Manipulating Single-Trial Motor Performance in Chronic Stroke Patients by Closed-Loop Brain State Interaction

Andreas Meinel, Jan Sosulski, Stephan Schraivogel, Janine Reis, and Michael Tangermann

***Abstract*— Motor impaired patients performing repetitive motor tasks often reveal large single-trial performance variations. Based on a data-driven framework, we extracted robust oscillatory brain states from pre-trial intervals, which are predictive for the upcoming motor performance on the level of single trials. Based on the brain state estimate, i.e. whether the brain state predicts a good or bad upcoming performance, we implemented a novel gating strategy for the start of trials by selecting specifically *suitable* or *unsuitable* trial starting time points. In a pilot study with four chronic stroke patients with hand motor impairments, we conducted a total of 41 sessions. After few initial calibration sessions, patients completed approximately 15 hours of effective hand motor training during eight online sessions using the gating strategy. Patients' reaction times were significantly reduced for *suitable* trials compared to *unsuitable* trials and shorter overall trial durations under suitable states were found in two patients. Overall, this successful proof-of-concept pilot study motivates to transfer this closed-loop training framework to a clinical study and to other application fields, such as cognitive rehabilitation, sport sciences or systems neuroscience.**

Manuscript received February 23, 2021; revised July 1, 2021; accepted August 19, 2021. Date of publication August 26, 2021; date of current version September 10, 2021. This work was supported in part by BrainLinks-BrainTools Cluster of Excellence funded by the German Research Foundation (DFG) under Grant EXC 1086, in part by the project SuitAble, DFG, under Grant 387670982, in part by the State of Baden-Württemberg, Germany, through bwHPC and DFG under Grant INST 39/963-1 FUGG, in part by the Baden-Württemberg Ministry of Science, Research and Art (article processing charge), and in part by the University of Freiburg through the funding program Open Access Publishing. *(Corresponding author: Michael Tangermann.)*

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by Ethics Commission of the University Medical Center Freiburg.

Andreas Meinel was with the Department of Computer Science, University of Freiburg, 79110 Freiburg im Breisgau, Germany. He is now with Haufe Group SE, 79111 Freiburg, Germany.

Jan Sosulski is with the Department of Computer Science, University of Freiburg, 79110 Freiburg im Breisgau, Germany.

Stephan Schraivogel was with the Department of Computer Science, University of Freiburg, 79110 Freiburg im Breisgau, Germany. He is now with the ARTORG Center for Biomedical Engineering Research, University of Bern, 3012 Bern, Switzerland.

Janine Reis was with the Department of Neurocenter, University Medical Center Freiburg, 79106 Freiburg im Breisgau, Germany. She is now with Neuropraxis Tuttlingen, 78532 Tuttlingen, Germany.

Michael Tangermann was with the Department of Computer Science, University of Freiburg, 79110 Freiburg im Breisgau, Germany. He is now with the Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 GL Nijmegen, The Netherlands (e-mail: michael.tangermann@donders.ru.nl).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2021.3108187, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2021.3108187

***Index Terms*— Neurorehabilitation, stroke, brain state dependent gating.**

I. Introduction

MACHINE learning methods allow for the single-trial decoding of brain recordings like the electroencephalogram (EEG) to drive real-time applications [1] in brain-computer interface (BCI) systems. Typically BCI's implement a *direct* decoding of, e.g., left and right hand tasks [2] or of attended target stimuli and ignored non-target stimuli [3] to control applications. However, BCI were also suggested to extract information about background brain states [4]–[6]. This is closely related to the research field of passive BCIs [7] where the user's brain state is used as an additional input modality for a technical system. A novel closed-loop system contributing to this branch of research will be presented hereafter.

Focusing on the field of post-stroke motor rehabilitation, a variety of BCI systems have been proposed and their efficacy—as well as the efficiency compared to non BCI-supported baseline methods—is still under intense investigation [8]–[11]. In most applications, the BCI system exploits brain signatures, which are directly informative about an attempted, executed or imagined movement. Commonly, bandpower features of the EEG such as event-related de-/synchronization (ERD/ERS) are used to close the feedback loop for the patient by triggering either a simulated hand movement [12] or a passive movement via an external robotic device, an active orthesis [13], [14] or functional electric stimulation [15], [16].

A patient's ongoing brain signal recording might however provide information beyond the intended movement, and this complementary information could add value for the design of repeated motor tasks as deployed in post-stroke rehabilitation, e.g., using the sequential visual isometric pinch task (SVIPT) for hand motor training [17]. An example of such complementary information has been reported in our earlier study with healthy subjects performing the SVIPT [18], where intra-session performance variations were observed on two different time scales. While inter-session trends mostly reflected motor skill acquisition, we found that trial-by-trial performance variations on the scale of seconds could partially be explained by the power of pre-trial oscillatory activity. It was proposed, that oscillatory power fluctuations are correlated with the interaction of various networks—involving the visual, premotor and motor cortex as well as subcortical and spinal structures [19]–[21].

TABLE I
DEMOGRAPHIC AND IMPAIRMENT DETAILS OF FOUR CHRONIC STROKE PATIENTS INCLUDED IN THE PILOT STUDY

Patient	$P1$	$P2$	$P3$	$P4$
Age (years)	44	64	56	52
Gender (male/female)	m	m	f	m
Affected limb (right/left)	r	r	r	l
Time after stroke (months)	37	83	24	59
Initial Upper Extremity Fugl-Meyer (UEFM)	27	52	53	58
Naïve SVIPT user (yes/no)	y	y	y	n
Naïve BCI user (yes/no)	y	n	y	y

The ability to decode *predictive* oscillatory brain states in quasi real-time can in a next step be used to enhance motor training applications, e.g., with the goal to augment post-stroke motor learning. The prospective use of BCIs as complementary tools in clinical applications calls for a *reliable* decoding of *functionally relevant* features. Specifically, the decoding must be effective also under the challenging conditions posed by the work with patients. To meet these requirements, we propose to use the established work flow for individual single-trial motor performance prediction based on oscillatory brain signal components [18], but to improve it using an additional regularization strategy [22]. For the identification of functionally relevant spatial components, we propose the data-driven regression approach SPoC (source power comodulation [23]), and enhance the selection of components by a mining approach [24].

We hypothesize, that selecting *suitable* and *unsuitable* starting time points defined by individual oscillatory components can induce performance variation for a repetitive hand motor paradigm. We evaluated this hypothesis in a pilot study with four chronic stroke patients.

II. METHODS

A. Subjects

The demographic and impairment related data of the four included chronic stroke patients (abbreviated by $P1$ to $P4$) is reported in Tab. I. Here, the term *chronic* implies that the stroke dated back at least three months prior to participation in our pilot study [25]. The inclusion criteria for participation specified patients between 18 and 80 years that had a first-ever, unilateral ischemic stroke resulting in a hemiparesis and who had sufficient cognitive function to allow for written informed consent and to comply with the task instructions. Patients were excluded when they had any of the following conditions: hemorrhagic stroke, cerebellar infarction, hemiplegia, severe aphasia or neglect, implanted medical devices or intracranial ferromagnetic objects, skull lesions, epileptic seizures in the anamnesis, cognitive impairment and medical, neurological or psychiatric disorders interfering with consistent participation, extreme ametropia, complete paresis of extremity without residual function. The hemiparesis was mild to moderate for $P2$, $P3$ and $P4$, and stronger for $P1$. Three of the four patients were naïve to SVIPT. Patient $P4$ had prior experience with SVIPT due to the participation in the control group of an earlier study [25]. Patient $P2$ had prior experience with a BCI, as he completed a 30 hour language training with feedback based on task-relevant EEG activity [26].

