Beamforming in Noninvasive Brain–Computer Interfaces

Moritz Grosse-Wentrup*, *Member, IEEE*, Christian Liefhold, Klaus Gramann, and Martin Buss, *Member, IEEE*

***Abstract*—Spatial filtering (SF) constitutes an integral part of building EEG-based brain–computer interfaces (BCIs). Algorithms frequently used for SF, such as common spatial patterns (CSPs) and independent component analysis, require labeled training data for identifying filters that provide information on a subject's intention, which renders these algorithms susceptible to overfitting on artifactual EEG components. In this study, beamforming is employed to construct spatial filters that extract EEG sources originating within predefined regions of interest within the brain. In this way, neurophysiological knowledge on which brain regions are relevant for a certain experimental paradigm can be utilized to construct unsupervised spatial filters that are robust against artifactual EEG components. Beamforming is experimentally compared with CSP and Laplacian spatial filtering (LP) in a two-class motor-imagery paradigm. It is demonstrated that beamforming outperforms CSP and LP on noisy datasets, while CSP and beamforming perform almost equally well on datasets with few artifactual trials. It is concluded that beamforming constitutes an alternative method for SF that might be particularly useful for BCIs used in clinical settings, i.e., in an environment where artifact-free datasets are difficult to obtain.**

***Index Terms*—Beamforming, brain-computer interfaces, common spatial patterns, electroencephalography, motor imagery, spatial filtering.**

I. Introduction

NONINVASIVE brain–computer interfaces (BCIs) are devices that infer a subject's intention from noninvasive measurements of brain activity. BCIs thereby enable subjects to communicate without utilizing the peripheral nervous system. This is of particular interest to subjects with damage to the peripheral nervous system, e.g., caused by amyotrophic lateral sclerosis (ALS) or brain stem stroke, for which normal communication is impaired or even impossible. A general introduction to research on noninvasive BCIs is given in [24]. In principle, any noninvasive recording modality of brain activity, such as EEG, magnetoencephalography (MEG), or functional MRI (fMRI), can be used to construct a noninvasive BCI. Of these modalities, EEG is the most affordable and most widely available. Accordingly, the work presented in this paper focuses on EEG recordings. However, all results can be adapted to MEG recordings with relative ease.

One of the largest obstacles to constructing powerful BCIs based on EEG is the low signal-to-noise ratio (SNR) of EEG recordings. The components of the EEG providing information on the user's intention are usually heavily cloaked by ongoing background activity of the brain, hindering an effective inference of the user's intention. One commonly employed strategy to improve the SNR is linear spatial filtering (SF). Here, EEG measurements from multiple sites on the scalp are linearly combined in order to optimally attenuate EEG sources not providing information on the user's intention. Constructing good spatial filters, however, is a difficult problem since it is, in general, unknown which characteristics of the EEG provide maximum information on the user's intention, i.e., how the user's intention is encoded in the electric field of the brain. It is known, however, that subjects are capable of intentionally inducing changes in the power of spectral components of the electric field of the brain. For example, motor imagery of different limbs can be used to induce event-related synchronization/desynchronization (ERS/ERD) in those areas of the motor cortex representing the specific limbs (as reviewed in [19]). As first demonstrated in [20], this can be used to construct a noninvasive BCI. Most BCIs based on EEG are currently based on motor-imagery paradigms [13], which are also the type of paradigm used in the experimental evaluation of this paper. In this context, linear spatial filters are considered optimal if they maximally attenuate the variance of those EEG sources that are not modulated by motor imagery.

One of the most successful algorithms for SF in noninvasive BCIs based on motor imagery is the common spatial patterns (CSPs) algorithm, introduced to the BCI community in [21]. CSP is a supervised algorithm designed for two-class paradigms. It constructs linear spatial filters that maximize the ratio of class-conditional variances of extracted EEG sources. Excellent classification results have been reported using CSP, e.g., in one of the winning entries to the BCI competition 2003 [2]. Furthermore, there is evidence that CSP is optimal in terms of maximizing mutual information of extracted features and the subject's intention [9]. However, being a supervised algorithm, CSP suffers from overfitting phenomena [3]. Instead of extracting sources providing information on the subject's intention, CSP often focuses on artifactual components. This is due to the fact that the variance of artifactual EEG components often exceeds that of endogenous components of the brain. If a certain

Manuscript received April 30, 2008; revised August 19, 2008 and October 1, 2008. First published December 2, 2008; current version published May 6, 2009. *Asterisk indicates corresponding author.*

*M. Grosse-Wentrup was with the Institute of Automatic Control Engineering (LSR), Technische Universität München, Munich 80290, Germany. He is now with the Department of Empirical Inference, Max-Planck-Institute for Biological Cybernetics, Tübingen, 72076, Germany (e-mail: moritzgw@ieee.org).

C. Liefhold is with Max-Planck-Institute for Biological Cybernetics, Tübingen 72076, Germany (e-mail: christian.liefhold@mytum.de).

K. Gramann is with Swartz Center for Computational Neuroscience, University of California, San Diego, CA 92037 USA (e-mail: klaus@sccn.ucsd.edu).

M. Buss is with the Institute of Automatic Control Engineering (LSR), Technische Universität München, Munich 80290, Germany (e-mail: m.buss@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2008.2009768

type of artifact is more pronounced in the EEG training data of one class than in the EEG data of the other class, the ratio of class-conditional variances is maximized by extracting the artifactual EEG component that does not provide information on the subject's intention. Note that in the context of CSP, overfitting should not be understood as necessarily overfitting on the 0–1 loss function used for classification, but rather as overfitting on the CSP loss function of the ratio of class-conditional variances.

