.

6.4. Co-adaptive learning in real data

In order to understand the potential of SBA as a co-adaptive method, real datasets in a subject-specific cross-session MI-BCI scenario were used. For each dataset, the first session was used as the calibration data, from where the decoding algorithm was trained. CSP+LDA was used by setting the number of spatial filters in CSP equal to six, as typically defined in the field [21]. The second session was used as the application session, from where the already trained decoding algorithm was used with SBA applied at each testing trial. To mimic real-life implementation, the regularization parameters were selected for each subject-data by cross-validation, using their corresponding recalibration data. That is, the first 20 trials of the new session were used to find both the source subset as well as the best regularization parameters.

During the application session, the trial distinctiveness as well as the the algorithmic support h was calculated for each testing trial. The session distinctiveness per each class was computed across the training and testing data per each subject in each database.

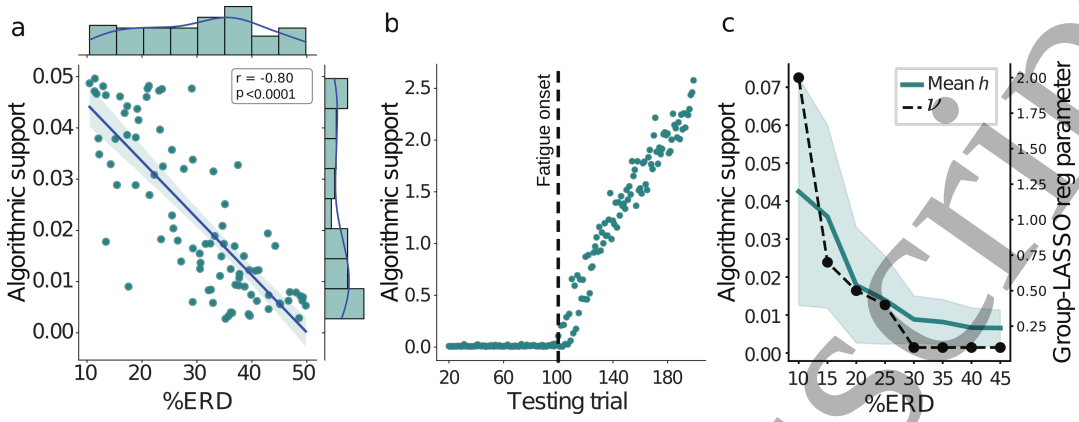

Figure 5: SBA algorithmic support informs about MI-EEG discriminability. **a.** Algorithmic support h as a function of the ERD% in a testing session with uniform ERD% distribution. Text aggregates Pearson correlation analysis between the algorithmic support h and the ERD%. **b.** Algorithmic support h in a testing session with linearly increasing fatigue levels in the second half of the trials. The dashed vertical line indicate the fatigue onset. Each dot represent a trial. **c.** Algorithmic support h at different testing sessions with different ERD% when the regularization parameters are selected by cross-validation. The thick continuous line shows the mean h value across trials, while the shaded area shows standard deviation. The value of the group-LASSO regularization parameter at each tested session is also shown. A connecting line is drawn between points to illustrate the tendency.

Accuracy was used to assess decoding performance for the classifier with and without adaptation, namely SC and SBA, respectively.

Figure 6a shows the algorithmic support h, trial distinctiveness D_ℓ and accuracy of each subject in Dataset-1 sorted from higher to lower median h value. In general, the bigger the algorithmic support, the higher the trial distinctiveness; indicating that the model needed more effort in the adaptation as more dissimilar the trial to their corresponding class training data was. But, while decoding accuracy is important to ensure a proper BCI control, that does not always reflect the MI user intention. In this way, it is interesting to observe the accuracy gained when performing the adaptation, shown in top panel of Figure 6a. As can be seen, a large classification performance improvement did not always imply a huge algorithmic support. Based on the results presented in simulated data, this phenomena could indicate that the subject was performing the indicated mental task correctly but with weak modulation (low ERD%). Similarly, cases at which strong algorithmic support was measured but with null or small relative classification performance increments between SC and SBA, can be explained by failed trials, that is, generated EEG patterns by the user that did not correspond to the indicated mental task by the system.