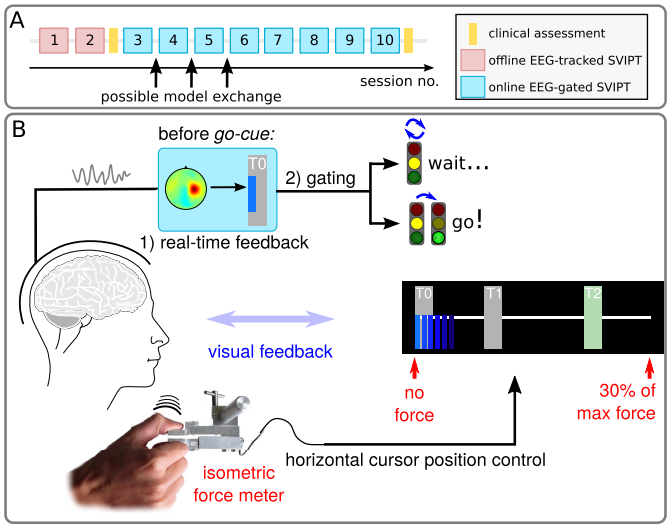

Fig. 1. (A) Study protocol for testing motor performance separability across eight online sessions. Before and after the online training, a clinical assessment was performed. (B) Scheme for the EEG-gated SVIPT. Prior to the start of a trial, the patient received real-time feedback about their brain state by changing vertical cursor positions. The gating strategy determined the *go-cue* time point of every trial. For details on the gating strategy, see Fig. 2.

Following the declaration of Helsinki, all subjects provided written informed consent prior to participation. The study was approved by the local ethics committee of the University Medical Center Freiburg.

B. Experimental Setup

Each pilot patient completed a high-intensity training with their affected hand: about 15 hours of effective training was conducted within three consecutive weeks, comprising ten (patients P2, P3, P4) or eleven sessions (patient P1) of the EEG-tracked SVIPT (Fig. 1 (A)). Among these, two (P2, P3, P4) or three (P1) initial sessions were conducted offline, i.e., without brain-state dependent gating.

SVIPT requires isometric force control of thumb and index finger. After a go-cue has been provided, the user is required to control the horizontal position of a blue cursor (Fig. 1 (B)) through a sequence of target fields as fast and as accurately as possible. Inaccuracies, i.e., overshoots of the target fields, result in adding a penalty second to the trial duration per mistake. Note that the patients trained a specific BCI-SVIPT version with only two target fields, for details about the setup see [18].

Except for $P1$, who had to cope with muscular fatigue and thus could execute 5–9 runs only in six of his sessions, a session generally comprised ten runs of 20 trials each. EEG signals were registered from 63 passive Ag/AgCl electrodes placed according to the extended 10–20 system recorded using a BrainProducts BrainAmp DC. If not explicitly noted differently, further experimental details of a single EEG-tracked SVIPT session were identical to those of an earlier offline study [18]. After each run, an individual high score with the

best average trial duration per run was displayed to motivate patients to further improve their performance. This trial duration also included the penalty seconds for inaccuracies, i.e., overshoots of the target fields, and therefore reflects both the accuracy and the speed of the patient.

The initial offline sessions were used to calibrate an individual performance prediction model, see Sec. II-B1. During the SVIPT's *get-ready* phase a light blue cursor is presented in the leftmost target field $T0$ while patients were asked to fixate the center of $T0$. At online sessions only, the system aimed to influence the patient's upcoming performance during this *get-ready* phase by: (1) Providing continuous visual feedback—a continuously updated vertical cursor position—about the ongoing brain state as sketched in Fig. 1 (B). (2) Depending on the current brain state estimate, a temporal gating strategy was realized: A *go-cue* was elicited either if a user-specific prediction model indicated a desired brain state or when a timeout criterion was met (for details see Sec. II-B2).

In offline sessions, patients unknowingly received pseudo-feedback, as the varying vertical cursor positions were independent of the true ongoing brain state and instead reflected brain state estimates recorded earlier from a healthy subject.

1) Performance Prediction Models: The prediction of single-trial motor performance from oscillatory EEG can be realized with the supervised spatial filtering algorithm named source power comodulation (SPoC) introduced by Dähne and colleagues [23]. Given N_e band-pass filtered multichannel data epochs $\mathbf{X}(e) \in \mathbb{R}^{N_c \times N_s}$, with N_s sample points per epoch and N_c EEG channels, as well as corresponding continuous labels $z(e)$ for each epoch e, SPoC learns an optimal spatial filter $\mathbf{w}^* \in \mathbb{R}^{N_c}$ with N_c recorded channels which maximize the epoch-wise co-modulation between the bandpower of a source $\hat{s}(e)$ and the given target variable $z(e)$. Given a single spatial filter $\mathbf{w}$, the epoch-wise bandpower $\Phi(e)$ of the corresponding source $\hat{s}$ can be approximated by its variance:

$$\Phi_{\hat{s}}(e) = \mathrm{Var}[\hat{s}(t)](e) = \mathrm{Var}[\mathbf{w}^\top \mathbf{x}(t)](e) = \mathbf{w}^\top \Sigma(e)\, \mathbf{w}, \quad (1)$$

where $\Sigma(e) = (N_s - 1)^{-1} \mathbf{X}(e)^\top \mathbf{X}(e)$ denotes the epoch-wise spatial covariance matrix.

As in our patient scenario training data was extremely limited, we deployed NTik-SPoC, a variant which uses Tikhonov regularization in the objective function [22]. By adding an L2 penalty to the weight vector, the estimated spatial filters have small weights which prevents overfitting. This form of regularization has also been successfully applied to the common spatial pattern (CSP) algorithm [27]. Additionally, in order to avoid the problem that covariance matrices are not scale invariant, they are first normalized towards the matrix trace [28]. The algorithm was trained individually per patient. Training was conducted on the pooled patient-specific offline sessions using the first 80% (in chronological order) per session. The remaining offline session data served to validate the obtained models and the online performance.[1] According to the offline analyses reported by Meinel and colleagues on data of normally aged subjects [18], reaction time (RT) was the performance metric which allowed for the highest decoding accuracy. In addition, it was found that RT is correlated to other motor performance metrics of the SVIPT task. In chronic stroke patients, RT generally is lower than in controls, e.g., Sheng and Wan observed, that the RT obtained during wrist flexion and extension tasks correlates negatively with the Wolf Motor Function Test [29]. Thus RT was selected to provide the continuous labels to train NTik-SPoC. Accordingly, we aimed to primarily modulate the upcoming RT in the gated online SVIPT sessions. As proposed by [24], we embraced the unavoidable spatial filter variability under different hyperparameter configurations by training NTik-SPoC components in a large configuration space. The SPoC spatial filters were trained on narrow-band filtered data in 2-5 Hz wide intervals between 1 and 48 Hz. We additionally evaluated time intervals of different lengths in intervals between -1 s and $+1$ s relative to the go cue and different strengths of regularization using the NTik-SPoC algorithm, as expressed by α values ranging between 10^{-10} and 10^{-2}. The resulting set of components contains many similar ones. We obtained functionally reliable representative components from this set by applying the DBSCAN clustering approach [30] as proposed for such component data by [24]. The resulting clusters contained components with, e.g., similar patterns and envelope dynamics, and the corresponding cluster representatives were considered candidates for good components.