Another popular approach to SF in noninvasive BCIs is independent component analysis (ICA) (see [5] for a general introduction to ICA and [4] for an introduction to SF in BCIs by ICA). ICA computes spatial filters in an unsupervised manner by decomposing the observed EEG into statistically independent components (ICs). However, after computing the ICs, it is necessary to identify those that provide maximum information on the subject's intention. To the best of the authors' knowledge, this has only been demonstrated using labeled training data, which make SF by ICA susceptible to overfitting as well.

In general, supervised algorithms such as CSP and ICA perform well if the recorded EEG data are not contaminated by artifacts. For noisy datasets, supervised methods tend to focus on artifactual components, which often results in unsatisfactory classification results. While this is not of primary concern in research environments where experiments can be carried out with healthy subjects under optimal conditions, EEG data recorded from patients in clinical environments are usually heavily contaminated by artifacts, e.g., as caused by electric devices used for life support or monitoring purposes or by medical care during the recording session. It is hence desirable to develop algorithms for SF that perform well on noisy datasets, i.e., that are more robust against artifactual components of the EEG.

One way to render SF more robust against artifactual components is to focus on unsupervised methods that do not rely on labeled training data. In motor-imagery paradigms, it has been demonstrated that Laplacian spatial filtering (LP) substantially increases classification accuracy without being prone to overfitting [14]. Laplacian spatial filters, however, assign weights to each electrode in a rather *ad hoc* manner that cannot be regarded as optimal. Furthermore, in LP, only few electrodes are used, thereby discarding potentially useful information recorded at other locations on the scalp.

In this paper, a different approach to unsupervised SF is proposed. In many experimental paradigms, neurophysiological knowledge is available on which regions of the brain provide information on a subject's intention. For example, it is well known that in motor-imagery paradigms, EEG components originating in those areas of the motor cortex that represent the specific limbs provide information on the user's intention [10], [19]. In this study, linear spatial filters are presented that utilize this *a priori* knowledge by optimally attenuating the variance of all EEG sources not originating in chosen regions of interest (ROIs) within the brain. By choosing ROIs according to neurophysiological *a priori* knowledge for a given paradigm, it is possible to construct linear spatial filters that: 1) optimally attenuate EEG sources that do not provide information on the subject's intention and 2) are robust against artifactual EEG components due to their unsupervised nature.

In EEG/MEG analysis, spatial filters extracting sources from certain regions within the brain are commonly known as beamformers (reviewed in [8]). In fact, the beamforming approach presented here is similar to the MaxSNR beamformer well known in the area of array signal processing (cf., [23]). However, to the best of the authors' knowledge, this paper is the first to apply the concept of beamforming in the context of noninvasive BCIs. However, note that a preliminary version of this paper has been presented in [11].

This paper is organized as follows. In Section II-A, the notation used throughout this paper is introduced. The beamforming approach to linear SF is presented in Section II-B and the properties of the obtained beamformer are discussed in Section II-C. In Section III, experimental results from a two-class motor-imagery paradigm of ten healthy subjects are presented. Classification results obtained with beamforming, CSP, and LP are compared, and the feasibility of BCIs with real-time feedback based on beamforming is demonstrated. This paper concludes with a discussion of the results in Section IV.

II. Methods

A. Notation

Throughout this paper, vectors are denoted by bold letters and matrices by capital letters. Accordingly, $\boldsymbol{x}(t) \in \mathbb{R}^M$ refers to one sample of EEG data recorded at time t at M electrodes. If the time index is dropped, $\boldsymbol{x}$ is treated as an M-dimensional stationary random variable. A spatial filter is denoted by $\boldsymbol{w} \in \mathbb{R}^M$, and the spatially filtered EEG data by $y(t) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}(t) \in \mathbb{R}$. Spatial covariance matrices are denoted by $R_{(\cdot)}$.

B. SF by Beamforming

In this section, a spatial filter is derived that optimally attenuates the variance of EEG sources outside a predefined ROI. In general, it is desirable to completely eliminate EEG sources originating outside the ROI. However, this is not possible due to the ill-posed nature of the inverse problem of EEG. In EEG recordings, the continuous current distribution within the brain, which gives rise to the EEG, is mapped onto a finite number of measurement electrodes. This corresponds to a mapping from an infinite- to a finite-dimensional space. Accordingly, estimating EEG sources originating in a certain ROI constitutes an underdetermined problem. The best one can do is to find a spatial filter that, in some sense, optimally attenuates all sources outside the ROI. Since it is assumed here that only variance changes provide information on the subject's intention, optimal attenuation is defined as maximizing the ratio of variances of EEG sources originating inside and outside the ROI. In mathematical terms, the goal is to compute a spatially filtered EEG signal

$$y(t) = \boldsymbol{w}^{*\mathrm{T}}\boldsymbol{x}(t) \tag{1}$$

with

$$\boldsymbol{w}^* = \underset{\boldsymbol{w}\in\mathbb{R}^M}{\operatorname{argmax}} \left\{ \frac{\boldsymbol{w}^{\mathrm{T}} R_{\mathrm{ROI}} \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{T}} R_{\mathrm{Out}} \boldsymbol{w}} \right\} \tag{2}$$

and $R_{\mathrm{ROI/Out}} \in \mathbb{R}^{M\times M}$ the spatial covariance matrices of those components of the EEG originating within/outside the

ROI and measured at the M electrodes. Since (2) is in the form of the well-known Rayleigh quotient, solutions to (2) are given by the eigenvectors of the generalized eigenvalue problem

$$R_{\mathrm{ROI}}\boldsymbol{w} = \lambda R_{\mathrm{Out}}\boldsymbol{w}. \tag{3}$$

Since it further holds that for an eigenvalue λ^* with associated eigenvector $\boldsymbol{w}^*$

$$\lambda^* = \frac{\boldsymbol{w}^{*\mathrm{T}} R_{\mathrm{ROI}} \boldsymbol{w}^*}{\boldsymbol{w}^{*\mathrm{T}} R_{\mathrm{Out}} \boldsymbol{w}^*} \tag{4}$$

the eigenvector of (3) with the largest eigenvalue constitutes the desired beamformer.