A similar analysis can be done for the results presented in Figure 6b for Dataset-2. In this case, due to the large number of subjects comprising this database, the

median of h and D_ℓ value across trials is depicted per subject. Subjects with higher algorithmic support are highlighted with darker colors. Figure 6c aggregates accuracy information with and without data adaptation, showing, here again, that there is no clear correspondence in between algorithmic support and classification improvement.

Considering real-life applications, at which feedback should be delivered as soon as possible, we measured the computational time required to adapt and compute the logarithmic support metric of a given trial by means of SBA. Experiments were run on a standard PC with Intel® Core^{TM} i7-6700K CPU. For both datasets, the average SBA computing time was of about 40 ms, around 2.5 times faster than the average time needed to compute the Riemannian trial distinctiveness metric.

Given that experiments were conducted in a cross-sessions scenario, it was important to understand whether correlation existed in between data similarity across classes and sessions, and the algorithmic support measured by means of h. Figure 6d and 6e show for Dataset-1 and Dataset-2, respectively, the median value across trials of h per each subject versus the session distinctiveness value found per each class. It is notable how positive and significantly correlated these two metrics were. This supports the idea that h can be used to measure the MI-BCI user skills.

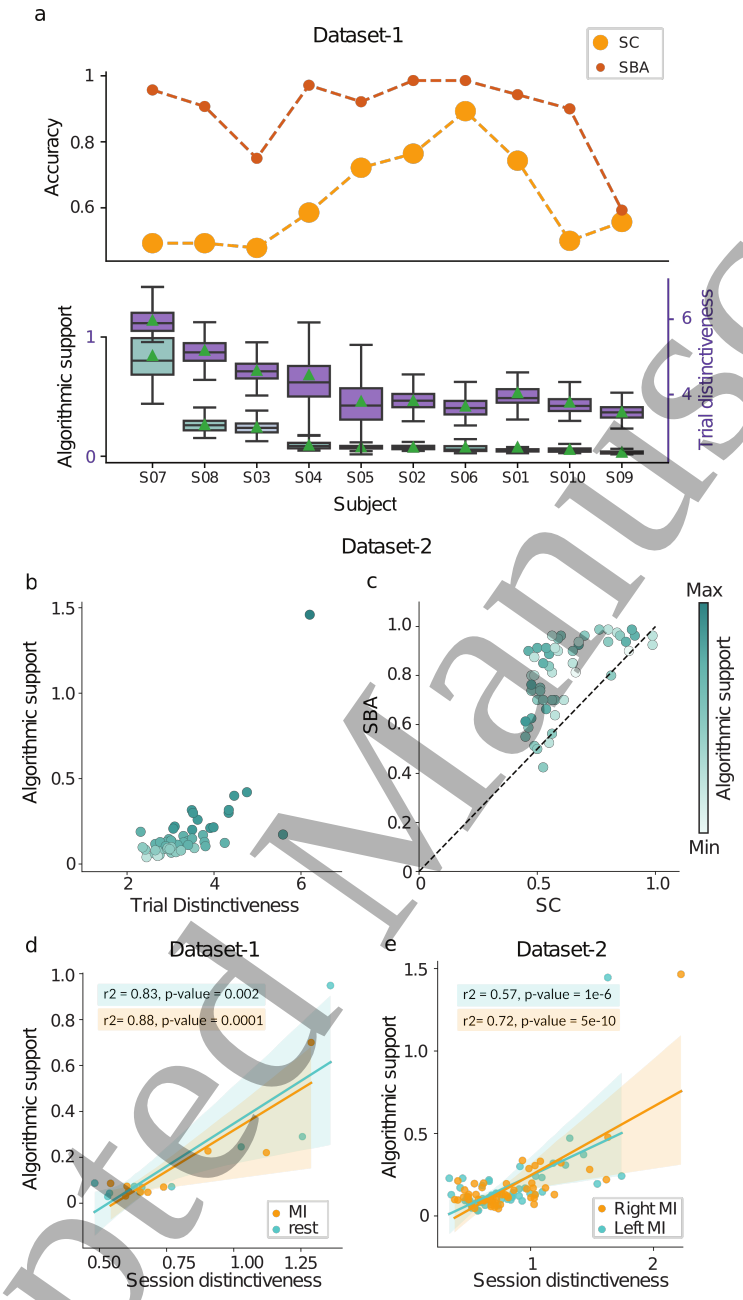

Figure 6: SBA algorithmic support is strongly correlated to Riemannian-based metrics. **a.** Algorithmic support, trial distinctiveness, and accuracy for each subject of Dataset-1 sorted in descending median h value. Lines in the accuracy plots are drawn to help results readability. Boxplots on the button reflect across trial computed metrics, with the left and right y-axes showing range values for the algorithmic support and the trial distinctiveness, respectively. **b.** Algorithmic support as a function of trial distinctiveness, where each dot represent the median value for each subject in Dataset-2. Subjects were colored-coded accordingly to their median h value, where darker colors correspond to higher algorithmic support values. **c.** Accuracy reached by SC and SBA per each subject in Dataset-2. Dots' color follows the increasing order of the corresponding h value per each subject. **d-e.** Algorithmic support as a function of session distinctiveness per each class for Dataset-1 and Dataset-2, respectively. Text aggregates Pearson correlation coefficient found and its p-value per class.

7. Discussion

Motor imagery-based BCIs are promising technologies for neurorehabilitation [34, 35]. While around 25% of the users cannot command MI-BCIs [36], this may not necessarily indicate that they are not working towards being able to self-regulate their brain activity. Nevertheless, generating stable and distinguishable EEG patterns takes time. The session-to-session learning process must be adequately supported by the system, adapting to data changes. In the present work, we propose an index to assess the discriminability and stability of the user generated patterns across-sessions of BCI use. This measure, referred to as *algorithmic support*, is based on the cost of adapting the EEG trial to its corresponding training peers within a supportive backward adaptation (SBA) process. We showed that the algorithmic support h can be used for online assessing the discriminability and stability of EEG patterns, and thus, SBA has the potential to become a valuable tool for building co-adaptive MI-BCIs.