Although RT provides continuous labels, we were mainly interested in determining suitable and unsuitable brain states, i.e., the gating strategy involves a two-class problem. While this corresponds to a classification problem, we nevertheless opted for the regression approach provided by SPoC, as first it is able to deal with the continuous performance labels, and second the continuous output allows to, e.g., differentiate the good from the very good trials in a post-hoc analysis. However, a discretization of the continuous SPoC outputs to the suitable/unsuitable classes finally had to take place, as described in Sec. II-B3. To realize this, we split the offline RT data at the median into an upper and lower half. Based on this, we could test how well the corresponding two power distributions of any component would predict the class. The obtained measure is called z-AUC and has a chance level of 0.5 [18]. All candidate components considered for the clustering step were required to have z-AUC$_{min} \geq 0.6$ on the validation data. Furthermore we required, that they were of neural rather than artifactual origin, which was expressed by an artifact probability rating $p_{art} \leq 10^{-5}$ derived by the automatic component classification approach MARA (cf. Sec. II-B4 for details). In a final step, a manual inspection was performed among the remaining candidates to select *one* specific oscillatory component $\mathbf{w}$ per patient for the closed-loop interaction in the following online sessions. Hereafter, this will be referred to as the *selected* oscillatory component. For the manual selection, the following criteria were considered: **Rich envelope dynamics**: Driven by the findings of [24], oscillatory components revealing a rich within-trial envelope dynamics were preferred. An ERD effect upon *get-ready* events and/or

[1]For patient $P1$, the model was trained on the first two sessions and evaluated on the third one.

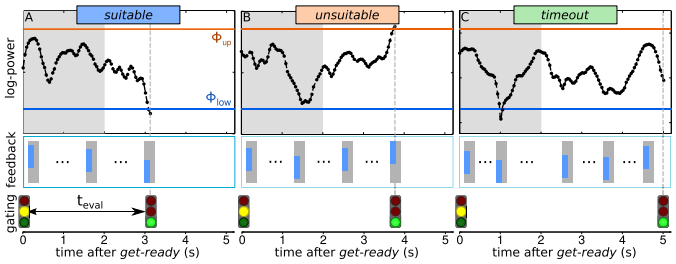

Fig. 2. Online gating: three examples of *get-ready* periods. The two upper rows show the continuously sampled power and the displayed cursor feedback. The lower row illustrates the course of events showing the *get-ready* and *go-cue* time points of these trials. Please note that the patient only saw the visualization depicted in the second row. Initiated by the *get-ready* event (yellow light), the power of the selected oscillatory component was continuously estimated (black dots) and translated into a vertical cursor position. The *go-cue* time point (green light) was triggered earliest 2 s after *get-ready* and by distinguishing between three cases: (A) if the power $\Phi(t)$ fell below the threshold Φ_{low}, a *suitable* trial was elicited. (B) If $\Phi(t)$ exceeded the threshold Φ_{up}, an *unsuitable* trial was initiated. (C) A *timeout* trial was started, if the power did not exceed either of the thresholds in the interval $[2, 5]$ s.

the *go-cue* as well as an ERS upon the events *hit 3* or *hit 4* substantiate the neurophysiological plausibility of individual components. **Motor-relatedness**: Priority was given to motor components, specifically if their patterns were lateralized over the patient's affected hemisphere. **Stability**: Similarity of patterns (visually inspected) across offline sessions was required to ensure stability of the target component.

After each of the first three online sessions, our protocol allowed to refine the selected component. For this purpose both, the currently selected and potentially novel component candidates were evaluated on the most recent unseen data. In case the prediction performance of a candidate component was clearly outperforming the currently selected one and if it simultaneously fulfilled the stated criteria, the selected component and decoding model was updated for upcoming online sessions.

2) Brain State-Dependent Gating Strategy: The selected subject-specific spatial filter $\mathbf{w}$ was now used in a closed-loop setting, see Fig. 1 (B). During the *get-ready* phase and prior to the *go-cue* of a SVIPT trial, the component's log-bandpower $\Phi(t)$ was evaluated every 40 ms from data sampled at 1 kHz. For brevity, power estimation hereafter always refers to log-bandpower estimation. For this, window lengths of 400 ms (alpha) and 300 ms (beta) had been determined based on data of healthy subjects (not shown), representing trade-offs between sufficient data to estimate the ongoing power and short delays to not miss *go-cue* time points for gating. The online data were bandpass filtered by a Butterworth filter to the same frequency band as the training data of the selected model.

During the closed-loop interaction the estimated ongoing brain state, i.e., the selected component's power, was translated into a vertical cursor position (Fig. 2). The online gating strategy was applied with the goal to manipulate the patient's performance during each upcoming trial: *go-cue* events were triggered if (according to the spatial filter model) either a particularly short or long reaction time was expected. Given the estimate $\Phi(t)$, the *go-cue* event was triggered by a threshold scheme as sketched in Fig. 2: a *suitable* trial was elicited if the component's power would deceed a lower threshold Φ_{low}, while an *unsuitable* trial was triggered if $\Phi(t)$ exceeded an upper threshold Φ_{up}.[2] If none of the thresholds were passed within 5 *s*, the *go-cue* was elicited and the trial was labeled as *timeout*. Note that in offline sessions the time t_{eval} between *get-ready* and *go-cue* was sampled from the interval $[2, 5]$ s, based on statistics of pilot online sessions with healthy users. For consistency, the same timeout limits were applied also for the online sessions with patients. Any threshold crossed within the first two seconds was ignored (see Fig. 2 (C)). To make offline and online sessions as similar as possible, the vertical cursor position was varied also during offline session, but based on pre-recorded data.

In online sessions, patients were motivated to explore strategies to lower the vertical cursor position, i.e., to obtain a more suitable brain state by modulating their brain signals during the *get-ready* phase. They were informed that the cursor position reflected the quality of their general upcoming motor performance but did not know that RT was the targeted metric.

As single sessions were strictly limited to 200 valid trials (corresponding to ≈90 min), an online artifact detection was applied throughout the *get-ready* phase to reduce the rate of artifactual trials [31], [32]. Therefore, a min-max threshold of 100 μV for frontal EEG channels' activity was applied after bandpass-filtering to [0.7,45] Hz. In case of a threshold violation, the trial was immediately aborted. The patients received visual feedback about this, and the trial was restarted after 2 s. Aborted trials did not enter the post-hoc analysis.

3) Online Adaptation of the Prediction Model: We aimed at triggering $p^*_S = 55\%$ *suitable* trials and $p^*_U = 35\%$ *unsuitable* trials. This ratio prefers successful, rewarding trials for the patient while ensuring sufficient data of both conditions for statistical comparison. In addition, we expected 10% *timeout* trials in total. The ratio was controlled by a careful online adaptation strategy: Switching from offline sessions to an online session, we expected that a component's power may reveal a different distribution [33], an effect commonly known as *covariate shift* [34]. Various approaches have been proposed to account for non-stationarity characteristics of neural signals in closed-loop BCI applications [35]–[38]. In this pilot study, we decided to fix the selected filter $\mathbf{w}$ and to counter non-stationarity by adapting the decision boundaries, namely the gating thresholds Φ_{low} and Φ_{up} during online sessions.