It then remains to determine the covariance matrices $R_{\mathrm{ROI/Out}}$. These cannot be computed directly from measured data and thus have to be approximated. Toward this, first note that the EEG generated by the brain and measured at M locations on the scalp is given by [17]

$$\boldsymbol{x}(t) = \int_V L(\boldsymbol{r}, \boldsymbol{r}')P(\boldsymbol{r}', t)\mathrm{d}V(\boldsymbol{r}') \tag{5}$$

with V being the volume of the brain, $P : \mathbb{R}^3 \times \mathbb{R} \mapsto \mathbb{R}^3$ the tissue dipole moment (source strength) at position $\boldsymbol{r}'$ and time t in x-, y-, and z-direction, $\boldsymbol{r} \in \mathbb{R}^{3M}$ the vector describing the x-, y-, and z-position of the M sensors on the scalp, and $L : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^{M\times 3}$ the so-called leadfield equation, describing the projection strength of a source with dipole moment in x-, y-, and z-direction at position $\boldsymbol{r}'$ to the measured electric potentials at the sensor locations $\boldsymbol{r}$. Note that the leadfield equation incorporates all geometric and conductive properties of the head. Without loss of generality, it is assumed that $\boldsymbol{x}$ has zero mean. The integral in (5) can be split up into the contributions to the EEG from within and outside the ROI, resulting in

$$\begin{aligned}\boldsymbol{x}(t) &= \int_{\mathrm{ROI}} L(\boldsymbol{r}, \boldsymbol{r}')P(\boldsymbol{r}', t)\mathrm{d}V(\boldsymbol{r}') \\ &\quad + \int_{V\setminus\mathrm{ROI}} L(\boldsymbol{r}, \boldsymbol{r}')P(\boldsymbol{r}', t)\mathrm{d}V(\boldsymbol{r}') \\ &= \boldsymbol{x}_{\mathrm{ROI}}(t) + \boldsymbol{x}_{\mathrm{Out}}(t).\end{aligned} \tag{6}$$

Assuming stationarity of the EEG and uncorrelatedness of EEG sources within and outside the ROI, the covariance matrix of the EEG recordings is given by

$$R_{\boldsymbol{x}} = R_{\mathrm{ROI}} + R_{\mathrm{Out}}. \tag{7}$$

Inserting (7) into (3) then results in

$$R_{\mathrm{ROI}}\boldsymbol{w} = \tilde{\lambda} R_{\boldsymbol{x}}\boldsymbol{w} \tag{8}$$

with $\tilde{\lambda} = \lambda/(1+\lambda)$. Since $R_{\boldsymbol{x}}$ can be estimated from recorded EEG data, only R_{ROI} remains to be determined. This can be approached by first approximating the integral of the contribution of sources within the ROI to the measured EEG in (6) as

$$\boldsymbol{x}_{\mathrm{ROI}}(t) = \alpha \sum_{j=1}^{J} L(\boldsymbol{r}, \boldsymbol{r}'_j)P(\boldsymbol{r}'_j, t) \tag{9}$$

with $\boldsymbol{r}'_j, j = 1, \ldots, J$, being the locations of an equally spaced grid with J points within the ROI and α some numerical constant. The electric field at the M electrodes on the scalp due to sources within the ROI can thus be approximated as

$$\boldsymbol{x}_{\mathrm{ROI}}(t) = \alpha L\boldsymbol{p}(t) \tag{10}$$

with the leadfield matrix $L \in \mathbb{R}^{M\times 3J}$ describing the projection strength in x-, y-, and z-direction of the sources at the J grid points to the M electrodes, and $\boldsymbol{p}(t) \in \mathbb{R}^{3J}$ representing the dipole moments of the J sources. Since $\boldsymbol{x}$ has zero mean and the EEG is assumed to be stationary, the covariance matrix of $\boldsymbol{x}_{\mathrm{ROI}}$ can be written as

$$R_{\mathrm{ROI}} = \alpha^2 L R_{\boldsymbol{p}} L^{\mathrm{T}} \tag{11}$$

with $R_{\boldsymbol{p}}$ being the source covariance matrix of sources within the ROI. Inserting (11) into (8) and letting $\hat{\lambda} = \tilde{\lambda}/\alpha^2$, the desired spatial filter is finally obtained as the eigenvector with the largest eigenvalue of the generalized eigenvalue problem

$$L R_{\boldsymbol{p}} L^{\mathrm{T}}\boldsymbol{w} = \hat{\lambda} R_{\boldsymbol{x}}\boldsymbol{w}. \tag{12}$$

The leadfield matrix L describes the projection of sources within the ROI to the EEG electrodes, and thus, implicitly defines the ROI. It has to be computed using a suitable model of EEG volume conduction (reviewed in [1]). In this study, a four-shell spherical head model is utilized [22]. Furthermore, the covariance matrix of EEG sources within the ROI has to be specified. In absence of any prior knowledge, it is assumed that $R_{\boldsymbol{p}}$ equals the identity matrix. The eigenvector of (3) with the largest eigenvalue, which constitutes the desired beamformer $\boldsymbol{w}^*$, can then be computed with standard tools for numerical computation (e.g., with the command *eig* in MATLAB).

C. Beamformer Properties

In the derivation of the beamformer, several assumptions are made that warrant further discussion.