The algorithmic support metric measures the effort made by SBA when performing backward optimal transportation in a supervised manner. We studied SBA as a co-adaptive method for MI-BCI. SBA is based on BOTDA [11], an adaptive method designed to match the data distribution of a new session with the data distribution of a calibration session, avoiding classifier retraining. Considering MI-BCI for motor rehabilitation, the domain adaptation is done in a trial-basis using the indicated mental task provided by the system to guide the adaptation. In this work we showed that the backward adaptation based on the system information (cued label) is successful when the user-generated EEG patterns reflect the indicated mental task. That is, data alignment between sessions leads to high accuracy only when the user intention matches the cued label information. Most importantly, we showed that the effort the model exerts in performing such adaptation informs how distinguishable and stable the EEG patterns are.

Experiments on simulated data revealed the correspondence between patterns discriminability and the system information when performing SBA. Throughout extensive simulations, we proved that the adaptation is able to match the data distribution of the new session to the old one only when the EEG patterns belong to the indicated mental task, regardless of the level of ERD%, as shown in Figure 3. These results indicate that the system information will only guide the adaptation to be successful when the generated EEG patterns by the user match the cued mental task. In other words, when the user self-modulate their brain activity in a discriminative and task-related manner. Results on simulated data showed that the backward adaptation was possible also in extreme scenarios at which the simulated ERD% values were as low as 10% (see Figure 2b). While the accuracy of the adaptive model reached perfect classification, we found that the algorithmic support increases as the ERD% decreases (Figure 5c).

A relevant aspect to analyze when performing a backward adaptation is the quality, from the discriminative viewpoint, of the calibration data. Our results showed that

data used to calibrate the model (source domain), should reflect data discriminabilty, as measured by the class distinctiveness and the accuracy in training data (see Figure 4). In this way, only "good" calibration sessions are reliable to act as source domain, that is, domains to where perform the backward data adaptation. This finding indicates that for multi-session MI-BCIs, the user might need to perform a good calibration session, as measured by both the training accuracy and class distinctiveness before passing to the application session. In real-life implementations, this could help avoiding user frustration during sessions with feedback. Considering most naive BCI users fail to perform MI in the very first session, different strategies could be asked to be done in different session days until reaching good calibration performance. Only then, user could pass to next stages of the MI-BCI controlling experience.