In the first online session, the gating thresholds were initialized by the 5th and 95th percentiles of the power distribution of the selected component $\mathbf{w}$ as determined on the offline training data. Further on, we performed an adaptation every five trials, a so-called *update block* of trials. To reach an

[2]This example assumed a negative correlation of power $\Phi(e)$ with motor performance $z(e)$. The sign of the correlation $R(z, \Phi(\mathbf{w}))$ is obtained from training data.

intended gating ratio while also coping with non-stationarities, we implemented two update mechanisms:

a) Coarse unsupervised adaptation: Aimed to counter potential large fluctuations of oscillatory power between sessions (or after a long break within a session). Typically, it was used during the first run of a novel session. We assumed a stable power fluctuation width across sessions and that only the average power level shifts from session j to $j+1$, as observed for CSP components [39]. Based on the median of the sampled power $\Phi_m(j+1)$ of the new session, the last available upper threshold from the previous session j was updated by:

$$\Phi_{up}(j+1) = \Phi_{up}(j) + (\Phi_m(j+1) - \Phi_m(j)) \qquad (2)$$

$\Phi_{low}(j+1)$ was updated correspondingly. The respective power medians in Eq. (2) were estimated from data kept in a session-specific ring buffer. It contained up to 1000 power estimates of the latest time windows. If the experimenter recognized a need for faster adaptation, its size could be reduced to between 500 and 100 entries, which became effective with the next session. In each novel session this ring buffer was initialized and filled, and the coarse unsupervised strategy was deployed throughout the first run of each online session. The adaptation took place after every update block only. The buffer was updated with samples collected from between 1.5 s after *get-ready* and the *go-cue* only, as earlier samples would have introduced a systematic bias due to *get-ready* triggering an ERD effect in most components.

b) Refined supervised adaptation: Started with the second run. Applied supervisedly after each block, it aimed at reaching the intended gating ratios during online sessions. To this end, the sampled probabilities p_S and p_U of the suitable and unsuitable gating conditions were evaluated on the latest 60 trials and used as labels. The iterative adaptation was proportional to the signed square of the condition-related probability deviation $\delta p_U = (p_U - p_U^*)$ to penalize strong deviations from the expected label distribution. In addition, the update was performed relative to the absolute distance of the threshold $\Phi_{up}(i)$ used in the last iteration i to the power median Φ_m based on the previously mentioned ring buffer of the current session $j+1$. These steps are executed analogously for p_S. Together, this translated to the following adaptation from update block i to $i+1$ for the upper gating threshold:

$$\Phi_{up}(i+1) = \Phi_{up}(i) + \eta \cdot \mathrm{sgn}(\delta p_U) \cdot (\delta p_U)^2 \cdot |\Phi_{up}(i) - \Phi_m| \qquad (3)$$

with a fixed learning rate $\eta = 2$ determined on earlier pilot data. $\Phi_{low}(i+1)$ was updated analogously. To account for noisy estimates, threshold updates were executed only, if the class-related probability deviations δp_U and δp_S exceeded an absolute tolerance level of $\delta p_{tol} = 0.05$. After each run, the experimenter received feedback about the component's within-session power time course and the selected thresholds. Experiencing severe non-stationarities, e.g., when noisy EEG channels had to be fixed, the experimenter could resort to an additional coarse unsupervised adaptation (see Eq. (2)). This happened in four sessions of P2 and four of P4.

4) Data Preprocessing: For post-hoc evaluation, linear Butterworth filters of fifth order were applied: the raw EEG signals were low-pass filtered at 100 Hz, sub-sampled to 500 Hz sampling rate before high-pass filtering at 1 Hz.

Per session, noisy channels were removed. First, the variance of single epochs and channels was computed. Based on the pooled statistics, channels were removed, if they exceeded the 90$^{\text{th}}$ percentile or undershot the 10$^{\text{th}}$ percentile by at least two times the [10, 90] range. Considering also the influence of outlier epochs, more channels could be removed in a second step: The N_{rej} out of all N_{all} epochs which exceeded the a min-max threshold on any channel were marked for rejection. Then channels responsible for (a) more than $0.1 \cdot N_{\text{rej}}$ outlier epochs and (b) for at least $0.05 \cdot N_{\text{all}}$ outlier epochs were removed. Finally, we updated the list of outlier epoch by re-running the min-max test and removed those that, despite the meanwhile reduced channel list, were still marked for rejection.

Artifact cleaning was done using an independent component analysis (ICA) decomposition of pooled data of the active trial phase—from *get-ready* to *trial end*—of each session. For the calculation of the ICA, we used the FastICA algorithm [40]. To restrict the computational effort for the ICA, only one randomly selected run was included per session.[3] The obtained ICs were rated for artifactual origin with the automated artifact detection framework MARA [41]. For the post-hoc offline analysis, based on MARA ratings, a maximum number of 10 artifactual ICs were removed from the EEG data before projecting it back into the original sensor space.

Based on the force sensor recordings, different single-trial motor performance metrics such as force jerk or force path length were extracted per trial, see [18] metrics. For the post-hoc analysis of behavioral data, the first ten trials per session were omitted due to the ramp-up phase of the threshold adaption at session start.

III. Results

A. Effect of Uncorrelated Feedback

We analyzed if brain state independent vertical cursor position at the *go-cue* could potentially suffice to explain observed reaction time changes. We could evaluate this for P3 and P4 only, as cursor positions during the offline sessions of P1 and P2 were not logged. Per patient, a two-sided Wilcoxon signed rank test determined whether the RT distributions were different between up or down cursor positions. Additionally we calculated the median difference, i.e., median reaction time at cursor up position minus the median reaction time at cursor down position. The difference for P3 was −26 ms with a mean absolute RT deviation (MAD) of 106 ms. For P4 we obtained a difference of +8 ms with a MAD of 56 ms. For neither P3 nor P4 did we find a significant difference in the RT distributions between cursor positions (P3: $p = 0.056$, P4: $p = 0.541$).

[3]For *P*1, the sessions 1 and 2 were excluded for ICA training due to instabilities in the resulting decompositions.

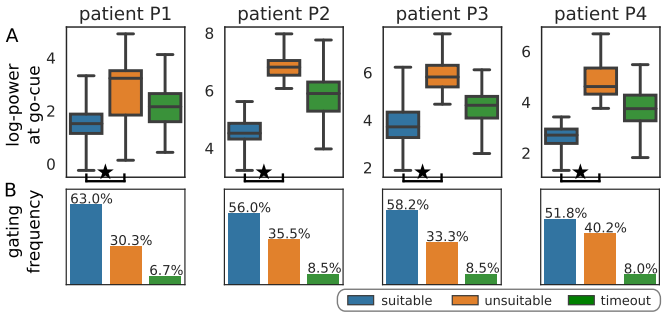

Fig. 3. (A) The box plot for each patient reports the component power (pooled over online sessions, but separately for gating conditions) that triggered a *go-cue* event. Stars mark statistically significant differences between the effects of the suitable (blue) and unsuitable (orange) gating strategies (two-sided Wilcoxon rank-sum test, $p < 0.01$). Each box shows the quartiles of the underlying data, whiskers refer to two interquartile ranges. (B) The achieved frequencies of gating conditions across all sessions.

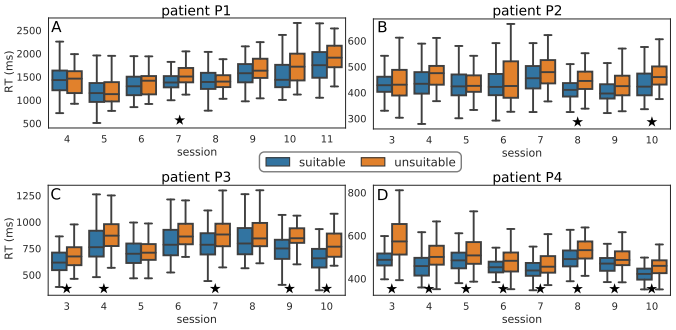

Fig. 4. Behavioral results: development of the RT distributions for *suitable* and *unsuitable* trials across all online sessions. A star refers to single sessions with a significant difference (two-sided Wilcoxon rank-sum test, $p < 0.05$ with Holm-Bonferroni correction) between conditions.

As this observation is based on limited data, it can only be a first indication that reaction time separation might not depend on the cursor position.

B. Brain State-Dependent Gating

Separately per patient, but pooled across their online sessions, Fig. 3 (A) reports the mean single-trial power distributions (after post-hoc artifact rejection) observed immediately before the *go-cue* for the selected oscillatory components. A Wilcoxon rank-sum test determined significance of the power contrast between suitable and unsuitable trials.