First, it is assumed that EEG sources within and outside the ROI are uncorrelated. This assumption is probably violated for sources outside but close to the ROI. Nevertheless, this is a useful assumption since it allows formulating the generalized eigenvalue problem in terms of the observed EEG covariance matrix. In this way, the beamformer can be adapted to recorded data. In principle, it is also possible to estimate the covariance matrix R_{Out} in the same manner as R_{ROI}, i.e., using a purely model-based approach, and compute the desired spatial filter directly from (3). However, if the beamformer is adapted to the recorded data, then the attenuation of sources focuses on regions of the brain outside the ROI that interfere most with sources inside the ROI. In this way, the beamformer is adapted to the subject- and task-specific current distribution within the brain.

Second, it is assumed that the EEG is stationary, implying that the covariance matrices of sources within and outside the ROI do not change over time. There is evidence, however, that EEG displays nonstationary behavior [18]. While this is neglected in the derivation of the beamformer, the nonstationary nature of EEG signals can be taken into account by updating the EEG covariance matrix $R_{\boldsymbol{x}}$ in certain intervals and computing a new spatial filter using the updated covariance matrix. In this way,

TABLE I
PERCENTAGE OF TRIALS CONTAMINATED BY ARTIFACTS

S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
23.7%	55.3%	5.0%	6.3%	23.3%	14.3%	14.6%	11.33%	15.0%	74.0%

a quasi-static beamformer can be realized. The performance of such an update scheme is investigated in Section III.

Third, the source covariance matrix $R_{\boldsymbol{p}}$ is assumed to equal the identity matrix. This implies that all EEG sources within the ROI are uncorrelated and of equal strength. This surely constitutes an unrealistic assumption. However, in absence of any knowledge on the actual source covariance matrix for a given dataset, this is the most simple prior. It should be noted, however, that the absolute strength of sources within the ROI is irrelevant since any scaling of $R_{\boldsymbol{p}}$ is absorbed into the eigenvalues in (12), and thus, has no influence on the obtained beamformer.

Finally, it should be emphasized that any model of EEG volume conduction can be used to compute the leadfield matrix in (12). For reasons of simplicity, only a four-shell spherical headmodel is considered in this study. It can be expected that more realistic models, such as boundary element or finite-element models (BEM/FEM) [1], also lead to more accurate beamformers.

III. EXPERIMENTAL RESULTS

In this section, SF by beamforming is compared with CSP and LP on experimental data from a two-class motor-imagery paradigm. Two different beamforming schemes are investigated, termed *static* and *block-adaptive* beamforming (SBF-BBF). CSP and LP are chosen for comparison with beamforming due to their excellent performance in motor-imagery paradigms and popularity in the BCI community. Furthermore, preliminary results from a study with real-time feedback are presented.

A. Experimental Paradigm

The experimental paradigm adopted in this study was as follows. Each subject was seated in a dimly lit and shielded room, approximately 2 m in front of a silver screen. A trial started with the central display of a white fixation cross. After 3 s, a white arrow was superimposed on the fixation cross, either pointing to the left or the right. Subjects were instructed to perform haptic motor imagery of the left or the right hand during display of the arrow, as indicated by the direction of the arrow. After another 7 s, the arrow was removed, indicating the end of the trial and start of the next trial. While subjects were explicitly instructed to perform haptic motor imagery with the specified hand, i.e., to imagine feeling instead of visualizing how their hands moved, the exact choice of which type of imaginary movement, i.e., moving the fingers up and down, gripping an object, etc., was left unspecified. A total of 150 trials per condition were carried out by each subject, with trials presented in pseudorandomized order. Note that in the employed experimental paradigm, subjects were not free to choose when to initiate a certain motor imagination. Hence, this study is restricted to synchronous BCIs.

B. Experimental Data

Ten healthy subjects (S1–S10) participated in the experimental evaluation. Of these, two were females, eight were right handed, and their average age was 25.6 years with a standard deviation of 2.5 years. Subject S3 had already participated twice in a BCI experiment, while all other subjects were naive to BCIs. EEG was recorded at $M = 128$ electrodes placed according to the extended 10–20 system. Data were recorded at 500 Hz with electrode Cz as reference. Four BrainAmp amplifiers were used for this purpose, using a temporal analog high-pass filter with a time constant of 10 s. The data were re-referenced to common average reference offline. Electrode impedances were below 10 kΩ for all electrodes and subjects at the beginning of each recording session. No trials were rejected and no artifact correction was performed. For each subject, the locations of the 128 electrodes were measured in three dimensions using a Zebris ultrasound tracking system and stored for further offline analysis.

After the recording sessions, the recorded EEG of each subject was visually inspected for eye blinks, movement artifacts, and artifacts caused by interference of electric devices by an experienced EEG user. This was not done to reject artifactual trials, but rather to assess the percentage of trials contaminated by artifacts. The percentage of trials for each subject determined to contain substantial amounts of artifacts during the actual motor-imagery phase are summarized in Table I. Note that while manual labeling of artifactual trials by an experienced EEG user is a subjective measure, we believe this to constitute a more sensitive measure than an automatic identification of artifactual trials. The recorded EEG data and trials marked as artifactual can be made available upon request. As can be seen from Table I, subjects S3 and S4 displayed very few artifactual trials (below 10%), subjects S6–S9 displayed a moderate amount of artifactual trials (between 10% and 20%), and subjects S1, S2, S5, and S10 showed a large amount of artifactual trials (above 20% and up to 74%). Subsequently, subjects S3 and S4 will be referred to as clean subjects, subjects S6–S9 as moderate subjects, and subjects S1, S2, S5, and S10 as noisy subjects.