Frustration with respect to BCI controlling capabilities development can also be diminished if the feedback provided to the user is not solely based on the predicted class by the decoding algorithm. It has been argued that with a feedback that informs the user about the quality of the generated EEG patterns with respect to the indicated mental task, the participant might better modulate their brain activity and thus, improve their BCI controlling capabilities [4, 12]. Here, we showed that the effort put by SBA when performing the adaptation informs about the goodness on the generated EEG patterns. Results through simulations showed that the algorithmic support in performing the adaptation increases as the data discriminative diminishes, as simulated by low ERD% values in the data (see Figure 5a). Data discriminability can also be affected by other brain process that could co-exist during a MI-BCI session. This is the case of mental fatigue, that could appear during MI-BCI sessions perpetuating the underlying task-related brain oscillations [37]. We showed that, as long as MI-related activity is preserved, data adaptation can be performed but with higher algorithmic support as the fatigue level increases (see Figure 5b). Although it is hard to account for online fatigue biomarkers, this result sheds light on the necessity to monitor fatigue levels within a MI-BCI session.

As claimed by the authors in [4], metrics based on the Riemannian geometry can help identifying user's MI ability beyond the classifier output. This involves measuring the stability and discriminability of the EEG patterns using a parameter-free and machine-independent algorithm. Such an approach can help delineate the decoding roles between the classifier and the user more effectively; however, it strongly relies on the estimation of covariance matrices, which are prone to issues in low electrode counting scenarios. In the current study we propose a novel way of assessing the user MI-BCI skills within an adaptive algorithm. It works at the feature space level, as extracted by any machine learning model or feature engineering process. Real data results show that the algorithmic support, as measured by h, is positive and significantly correlated to well-established metrics based on the Riemannian geometry. Given that class distinctiveness was proven to be positive correlated to mental rotation capabilities [4, 14], the found correlation across a large number of subjects' data suggest that h has the potential to become a valuable tool to measure, trial-by-trial, the effort made by the system to

perform the adaptation, and thus explicitly indicate to the user how well the indicated mental task was performed. In this way, it is envisioned that by means of SBA not only across-sessions adaptation is performed but also meaningful feedback be delivered to the user, and thus support their MI-BCI learning process.

We envision SBA as powerful tool for users mastering their BCI control abilities. We aim at fostering independence and the sense of agency during MI-BCI sessions. Nevertheless, it is important to recognize a trade-off: while providing support may ease the user's experience, it can also lead to a lack of true learning and diminished engagement in actively controlling the MI-BCI. Conversely, if users navigate their sessions without adequate guidance, they may struggle to develop the necessary skills and self-awareness required for effectively command the system. With this in mind, and following what is commonly done in the literature [38], in multiple-sessions MI-BCIs, positive feedback can first be imposed. In the context of SBA, such positive bias can be performed by manually setting the group-LASSO regularization parameter to a high value. Then, as sessions progress, if the average algorithmic support decreases from session to session and offline quantifiers of ERD activity show better modulation capabilities, the regularization parameter can be adjusted at the beginning of the online session, managing the frustration-motivation ratio. In this line, our results in simulated data showed that there is a positive relationship between the selected regularization parameter and the algorithmic support, where both increase as long as the sessions become more and more difficult from the class-discriminative viewpoint (see Figure 5c). This last finding may indicate that the data-driven selected regularization parameter has the potential to inform about the cross-session adaptation complexity, where harder adaptive scenarios are more likely to have higher values of the group-LASSO regularization parameter. In closed-loop implementations, the data-driven selected regularization parameter can directly be used as users become more experienced, and thus no positive bias is needed to be imposed by the experimenter.

While the predicted class by the decoding algorithm affects the control efficiency of the BCI, it does not necessary resemble the generated pattern by the MI-BCI user [4]. In this way, the algorithmic support can provide another domain of quantification for a given trial, beyond predicted class category. As seen in Figure 6, improvements in accuracy after applying data adaptation not always indicate a huge algorithmic support, and vice-versa. Delivering to the user feedback based on not only the predicted class but also on the algorithmic support is expected to lead to more transparency on how the model decoding works, and hopefully guide and support the user throughout the MI-BCI learning process. In this way, visual neurofeedback in a form of an energy bar can be delivered in the closed-loop stimulation protocol. Based on the algorithmic support such energy bar can be colored and updated trial by trial. It is expected that users realize that the better the performed mental task, the lower algorithmic support exerted by the model is; and thus, use that feedback to improve, change or maintain their modulation strategy within a session. The implementation of such a closed-loop stimuli is part of our future plans.