As intended by our gating strategy, a significant split between the *suitable* and *unsuitable* power distributions could be observed for each of the four patients. In addition, the mean log-power values of *timeout* trials ranged between those of the suitable and unsuitable conditions. Fig. 3 (B) reports the achieved trial ratios across gating conditions. As targeted by the adaptation parameters p_S^* and p_U^* (see Sec. II-B3), all four patients could be trained with a majority of *suitable* brain states. Furthermore, the gating statistics of $P2$ to $P4$ demonstrate that the desired gating ratios were reached up to small deviations.

C. Single-Trial Motor Performance Caused by Different Gating Strategies

1) Effect on Single-Trial Reaction Time (RT): Fig. 4 visualizes the behavioral results separately for suitable and unsuitable brain states at *go-cue* per patient and across all online sessions. In two of four pilot patients we found that RT distributions separate significantly in at least half of the online sessions. Interestingly, in the best case ($P4$) a significant difference throughout all eight online sessions was achieved. Even though the online framework was tested with a heterogeneous patient group regarding their initial hand motor impairment (UEFM score ranges: 27–58), on pooled data across all online sessions a significantly shorter RT for *suitable* trials could be observed for all of them (see Fig. 5 (A)).

TABLE II
ESTIMATES OF EXPLAINED IQR (A MEASURE OF EFFECT SIZE, SEE FULL-TEXT FOR DETAILS) FOR EACH PATIENT USING TWO DIFFERENT METHODS FOR CALCULATING THE INTERQUARTILE RANGE

	P1	P2	P3	P4
explained IQR (estimate 1) [%]	24.48	26.96	38.05	47.64
explained IQR (estimate 2) [%]	23.56	25.91	36.59	44.23

To quantify the effect size of an intervention, one typically relates the differences between conditions' means to the standard deviation of the distribution(s), the so-called standardized mean difference (SMD). As the RT is not normally distributed, we prefer to use *explained IQR* over SMD. It replaces the mean estimates of the *suitable* and *unsuitable* conditions by the more robust median estimates and the standard deviation by the interquartile range (IQR) of observed RT values, i.e., the distance between the 25$^{\text{th}}$ and the 75$^{\text{th}}$ percentile.

It is not straight forward to estimate the true RT distribution or its IQR, as applying the gating strategy corresponds to an intervention which may influence the RT values. While timeout trials are intervention free and could deliver the IQR in principle, their number is too small for a robust estimate. We mitigate this problem by estimating the IQR in two ways: *IQR estimate 1* is calculated by first subtracting the median of each condition before pooling the data of both conditions. It most likely underestimates the true IQR and thus overestimates the explained IQR. *IQR estimate 2* directly pools the data of both conditions. It will overestimate the true IQR and underestimate the explained IQR. Thus the true explained IQR will likely be between these two estimates. Luckily we observed, that the over- and underestimated values differ slightly only (range: 0.92%–3.41%) when computed for each session individually and then averaged across sessions (see Table II).

2) Indirect Transfer to Other Motor Performance Metrics: The NTik-SPoC model was optimized for predicting the metric RT based on data collected immediately before the *go-cue*. It is interesting to see, if brain states predictive for RT may also be predictive for other motor performances, specifically for metrics that integrate over longer within-trial periods.

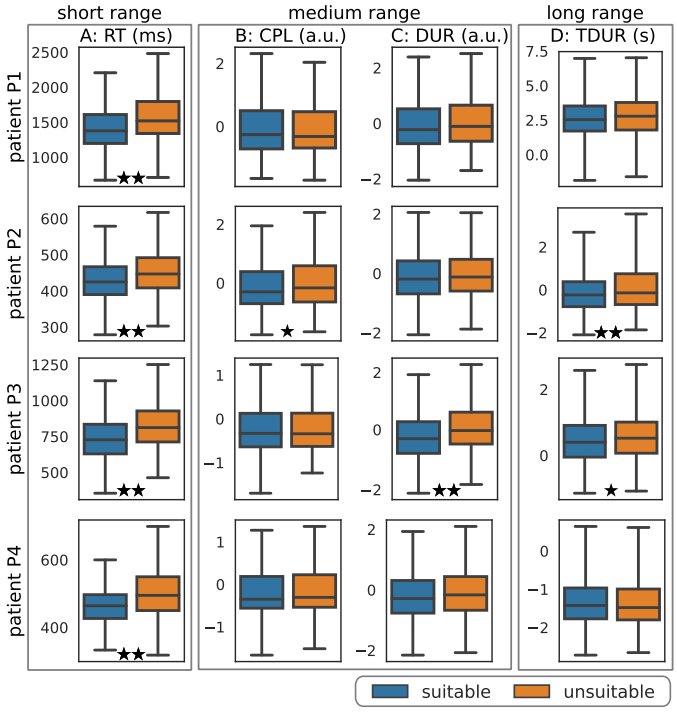

Fig. 5. Comparison of four behavioral performance metrics. From left to right, the temporal integration ranges after *go-cue* increase from short range (few hundred milliseconds) to long range (few seconds): (A) reaction time, (B) cursor path length up to *hit 1*, (C) time to *hit 1* and (D) full trial duration (session-wise median subtracted). Each row refers to a single patient. Per patient, data of the online sessions has been pooled. For each of the four metrics, significant differences between conditions suitable (blue) and unsuitable (orange) are reported after Holm-Bonferroni correction for values of $p < 0.01$ (**) and $p < 0.05$ (*).

On pooled data of all online sessions, Fig. 5 reports the individual distributions of different single-trial performance metrics. Their arrangement from (A)–(D) show an increasing temporal integration along the trial. While the metric RT typically takes into account a few hundred milliseconds after *go-cue*, the cursor path length (CPL) and duration (DUR) were calculated based on the time interval required to fulfill *hit 1*, which corresponds to approximately 40% of the full trial time. Both metrics were standardized per session and hit sequence and subsequently pooled. In a previous study, the metrics RT, CPL and DUR have been found to be mostly uncorrelated [42]. For the trial duration metric (TDUR), which integrates multiple seconds, the session-wise median was subtracted before pooling to eliminate session-to-session differences. Significance between conditions was tested by a two-sided Wilcoxon rank-sum test with Holm-Bonferroni correction for multiple tests.

While on pooled individual RT data a homogeneous picture in terms of a distinct separation was noticed in all four patients, the brain state-dependent gating translates only partially to the metrics which integrate performance over longer time intervals. Interestingly, for two out of four pilot patients a significant effect on trial duration was induced, even though the decoding model of the BCI system had not been trained on these labels.

TABLE III

PER PATIENT, THE HYPERPARAMETERS AND THE INITIAL DECODING ACCURACY ARE REPORTED FOR THE *Selected* SPATIAL FILTER MODEL. AS FOR P1 AND P3 THE MODEL WAS SWITCHED ONCE DURING ONLINE TRAINING, INFORMATION FOR BOTH MODELS IS PROVIDED IN SEPARATE ROWS, WITH BRACKETS CONTAINING THE RANGE OF SESSIONS A MODEL WAS APPLIED

patient	$P1$	$P2$	$P3$	$P4$
central frequency f_0[Hz] (session)	9.5 (4)	21.8	18.8 (3-5)	9.7
	8.7 (5-11)		10.9 (6-10)	
reg. parameter α	$1.0 \cdot 10^{-2}$	$4.4 \cdot 10^{-5}$	$8.7 \cdot 10^{-6}$	$1.0 \cdot 10^{-2}$
	$8.5 \cdot 10^{-4}$		$3.4 \cdot 10^{-3}$	
training epochs N_{train}	342	344	289	347
decoding accuracy z-AUC	0.62	0.62	0.65	0.69
	0.67		0.69	
MARA reject prob. p_{rej}	$1.43 \cdot 10^{-9}$	$1.22 \cdot 10^{-6}$	$6.81 \cdot 10^{-6}$	$7.44 \cdot 10^{-9}$
	$5.56 \cdot 10^{-9}$		$3.28 \cdot 10^{-7}$	

D. Across-Session Feature Stability

A spatial filter model to extract an oscillatory subspace had been defined during early sessions and was kept fixed during all (P2, P4) or most (P1, P3) of the online training sessions. We expected the subspace components to be stable across sessions with similar spectra and ERD/ERS characteristics. This robustness expectation can now be investigated by a post-hoc analysis of the recorded EEG data across training sessions, where the online sessions serve as unseen datasets to validate the robustness expectation. Details about the selected oscillatory components are summarized in Tab. III. Hereafter, various stability aspects regarding the used oscillatory subspace components are reported over the course of the training.