C. Evaluation Procedure

To evaluate the performance of the different algorithms for SF as a function of the amount of available training data, a bootstrapping procedure was employed. For each subject, n trials from each class were drawn randomly from the recorded data and used as the training set, while the remaining trials served as the test set. Then, SF, feature computation, and classification were performed as described next. For each size of the training set, this procedure was repeated ten times in order to obtain a sensible estimate of the classification accuracy. The size of the training set was varied from $n = 10$ to $n = 100$ trials per class in steps of ten trials.

Furthermore, two different time windows were investigated for all spatial filter and feature computations in order to assess the influence of different trial length on performance. The first time window, subsequently termed the long time window, ranged from 3.5–10 s within each trial, i.e., 500 ms after presentation of the instruction until the end of the trial. The second time window, subsequently termed the short time window, ranged from 3.5 to 6 s within each trial. Note that the length of the short time window corresponds to what is suggested for CSP in [3]. Time windows were chosen to start 500 ms after visual presentation of the motor-imagery instruction because subjects require several hundred milliseconds to initiate motor imagery of the specified hand. This is reflected in ERD/ERS onset roughly 500 ms after presentation of the instruction [19].

1) SF by CSP: SF by CSP was carried out with the parameters proposed in [3]. First, the recorded EEG data were bandpass-filtered between 7 and 30 Hz using a sixth-order Butterworth filter. Then, class-conditional covariance matrices were computed using data in the specified time window of the training set only. CSP spatial filters were computed by solving the associated generalized eigenvalue problem with diagonal loading to increase numerical stability. The obtained spatial filters with the three largest/smallest eigenvalues were then combined in the SF matrix $W_{\mathrm{CSP}} \in \mathbb{R}^{M\times 6}$.

2) LP: For LP, the large Laplacian spatial filter as described in [14] was employed. Specifically, electrodes C3/C4, situated over the left/right motor cortex, were chosen as the filter centers, and the four second closest electrodes to C3/C4 were used to compute the surface Laplacian, thereby forming the SF matrix $W_{\mathrm{LP}} \in \mathbb{R}^{M\times 2}$.

3) SF by SBF: In SBF, beamformers are computed once using a set of (unlabeled) training data. The beamformers are then applied to new data, i.e., the test set, without further update. Static beamformers were computed here by first high-pass filtering the recorded EEG with 0.5 Hz cutoff frequency using a sixth-order Butterworth filter in order to remove baseline drifts. The EEG covariance matrix $R_{\boldsymbol{x}}$ was then computed using the temporally filtered data of all trials in the training set and the specified time window. In order to direct two beamformers at the left and right motor cortex, respectively, two spherical ROIs of 1 cm radius located 1.9 cm radially below electrodes C3 and C4 were chosen. The associated leadfield matrices L_{C3} and L_{C4} were computed by placing radially oriented dipoles onto an equidistant grid of 2 mm grid distance within each of the ROIs, and calculating the projection strength to each of the $M = 128$ electrodes using a four-shell spherical head model [22]. Note that for each subject, the measured electrode locations were radially projected onto the outermost shell of the employed head model. Assuming a unit source covariance matrix $R_{\boldsymbol{p}}$, the two desired beamformers were then determined by computing the eigenvector with the largest eigenvalue of (12) for each of the two leadfield matrices L_{C3} and L_{C4}. Again, diagonal loading was used in the eigenvector computation to improve numerical stability. The two eigenvectors were finally combined to form the SF matrix $W_{\mathrm{SBF}} \in \mathbb{R}^{M\times 2}$.

4) SF by BBF: In BBF, beamformers are not computed from a fixed set of EEG data, but rather recomputed for each trial. In this way, beamformers can be adapted to nonstationarities in the recorded EEG. Here, the parameters employed for BBF were chosen identical to those used in SBF, as described earlier. For each trial, a separate EEG covariance matrix was then estimated using the (unlabeled) data from the specified time window, and a trial-specific SF matrix $W_{\mathrm{BBF}} \in \mathbb{R}^{M\times 2}$ was computed as described earlier.

5) Feature Computation and Classification: For all SF methods described earlier, the same feature computation and classification procedure was employed.

First, the SF matrix $W_{(\cdot)}$ was used to compute the spatially filtered EEG signal $\boldsymbol{y}(t) = W_{(\cdot)}^{\mathrm{T}}\boldsymbol{x}(t)$. Note that here $\boldsymbol{x}(t)$ refers to the raw EEG data, i.e., the original EEG recordings without previous temporal filtering. For each trial, a feature vector was then computed by first-bandpass filtering each component of the spatially filtered EEG signal in 20 frequency bands of 2 Hz width ranging from 1 to 41 Hz (again using sixth-order Butterworth filters), and afterwards, computing the log-bandpower in each frequency band using the specified time window. This resulted in a feature vector of 120 dimensions for CSP, and of 40 dimensions for Laplacian filtering, SBF, and BBF.

For actual classification, logistic regression with l_1 regularization, as described in [12], was employed. This linear classifier was chosen for two reasons. First, nonlinear classifiers do not significantly improve classification accuracy in noninvasive BCIs based on bandpower features while needlessly increasing complexity [7], [15]. Second, only some frequency bands provide information on the user's intention in motor-imagery paradigms, and these frequency bands vary across subjects [19]. It can thus be expected that most dimensions of the feature vector are irrelevant, but it is unknown which ones are relevant for a certain subject. For this class of classification problems, i.e., a high-dimensional feature space with mostly irrelevant features, it is proved in [16] that logistic regression with l_1 regularization possesses a sample complexity that only grows logarithmically in the number of irrelevant features. Rotationally invariant classifiers, such as support vector machines, have a worst case sample complexity that grows linearly in the number of irrelevant features. Hence, for this class of problems, logistic regression with l_1 regularization can be expected to display a faster convergence (in terms of the required amount of training data) to its minimum error than other state-of-the-art classification algorithms. Note that logistic regression using l_1 regularization leads to sparse weight vectors, and thus, can also be understood as a feature selection procedure. For each subject, training set, and algorithm used for SF, the l_1 regularization constant was determined by fivefold cross-validation on the training set. This regularization constant was then used to train another classifier using all features from the training set, and the resulting logistic regression model was used to classify the trials in the test set.