In the context of across-session adaptation, the proposed algorithmic support can also be influenced by non-physiological factors affecting the stability of the EEG. Currently, there is no form to disentangle how much the changes in data are due to modulation variability or non-BCI user-related factors, such as electrode-cap misalignment across the BCI sessions. Neither the algorithmic support nor the Riemannian distinctiveness metric calculated across sessions can inform whether the observed changes in a trial are related to variability due to MI-BCI user skills or any other type of source. Nevertheless, while we acknowledge that more research is needed in this direction, we argue that within a session, the algorithmic support value can only change due to user self-modulation, as other sources of cross-session variability remain constant.

The present study is based on two-sessions without feedback MI-BCI datasets. Undoubtedly, further research is needed to understand how to apply co-adaptation as more sessions become available. In addition, while SBA can be applied at any feature domain representation, the present study focused on the traditional CSP features, which rely on covariance matrices estimations, as Riemannian metrics do. Therefore, it remains to be elucidated whether the algorithmic support metric would exhibit similar behavior when applied in other feature spaces. Future plans involve studying the effects of applying SBA in more complex latent spaces, such as those encoded by deep learning architectures. Similarly, it is necessary to study the effect of the delivered feedback in both data adaptation and algorithmic support assessment. In this direction, our future research plans aim at studying the aforementioned study limitations as well as validating SBA as a co-adaptive method in multiple-sessions of MI-BCI with applications to motor rehabilitation. Next steps involve the use of SBA as an online adaptive algorithm as well as to provide meaningful feedback to the user in a trial-basis. We hope this induces a better user-machine experience, allowing more users to control a MI-BCI, and thus, ultimately potentiating the use of MI-BCI for motor rehabilitation.

8. Conclusion

To support the user's learning process during MI-BCI training, we introduced SBA, a method based on backward optimal transport for domain adaptation, which addresses across-sessions variability while assessing the quality of the user-generated EEG patterns. This study demonstrates the potential of SBA as a co-adaptive approach for MI-BCI systems, highlighting its envisioned capacity to enhance user experience during motor rehabilitation. By enabling online adaptation based on the data distribution of new sessions, SBA can minimize the need for classifier retraining while providing critical feedback on the user's ability to perform the indicated mental tasks. Our findings reveal that successful adaptation strongly depends on the discriminability and stability of the user-generated EEG patterns in a task-related manner, emphasizing the importance of high-quality calibration data. Furthermore, we established a quantitative measure of adaptation effort through the algorithmic support h index, which offers

insights into user performance. This dual functionality of SBA not only facilitates seamless session-to-session adaptation but also has the potential to provide users with meaningful feedback to enhance their MI-BCI control capabilities. As we advance the development of MI-BCI systems, integrating such adaptive methodologies is essential for fostering user confidence and engagement, ultimately improving clinical outcomes in motor rehabilitation.

Data availability statement

The simulated data created for this study is publicly available at Zenodo.

Acknowledgments

This work was partially supported by CONICET, through PIP 2021-2023 11220200100806CO and Universidad Nacional del Litoral, through CAI+D2020-50620190100069LI.

References

[1] C. Sannelli, C. Vidaurre, K.-R. Müller, and B. Blankertz, "A large scale screening study with a SMR-based BCI: Categorization of BCI users and differences in their SMR activity," *PloS one*, vol. 14, no. 1, p. e0207351, 2019.

[2] A. Roc, L. Pillette, J. Mladenovic, C. Benaroch, B. N'Kaoua, C. Jeunet, and F. Lotte, "A review of user training methods in brain computer interfaces based on mental tasks," 2 2021.

[3] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller, "Imagery of motor actions: Differential effects of kinesthetic and visual–motor mode of imagery in single-trial EEG," *Cognitive brain research*, vol. 25, no. 3, pp. 668–677, 2005.