Spectral and Spatial Feature Analysis: The robustness of individually *selected* components—under scarce training data and across multiple sessions—was an important characteristic for this pilot study. During the decoding and selection of components, we for this reason applied procedures that would prefer robust components and were known to compensate label noise and small training data sets well. Besides regularization of SPoC, these procedures comprised a clustering of overcomplete decompositions as described in detail in earlier contributions [22], [24] in order to select a spatial component. Note that during early sessions of the rehabilitation training for P1 and P3, we noticed that the selected component was not performing well enough. Therefore, we re-trained the model and selected a new component for the remaining rehabilitation sessions. For the re-training for P1 we used the data of the first three sessions, and validated the obtained component on the fourth session during the re-training, whereas for P3 we used the first 80% of the data obtained from sessions 1-5 for training and the remaining data for validation. We limit the analyses presented in this section to these *finally selected* components, as they had been used for at least the second half of the sessions.

To evaluate the robustness of these *finally selected* individual spatial filters $\mathbf{w}$, Fig. 6 displays their spectral content separately per session, including training and online sessions, and—in case of P1 and P3—also those initial online

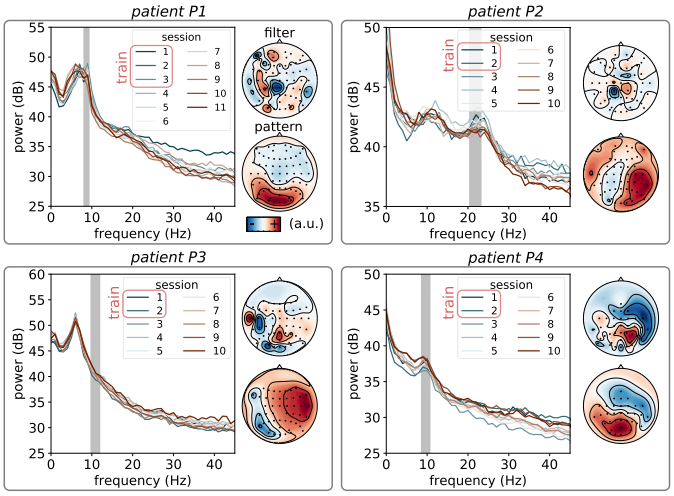

Fig. 6. Per patient, the pre-go power spectra of the *finally selected* subspace components are visualized. Please note, that the spectra are provided also for the offline sessions and for early online sessions of P1 and P3 which had used different components during the training. The gray shaded area in each power spectrum refers to the frequency band to which the data had been filtered to train the model. In addition, the individually selected spatial filters and corresponding pre-go spatial activity patterns averaged across all sessions are displayed.

sessions, during which another component had originally been selected.

The spectra were calculated on non-frequency filtered data but after projection to the *selected* subspace and using data segments prior to the *go-cue* only. In addition to the spectra, the filters and spatial patterns are provided.

We observed, that within each patient the power spectra are highly similar across sessions and reveal the well-known spectral $1/f$ decaying characteristic with distinct α- and/or β-modulations. This observation is a strong first indicator for the across-session robustness of the *finally selected* components.

An additional neurophysiological introspection is provided by assessing the spatial activation patterns of an oscillatory component and the average event-related (de-)synchronization of each component relative to events. As documented in Fig. 1 of the supplementary, the selected oscillatory components revealed mostly stable pre-go activity patterns over the course of the multi-session training. Figure 2 of the supplementary reveals a stable ERD-/ERS characteristic over sessions.

IV. Discussion

This work presented a novel brain state-dependent closed-loop interaction protocol. It was exemplified in a pilot study for a repetitive hand motor training task with four chronic stroke patients. Even though a total number of 41 experimental sessions were executed, we are aware of the limitation, that the following discussion will be based on data of four individuals only. In the future, a randomized controlled clinical trial shall allow to critically review these first observations.

The training protocol was designed to influence a patient's motor performance observed for repetitive trials of the training task. For this purpose, an individually optimized oscillatory feature was determined from EEG recordings, which allowed to discriminate between *suitable* and *unsuitable* brain states w.r.t. executing the hand motor task well. Using this information, the start of each training trial could be gated. The gating protocol was implemented in a closed-loop system, which operated at an update rate of 40 ms. In every trial of the online BCI-gated SVIPT motor task, the system strived for selecting specifically *suitable* or (less frequently) *unsuitable* starting time points.

A. Successful Performance Manipulation

Despite potentially patient-related challenges, we found clear evidence that the online gating strategy not only succeeded in separating the average oscillatory power between conditions but also that the proposed protocol manipulated single-trial reaction times in all four chronic stroke patients: Trials started during *suitable* brain states resulted in improved motor performance metrics compared to trials started during *unsuitable* brain states. Specifically, this intervention explained up to 44% of individual variability (measured by IQR) observed for the reaction time.

On the level of single sessions (Fig. 4), performance difference was observed for the 4×4 online sessions, which formed the second half of the online trainings. These 16 sessions uniformly display a shorter average reaction time under suitably gated trials, with 9 of these session showing significant difference. This might be driven by the time the patients required to learn the reliable down-modulation of the pre-trial power.

Comparing our results to the recent literature, it is worth noting that our protocol is capable to exploit spontaneous power fluctuations and that it has not explicitly trained patients up to use volitional modulation of pre-trial power. Norman and colleagues instead performed an explicit SMR training over multiple sessions prior to the online phase [43]. In their pilot robot-assisted hand motor training with 8 chronic stroke patients, only 4 of them achieved the required reliable SMR control. In their online training spanning 3 sessions, finally in 3 out of 8 patients shorter reaction times were found if pre-trial oscillatory activity had been low. Overall, the comparison of our study with [43] shows that we have performed more than the double amount of online motor training sessions—which supports a robust validation of our BCI-supported gating concept. In addition our protocol minimizes the number of preparatory sessions and maximizes a patient's effective training time on the motor task, however at the price of a potentially reduced volitional control over pre-trial component power. Which of the two strategies should be preferred with an optimized rehabilitation efficiency in mind remains subject to further studies.

Three earlier studies by groups around Norman, McFarland and Boulay [43]–[45] have introduced protocols for performance manipulation. They succeeded in at most half of the subjects, while our novel protocol achieved the performance manipulation in all four patients. This finding encourages us to explore the proposed data-driven framework for building

the performance predictors of future closed-loop training systems.

In accordance with three studies is our finding that a reduced SMR power immediately before the *go-cue*, correlates with shorter reaction times. Interestingly, this relation has by now been established in at least four different motor tasks (including SVIPT) and even though the four substantially varied in their underlying complexity.