D. Results

The obtained classification results are shown in Table II and Fig. 1 for the long time window, and in Table III and Fig. 2 for the short time window, both times as a function of the number

TABLE II
MEAN CLASSIFICATION ACCURACY FOR THE *Long* TIME WINDOW IN PERCENT AS A FUNCTION OF THE NUMBER OF TRAINING TRIALS PER CONDITION FOR SF BY CSPS, SBF, BBF, AND LP

Subject	SF	Number of training trials per condition										Mean
		10	20	30	40	50	60	70	80	90	100	
S1	CSP	52.4	56.0	55.8	55.7	55.5	57.7	60.4	58.6	58.2	59.5	**57.0**
	SBF	52.2	53.4	55.5	57.1	57.7	57.0	58.8	59.6	58.8	59.5	**57.0**
	BBF	50.3	52.0	53.9	52.7	54.2	55.7	55.2	57.5	58.8	57.6	54.8
	LP	52.1	53.3	52.1	53.4	55.6	55.1	52.9	57.2	55.0	56.7	54.3
S2	CSP	59.7	64.2	67.6	64.6	66.3	62.9	70.3	65.1	69.8	71.6	66.2
	SBF	83.1	88.6	90.7	92.5	92.2	93.7	91.1	93.6	94.2	95.2	91.5
	BBF	82.4	85.7	89.0	89.9	89.8	90.6	92.4	92.4	93.0	93.4	89.9
	LP	83.7	89.4	92.0	92.8	93.7	93.4	93.5	93.4	93.2	94.1	**91.9**
S3	CSP	78.6	92.3	96.0	96.7	98.1	97.9	98.1	98.9	98.2	98.9	95.4
	SBF	94.7	95.6	96.4	96.5	96.5	96.5	96.4	97.6	96.9	97.2	**96.4**
	BBF	95.6	94.9	96.4	96.2	95.1	96.8	96.9	97.1	97.4	97.0	**96.4**
	LP	89.2	92.1	93.2	91.7	93.1	92.7	93.1	94.5	93.6	94.8	92.8
S4	CSP	85.2	93.3	97.3	97.4	97.9	99.1	97.9	98.8	98.8	98.3	96.4
	SBF	95.0	96.3	96.0	96.0	96.8	97.2	96.1	97.1	97.8	97.2	96.6
	BBF	94.1	97.0	97.5	96.5	98.4	98.7	97.7	97.8	98.6	98.3	**97.5**
	LP	90.2	94.0	94.8	94.4	96.2	96.3	95.7	94.6	96.0	96.0	94.8
S5	CSP	76.7	75.5	82.3	78.3	79.8	86.5	88.4	89.8	89.9	88.3	83.6
	SBF	85.1	90.1	92.0	91.6	93.2	92.2	92.4	93.6	92.0	92.2	**91.4**
	BBF	82.4	84.8	87.5	88.6	89.5	89.3	89.7	90.3	88.0	89.8	88.0
	LP	81.4	87.3	86.8	89.4	90.4	90.2	90.2	90.1	89.2	88.3	88.3
S6	CSP	70.0	81.6	89.4	90.2	92.0	91.7	92.9	91.9	93.0	91.7	**88.4**
	SBF	75.1	85.0	86.2	89.1	89.4	89.7	89.9	90.0	89.5	88.2	87.2
	BBF	64.6	74.2	79.4	81.2	80.4	82.3	81.5	81.6	80.8	82.1	78.8
	LP	71.4	78.6	79.2	81.8	83.8	81.3	82.4	81.8	85.2	84.8	81.0
S7	CSP	50.1	51.2	51.3	50.5	51.7	49.6	49.8	51.5	50.0	49.1	50.5
	SBF	51.2	53.1	53.9	55.0	57.0	57.1	56.7	59.1	57.7	59.2	**56.0**
	BBF	51.6	54.1	53.9	54.5	56.2	56.5	57.9	56.4	57.2	59.1	55.7
	LP	50.0	53.7	54.3	54.4	56.5	52.8	55.9	54.6	57.0	55.2	54.4
S8	CSP	53.9	52.8	55.9	60.5	59.7	60.0	56.8	54.9	51.8	57.0	56.3
	SBF	56.9	69.2	72.6	76.0	75.4	76.9	76.1	78.1	76.1	75.5	73.3
	BBF	63.7	70.8	73.5	74.8	75.8	77.0	77.3	77.0	79.3	78.8	**74.8**
	LP	57.9	65.4	65.2	68.5	70.5	69.1	69.6	71.9	70.4	70.5	67.9
S9	CSP	57.0	61.7	66.9	69.8	69.4	63.4	63.4	67.1	72.4	67.1	65.8
	SBF	56.0	59.8	61.9	63.7	62.4	65.5	64.8	67.2	64.4	65.3	63.1
	BBF	57.5	62.6	63.5	66.5	68.4	69.2	68.7	68.4	67.4	68.0	66.0
	LP	60.9	64.2	66.5	67.2	68.5	67.6	67.9	70.9	69.1	68.8	**67.2**
S10	CSP	57.0	62.7	67.7	73.5	74.6	74.2	75.2	77.1	79.4	79.4	72.1
	SBF	77.8	79.7	84.2	84.6	87.1	86.7	87.7	87.7	86.4	88.1	**85.0**
	BBF	74.4	80.6	83.0	84.0	86.7	85.5	86.8	87.3	85.2	89.0	84.2
	LP	55.8	63.0	67.8	69.4	70.2	71.2	71.2	70.9	71.2	73.5	68.4