[4] F. Lotte and C. Jeunet, "Defining and quantifying users' mental imagery-based BCI skills: a first step," *Journal of neural engineering*, vol. 15, no. 4, p. 046030, 2018.

[5] J. S. Müller, C. Vidaurre, M. Schreuder, F. C. Meinecke, P. Von Bünau, and K.-R. Müller, "A mathematical model for the two-learners problem," *Journal of neural engineering*, vol. 14, no. 3, p. 036005, 2017.

[6] G. Huang, Z. Zhao, S. Zhang, Z. Hu, J. Fan, M. Fu, J. Chen, Y. Xiao, J. Wang, and G. Dan, "Discrepancy between inter-and intra-subject variability in EEG-based motor imagery brain-computer interface: Evidence from multiple perspectives," *Frontiers in neuroscience*, vol. 17, p. 1122661, 2023.

[7] N. Robinson, T. Chouhan, E. Mihelj, P. Kratka, F. Debraine, N. Wenderoth, C. Guan, and R. Lehner, "Design considerations for long term non-invasive brain computer interface training with tetraplegic cybathlon pilot," *Frontiers in Human Neuroscience*, vol. 15, 2021.

[8] W. Samek, M. Kawanabe, and K.-R. Müller, "Divergence-based framework for common spatial patterns algorithms," *IEEE Reviews in Biomedical Engineering*, vol. 7, pp. 50–72, 2013.

[9] G. Ditzler, M. Roveri, C. Alippi, and R. Polikar, "Learning in nonstationary environments: A survey," *IEEE Computational Intelligence Magazine*, vol. 10, no. 4, pp. 12–25, 2015.

[10] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[11] V. Peterson, N. Nieto, D. Wyser, O. Lambercy, R. Gassert, D. H. Milone, and R. D. Spies, "Transfer learning based on optimal transport for motor imagery brain-computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 2, pp. 807–817, 2021.

[12] F. Lotte, F. Larrue, and C. Mühl, "Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design," *Frontiers in human neuroscience*, vol. 7, p. 568, 2013.

[13] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K.-R. Müller, G. Curio, and T. Dickhaus, "Neurophysiological predictor of SMR-based BCI performance," *Neuroimage*, vol. 51, no. 4, pp. 1303–1309, 2010.

[14] C. Jeunet, B. N'Kaoua, S. Subramanian, M. Hachet, and F. Lotte, "Predicting mental imagery-based BCI performance from personality, cognitive profile and neurophysiological patterns," *PloS one*, vol. 10, no. 12, p. e0143962, 2015.

[15] E. M. Hammer, S. Halder, B. Blankertz, C. Sannelli, T. Dickhaus, S. Kleih, K.-R. Müller, and A. Kübler, "Psychological predictors of SMR-BCI performance," *Biological psychology*, vol. 89, no. 1, pp. 80–86, 2012.

[16] V. G. von Groll, N. Leeuwis, S. Rimbert, A. Roc, L. Pillette, F. Lotte, and M. Alimardani, "Large scale investigation of the effect of gender on mu rhythm suppression in motor imagery brain-computer interfaces," *Brain-Computer Interfaces*, pp. 1–11, 2024.

[17] L. J. Hagedorn, N. Leeuwis, and M. Alimardani, "Prediction of inefficient BCI users based on cognitive skills and personality traits," in *International Conference on Neural Information Processing*, pp. 81–89, Springer, 2021.

[18] S. Perdikis and J. d. R. Millan, "Brain-machine interfaces: a tale of two learners," *IEEE Systems, Man, and Cybernetics Magazine*, vol. 6, no. 3, pp. 12–19, 2020.

[19] N. Courty, R. Flamary, D. Tuia, and A. Rakotomamonjy, "Optimal transport for domain adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 9, pp. 1853–1865, 2017.

[20] L. V. Kantorovich, "On the translocation of masses," *Dokl. Akad. Nauk SSSR, 37*, vol. 37, no. 7-8, pp. 227–229, 1942.

[21] F. Lotte, "Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain–computer interfaces," *Proceedings of the IEEE*, vol. 103, no. 6, pp. 871–890, 2015.