The selected decoding model was optimized for reaction time, but we are aware, that reaction time may not be the metric clinically identified as most relevant for post-stroke motor learning even though shorter reaction time should serve as a useful useful basic building block for the successful execution of many higher-level sensorimotor tasks. However, we also observed, that the training under *suitable* brain states partially translated also into significant enhancement of additional metrics like the trial duration, which integrates SVIPT performance over a range up to multiple seconds (Fig. 5). This is interesting, as intuitively one may expect that the more behavioral information is integrated by a metric along the duration of a single trial, the less influence can be traced back to the brain state at *go-cue* time point. This finding is also in accordance with studies by [43], [44] in which a behavioral performance split for metrics with longer integration intervals was observed only in a smaller fraction of subjects. However, in our data we observed a separation on trial duration for those two patients ($P2$ and $P3$) which also showed the largest SVIPT motor learning effects along the training (see Fig. 4 in the supplementary). Again, however, data from more patients shall be acquired by future studies in order to investigate this.

B. Calibration of the Prediction Model

The calibration of a robust decoding model to predict upcoming single-trial motor performance is challenged by a low amount of available training data. We thus decided to use an online artifact detection pipeline already at the pre-go phase, which restricted the loss of data points for model training after EEG preprocessing to a minimum.

Earlier parameter studies on the regularized NTik-SPoC method [22] showed, that saturated decoding performance requires about $N_{train} \geq 200$ trials. Thus, two offline sessions of 200 trials each were conducted for the initial calibration with a chronological train/test-split to select an individual oscillatory component. A retraining of the model on all available labeled data after each session was not performed to maintain feature introspection and avoid coping with rank instabilities [18]. For patients $P3$ and $P4$ an initial training of the spatial filter models on all available channels resulted in a large fraction of artifactual components. As a mitigation strategy we removed frontal channels from the training of NTik-SPoC and observed an increased ratio of neural compared to artifactual components.

Even though we used single predictive and functionally relevant components, from a machine learning perspective it could be beneficial to combine multiple regularized SPoC features obtained from across multiple frequency bands—comparable to the filterbank CSP approach [46]. This fusion of features could be realized by an additional regression model and may allow for enhancing the trial-wise performance prediction. Similarly, the combination of predictors obtained for different performance metrics might serve to gain an overall enhanced predictive power or a model that serves a more generalized notion of SVIPT performance.

C. Careful Adaptation of the Prediction Model

As non-stationary effects in recorded brain activity can impede a robust decoding [39], a careful online adaptation was applied. While the selected spatial filter model was fixed after the initial calibration, we adaptively compensated for continuous power changes by adapting the gating thresholds. This allowed us to reach the targeted gating ratio across the full training for each individual patient despite the power fluctuations of the component over and within sessions. The control over this ratio is of specific importance if the influence of different gating strategies on post-stroke motor learning shall be studied in future experiments. As previously argued by Biasiucci and colleagues [47], a repeated re-calibration during the BCI training, such as realized by, e.g., Ang and colleagues [48], might over time translate into different oscillatory components. In case they reflect diverging efferent pathways, the change of these components may hinder training-induced plastic changes. Overall, the details and the degree of model adaptation over the course of a closed-loop training is an ongoing debate and requires further investigation.

D. In-Depth Introspection and Monitoring Over Sessions

Individually derived oscillatory components allow for monitoring a training progress over sessions. The combination of spectra, of ERD/ERS characteristics (e.g. contrast and latencies relative to events) and of spatial patterns can provide the clinical expert with in-depth introspection, which may complement functional assessments over the course of the training. As an example, see the envelope characteristics of P2 as shown in the supplementary in Fig. 3.

E. Applicability of Brain State-Dependent Gating

The proposed brain state-dependent gating concept is complementary to most existing BCI-based systems in the field of post-stroke motor rehabilitation [12], [13]. These systems focus on the direct decoding of movement intentions, while the gating concept is taking trial-wise brain state fluctuations into account which influence upcoming motor performance. Thus, the gating concept could be combined with and prove beneficial for most other repetitive motor paradigms in post-stroke rehabilitation.

Conceptually, the proposed gating concept can be seen as a sample application for brain state-dependent experimenting [5]. As our framework is based on a data-driven decoding model without prior assumptions about the underlying cortical network, it is not limited to a specific application. It only requires single-trial labels of behavioral variability

to identify a corresponding neural correlate. Thus, there are various application fields beyond motor rehabilitation. As an example, we foresee that the brain state-dependent gating concept could be beneficial in cognitive rehabilitation scenarios. Moreover, sports science could profit from the consideration of ongoing brain states [49], specifically for the development of training concepts in which single-trial performance should be optimized. These aspects might play an important role in disciplines such as archery, darts or ski jumping.

V. Conclusion

In a nutshell, this paper presents the feasibility study for an online temporal gating strategy to influence upcoming single-trial motor performance in a post-stroke training protocol. The performance influence was demonstrated in a pilot multi-session hand motor training with four chronic stroke patients which revealed strong trial-to-trial motor performance variations. Even under challenging conditions with patients, we could identify robust and predictive brain states that allowed the gating of *suitable* pre-trial starting time points of an upcoming motor task. Those elicited an improved motor performance. Particularly, single-trial reaction times were significantly reduced—ranging from 23 to 47% of the individual reaction time variations—for *suitable* trials compared with the *unsuitable* trials. As this framework on real-time brain state interaction is not exclusively designed for motor rehabilitation, the detection and exploitation of *un-/suitable* brain states can potentially be transferred to different applications such as cognitive trainings or sports sciences.

References

[1] K.-R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz, "Machine learning for real-time single-trial EEG-analysis: From brain–computer interfacing to mental state monitoring," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 82–90, Jan. 2008.

[2] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 41–56, Jan. 2008.

[3] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components—A tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, May 2011.

[4] J. M. Horschig, J. M. Zumer, and A. Bahramisharif, "Hypothesis-driven methods to augment human cognition by optimizing cortical oscillations," *Frontiers Syst. Neurosci.*, vol. 8, p. 119, Jun. 2014.

[5] O. Jensen *et al.*, "Using brain–computer interfaces and brain-state dependent stimulation as tools in cognitive neuroscience," *Frontiers Psychol.*, vol. 2, p. 100, Mar. 2011.

[6] J.-D. Haynes and G. Rees, "Neuroimaging: Decoding mental states from brain activity in humans," *Nature Rev. Neurosci.*, vol. 7, no. 7, pp. 523–534, Jul. 2006.

[7] O. Thorsten Zander and C. Kothe, "Towards passive brain–computer interfaces: Applying brain–computer interface technology to human–machine systems in general," *J. Neural Eng.*, vol. 8, no. 2, 2011, Art. no. 025005.

[8] M. A. Cervera *et al.*, "Brain-computer interfaces for post-stroke motor rehabilitation: A meta-analysis," *Ann. Clin. Transl. Neurol.*, vol. 5, no. 5, pp. 651–663, 2018.

[9] A. Remsik *et al.*, "A review of the progression and future implications of brain-computer interface therapies for restoration of distal upper extremity motor function after stroke," *Expert Rev. Med. Devices*, vol. 13, no. 5, pp. 445–454, May 2016.

[10] S. R. Soekadar, N. Birbaumer, M. W. Slutzky, and L. G. Cohen, "Brain-machine interfaces in neurorehabilitation of stroke," *Neurobiol. Disease*, vol. 83, pp. 172–179, Nov. 2015.

[11] R. Carvalho, N. Dias, and J. J. Cerqueira, "Brain-machine interface of upper limb recovery in stroke patients rehabilitation: A systematic review," *Physiotherapy Res. Int.*, vol. 24, no. 2, p. e1764, Jan. 2019.

[12] F. Pichiorri *et al.*, "Brain–computer interface boosts motor imagery practice during stroke recovery," *Ann. Neurol.*, vol. 77, no. 5, pp. 851–865, 2015.

[13] A. Ramos-Murguialday *et al.*, "Brain–machine interface in chronic stroke rehabilitation: A controlled study," *Ann. Neurol.*, vol. 74, no. 1, pp. 100–108, Jul. 2013.