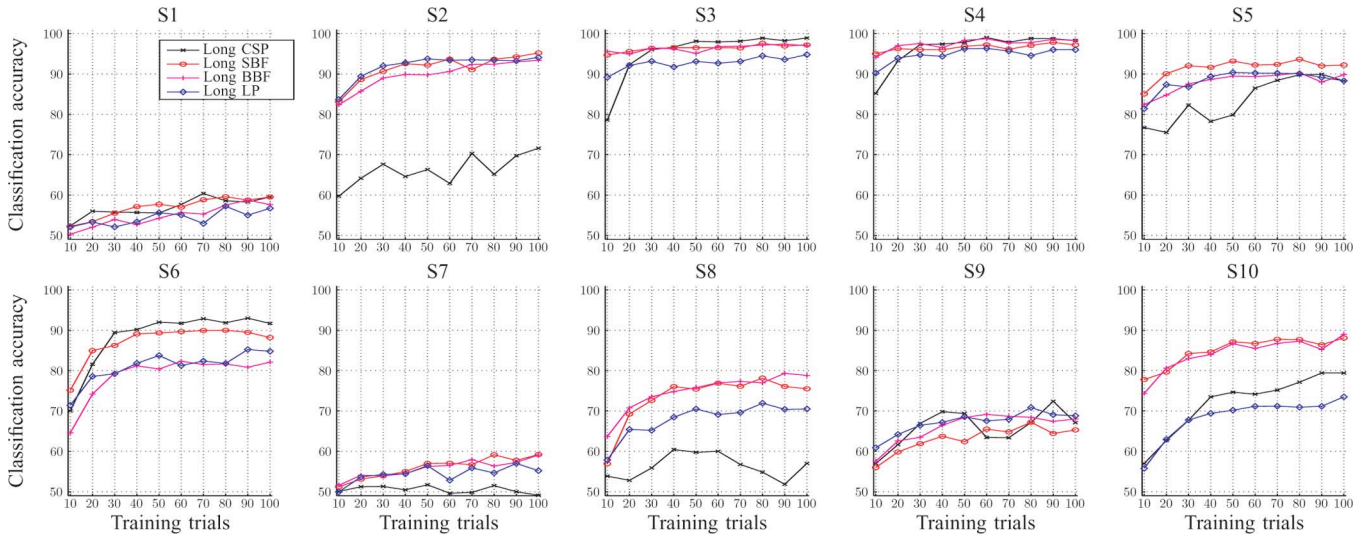


Fig. 1. Classification accuracies in percent for subjects S1–S10 using the *long* time window as a function of the number of training trials per condition for SF by CSPs, SBF, BBF, and LP.

of training trials per condition. The mean performance of all subjects across all training set sizes is shown in Table IV.

In general, the obtained classification results differ substantially between subjects. Subjects S3 and S4 achieve classification accuracies close to 100% for all SF methods, while subjects S1 and S7 do not perform substantially above chance level for any type of SF. The performance of the remaining six subjects ranges from mediocre to rather well, with substantial differences due to the algorithm used for SF. It is noteworthy that the capability of subjects to operate a BCI appears not to be

TABLE III
MEAN CLASSIFICATION ACCURACY FOR THE *Short* TIME WINDOW IN PERCENT AS A FUNCTION OF THE NUMBER OF TRAINING TRIALS PER CONDITION FOR SF BY CSPS, SBF, BBF, AND LP