[22] N. Courty, R. Flamary, and D. Tuia, "Domain adaptation with regularized optimal transport," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 274–289, Springer, 2014.

[23] F. Yger, M. Berar, and F. Lotte, "Riemannian approaches in brain-computer interfaces: a review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1753–1762, 2016.

[24] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, A. Barachant, S. Bonnet, M. Congedo, C. Jutten, and E. Eds, "Riemannian geometry applied to BCI classification," *Lva/Ica 2010*, pp. 629–636, 2010.

[25] T. Qu, J. Jin, R. Xu, X. Wang, and A. Cichocki, "Riemannian distance based channel selection and feature extraction combining discriminative time-frequency bands and riemannian tangent space for MI-BCIs," *Journal of Neural Engineering*, vol. 19, no. 5, p. 056025, 2022.

[26] C. Jeunet, B. N'Kaoua, and F. Lotte, "Advances in user-training for mental-imagery-based BCI control: Psychological and cognitive factors and their neural correlates," *Progress in brain research*, vol. 228, pp. 3–35, 2016.

[27] S. Kumar, H. Alawieh, F. S. Racz, R. Fakhreddine, and J. d. R. Millán, "Transfer learning promotes acquisition of individual BCI skills," *PNAS nexus*, vol. 3, no. 2, p. pgae076, 2024.

[28] M. S. Yamamoto, F. Lotte, F. Yger, and S. Chevallier, "Class-distinctiveness-based frequency band selection on the riemannian manifold for oscillatory activity-based BCIs: preliminary results," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3690–3693, IEEE, 2022.

[29] P. V. Ascenção, E. M. Santos, L. H. Lacerda, and F. J. Fraga, "Evaluation of performance metrics for users of brain computer interfaces during motor imagery," in *2019 IEEE International*

Conference on Systems, Man and Cybernetics (SMC), pp. 217–222, IEEE, 2019.

[30] C. M. Galván, R. D. Spies, D. H. Milone, and V. Peterson, "Neurophysiologically meaningful motor imagery EEG simulation with applications to data augmentation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024.

[31] V. Peterson, D. Wyser, O. Lambercy, R. Spies, and R. Gassert, "A penalized time-frequency band feature selection and classification procedure for improved motor intention decoding in multichannel EEG," *Journal of Neural Engineering*, vol. 16, no. 1, p. 016019, 2019.

[32] S. Marchesotti, R. Martuzzi, A. Schurger, M. L. Blefari, J. R. del Millán, H. Bleuler, and O. Blanke, "Cortical and subcortical mechanisms of brain-machine interfaces," *Human brain mapping*, vol. 38, no. 6, pp. 2971–2989, 2017.

[33] M.-H. Lee, O.-Y. Kwon, Y.-J. Kim, H.-K. Kim, Y.-E. Lee, J. Williamson, S. Fazli, and S.-W. Lee, "EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy," *GigaScience*, vol. 8, no. 5, p. giz002, 2019.

[34] R. Mane, T. Chouhan, and C. Guan, "BCI for stroke rehabilitation: motor and beyond," *Journal of neural engineering*, vol. 17, no. 4, p. 041001, 2020.

[35] W. Liao, J. Li, X. Zhang, and C. Li, "Motor imagery brain–computer interface rehabilitation system enhances upper limb performance and improves brain activity in stroke patients: a clinical study," *Frontiers in Human Neuroscience*, vol. 17, p. 1117670, 2023.

[36] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Co-adaptive calibration to improve BCI efficiency," *Journal of neural engineering*, vol. 8, no. 2, p. 025009, 2011.

[37] R. Foong, K. K. Ang, C. Quek, C. Guan, K. S. Phua, C. W. K. Kuah, V. A. Deshmukh, L. H. L. Yam, D. K. Rajeswaran, N. Tang, *et al.*, "Assessment of the efficacy of EEG-based MI-BCI with visual feedback and EEG correlates of mental fatigue for upper-limb stroke rehabilitation," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 3, pp. 786–795, 2019.

[38] J. Mladenović, J. Frey, S. Pramij, J. Mattout, and F. Lotte, "Towards identifying optimal biased feedback for various user states and traits in motor imagery BCI," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 3, pp. 1101–1110, 2021.