[14] A. A. Frolov *et al.*, "Post-stroke rehabilitation training with a motor-imagery-based brain-computer interface (BCI)-controlled hand exoskeleton: A randomized controlled multicenter trial," *Frontiers Neurosci.*, vol. 11, p. 400, Jul. 2017.

[15] M. Takahashi *et al.*, "Event related desynchronization-modulated functional electrical stimulation system for stroke rehabilitation: A feasibility study," *J. Neuroeng. Rehabil.*, vol. 9, p. 56, Aug. 2012.

[16] B. C. A. Osuagwu, L. Wallace, M. Fraser, and A. Vuckovic, "Rehabilitation of hand in subacute tetraplegic patients based on brain computer interface and functional electrical stimulation: A randomised pilot study," *J. Neural Eng.*, vol. 13, no. 6, Dec. 2016, Art. no. 065002.

[17] J. Reis *et al.*, "Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 5, pp. 1590–1595, Feb. 2009.

[18] A. Meinel, S. Castaño-Candamil, J. Reis, and M. Tangermann, "Pre-trial EEG-based single-trial motor performance prediction to enhance neuroergonomics for a hand force task," *Frontiers Hum. Neurosci.*, vol. 10, p. 170, Apr. 2016.

[19] O. Cohen, E. Sherman, N. Zinger, S. Perlmutter, and Y. Prut, "Getting ready to move: Transmitted information in the corticospinal pathway during preparation for movement," *Current Opinion Neurobiol.*, vol. 20, no. 6, pp. 696–703, Dec. 2010.

[20] H. van Dijk, J.-M. Schoffelen, R. Oostenveld, and O. Jensen, "Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability," *J. Neurosci.*, vol. 28, no. 8, pp. 1816–1823, 2008.

[21] J. M. Horschig, W. Oosterheert, R. Oostenveld, and O. Jensen, "Modulation of posterior alpha activity by spatial attention allows for controlling a continuous brain–computer interface," *Brain Topography*, vol. 28, no. 6, pp. 852–864, 2014.

[22] A. Meinel, S. Castaño-Candamil, B. Blankertz, F. Lotte, and M. Tangermann, "Characterizing regularization techniques for spatial filter optimization in oscillatory EEG regression problems," *Neuroinformatics*, vol. 17, no. 2, pp. 235–251, 2019.

[23] S. Dähne *et al.*, "SPoC: A novel framework for relating the amplitude of neuronal oscillations to behaviorally relevant parameters," *NeuroImage*, vol. 86, pp. 111–122, Feb. 2014.

[24] A. Meinel, H. Kolkhorst, and M. Tangermann, "Mining within-trial oscillatory brain dynamics to address the variability of optimized spatial filters," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 378–388, Mar. 2019.

[25] M. Hamoudi *et al.*, "Transcranial direct current stimulation enhances motor skill learning but not generalization in chronic stroke," *Neurorehabilitation Neural Repair*, vol. 32, nos. 4–5, pp. 295–308, Apr. 2018.

[26] M. Musso, D. Hübner, S. Schwarzkopf, C. Weiller, and M. Tangermann, "Evidence for language-specificity of a BCI-based language training," in *Proc. Acad. Aphasia 56th Annu. Meeting*, vol. 12, 2018, doi: 10.3389/conf.fnhum.2018.228.00069.

[27] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 355–362, Feb. 2011.

[28] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, Dec. 2000.

[29] B. Sheng and C.-X. Wan, "Comparison of the reaction time of wrist flexion and extension between patients with stroke and age-matched healthy subjects and correlation with clinical measures," *Chin. Med. J.*, vol. 126, no. 13, pp. 2485–2488, 2013.

[30] S. Erich *et al.*, "DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN," *ACM Trans. Database Syst.*, vol. 42, no. 3, pp. 1–21, 2017.

[31] M. M. N. Mannan, M. A. Kamran, and M. Y. Jeong, "Identification and removal of physiological artifacts from electroencephalogram signals: A review," *IEEE Access*, vol. 6, pp. 30630–30652, 2018.

[32] M. Jas, D. A. Engemann, Y. Bekhti, F. Raimondo, and A. Gramfort, "Autoreject: Automated artifact rejection for MEG and EEG data," *NeuroImage*, vol. 159, pp. 417–429, Oct. 2017.

[33] J. Dean Krusienski *et al.*, "Critical issues in state-of-the-art brain–computer interface signal processing," *J. Neural Eng.*, vol. 8, no. 2, 2011, Art. no. 025002.

[34] H. Shimodaira, "Improving predictive inference under covariate shift by weighting the log-likelihood function," *J. Statist. Planning Inference*, vol. 90, no. 2, pp. 227–244, 2000.

[35] H. Raza, H. Cecotti, Y. Li, and G. Prasad, "Adaptive learning with covariate shift-detection for motor imagery-based brain–computer interface," *Soft Comput.*, vol. 20, no. 8, pp. 3085–3096, Aug. 2016.

[36] M. Spüler, W. Rosenstiel, and M. Bogdan, "Principal component based covariate shift adaption to reduce non-stationarity in a MEG-based brain-computer interface," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, p. 129, Jul. 2012.

[37] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Co-adaptive calibration to improve BCI efficiency," *J. Neural Eng.*, vol. 8, no. 2, 2011, Art. no. 025009.

[38] P. Shenoy, M. Krauledat, B. Blankertz, R. P. N. Rao, and K.-R. Müller, "Towards adaptive classification for BCI," *J. Neural Eng.*, vol. 3, no. 1, pp. R13–R23, Mar. 2006.

[39] C. Vidaurre, M. Kawanabe, P. von Bánau, B. Blankertz, and K. R. Müller, "Toward unsupervised adaptation of LDA for brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 587–597, Mar. 2011.

[40] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, May 1999.

[41] I. Winkler, S. Brandl, F. Horn, E. Waldburger, C. Allefeld, and M. Tangermann, "Robust artifactual independent component classification for BCI practitioners," *J. Neural Eng.*, vol. 11, no. 3, Jun. 2014, Art. no. 035013.

[42] M. Tangermann, J. Reis, and A. Meinel, "Commonalities of motor performance metrics are revealed by predictive oscillatory EEG components," in *Proc. 3rd Int. Congr. Neurotechnol., Electron. Informat.*, 2015, pp. 32–38.

[43] S. L. Norman *et al.*, "Controlling pre-movement sensorimotor rhythm can improve finger extension after stroke," *J. Neural Eng.*, vol. 15, no. 5, 2018, Art. no. 056026.

[44] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw, "Effects of training pre-movement sensorimotor rhythms on behavioral performance," *J. Neural Eng.*, vol. 12, no. 6, Dec. 2015, Art. no. 066021.

[45] C. B. Boulay, W. A. Sarnacki, J. R. Wolpaw, and D. J. McFarland, "Trained modulation of sensorimotor rhythms can affect reaction time," *Clin. Neurophysiol.*, vol. 122, no. 9, pp. 1820–1826, Sep. 2011.

[46] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Hong Kong, Jun. 2008, pp. 2390–2397.

[47] A. Biasiucci *et al.*, "Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke," *Nature Commun.*, vol. 9, p. 2421, Jun. 2018.

[48] K. K. Ang *et al.*, "A randomized controlled trial of EEG-based motor imagery brain-computer interface robotic rehabilitation for stroke," *Clin. EEG Neurosci.*, vol. 46, no. 4, pp. 310–320, 2015.

[49] C. Jeunet, D. Hauw, and J. D. R. Millán, "Sport psychology: Technologies ahead," *Frontiers Sports Act. Living*, vol. 2, p. 13, Feb. 2020.