Subject	SF	Number of training trials per condition										Mean
		10	20	30	40	50	60	70	80	90	100	
S1	CSP	50.8	50.6	54.4	52.3	52.5	54.2	50.8	55.7	53.8	54.3	52.9
	SBF	50.0	53.8	51.0	53.4	54.8	53.8	53.3	54.6	54.2	54.3	53.3
	BBF	51.0	50.1	51.7	51.8	54.5	52.4	53.7	56.2	52.2	54.6	52.8
	LP	50.9	54.1	54.4	56.1	55.4	57.3	54.9	58.3	57.0	55.7	**55.4**
S2	CSP	63.1	69.1	74.8	76.6	79.0	80.5	86.1	85.4	86.6	88.7	79.0
	SBF	84.5	90.5	89.2	92.1	90.9	92.0	91.6	91.6	93.8	93.7	91.0
	BBF	82.0	85.0	87.3	87.9	88.9	89.4	90.0	90.8	91.8	91.6	88.5
	LP	88.5	90.5	92.4	92.5	92.4	93.2	92.1	92.9	94.8	94.0	**92.3**
S3	CSP	63.4	76.2	86.0	89.2	92.1	93.7	94.6	94.8	95.6	95.4	88.1
	SBF	86.7	90.5	92.6	93.6	93.9	94.3	94.4	94.8	94.6	94.4	**93.0**
	BBF	88.4	90.0	93.3	93.1	93.5	93.2	93.4	94.5	94.2	93.9	92.7
	LP	82.3	83.0	85.8	87.6	89.6	88.8	89.4	91.9	89.6	88.2	87.6
S4	CSP	68.4	90.3	91.0	94.1	95.8	96.1	95.8	96.1	96.7	97.4	92.2
	SBF	86.6	91.8	91.9	93.0	93.9	94.3	94.1	92.7	94.1	94.1	**92.6**
	BBF	84.3	91.6	92.2	92.8	93.2	93.6	93.4	93.2	94.3	94.6	92.3
	LP	86.1	90.4	91.2	91.4	91.3	91.8	92.0	90.6	91.7	92.2	90.9
S5	CSP	65.6	72.2	76.5	71.8	78.0	81.4	81.8	83.4	87.1	81.6	77.9
	SBF	75.2	82.8	86.7	86.5	86.5	88.9	89.0	89.3	89.4	88.9	**86.3**
	BBF	72.6	81.5	81.7	84.0	86.0	86.1	87.1	87.1	84.7	86.9	83.8
	LP	73.1	81.8	83.2	83.0	83.3	83.1	83.9	84.9	82.2	82.9	82.1
S6	CSP	62.7	64.5	76.9	79.6	80.7	83.2	85.8	85.3	86.8	85.1	79.0
	SBF	71.2	78.8	80.7	82.1	83.3	83.3	82.3	83.5	83.7	83.8	**81.3**
	BBF	59.5	72.0	74.8	73.6	80.7	79.1	80.6	80.6	82.8	82.4	76.6
	LP	63.3	70.5	74.0	75.3	74.8	75.3	74.5	75.1	76.5	74.0	73.3
S7	CSP	49.6	51.2	50.9	50.2	50.5	53.0	49.6	50.8	52.0	49.8	50.8
	SBF	50.9	51.5	51.2	54.1	53.5	54.7	51.6	52.8	53.2	53.5	52.7
	BBF	51.1	51.5	53.2	53.4	51.2	53.7	54.1	52.8	54.6	54.0	53.0
	LP	50.7	52.7	52.6	54.4	57.1	57.1	57.2	56.9	56.0	56.1	**55.1**
S8	CSP	53.0	53.0	58.2	61.0	58.8	60.9	57.2	57.1	58.1	63.5	58.1
	SBF	62.0	72.1	70.7	72.5	75.3	75.6	74.4	77.3	76.9	76.9	**73.4**
	BBF	58.9	63.2	65.5	67.0	69.0	71.6	68.0	69.9	70.5	69.7	67.3
	LP	54.4	59.3	59.6	59.6	61.7	62.1	61.8	62.9	63.5	63.7	60.9
S9	CSP	55.3	61.4	61.8	68.5	72.1	70.9	73.6	74.5	75.8	76.6	**69.0**
	SBF	59.6	64.2	62.8	65.8	67.8	68.4	67.5	70.3	69.4	69.9	66.6
	BBF	60.5	63.3	64.7	65.9	68.2	68.3	67.1	67.4	69.3	68.3	66.3
	LP	58.1	62.4	66.2	68.1	68.1	69.9	68.0	70.6	70.8	70.6	67.3
S10	CSP	54.7	61.5	60.2	68.9	69.8	69.1	72.5	73.4	76.4	76.0	68.2
	SBF	67.7	73.0	79.3	77.2	77.3	79.1	81.2	81.3	80.3	81.4	**77.8**
	BBF	64.4	72.9	77.2	78.2	79.5	80.2	79.9	80.1	79.1	81.5	77.3
	LP	56.3	61.8	63.1	64.4	66.0	66.7	64.3	65.8	67.6	68.1	64.4

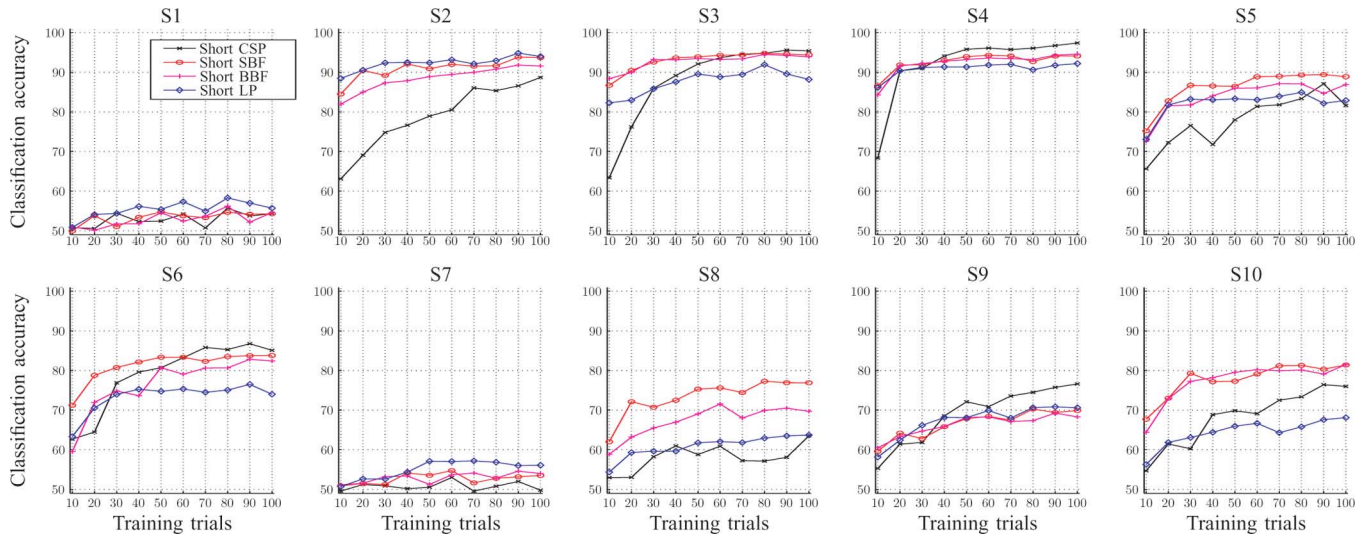


Fig. 2. Classification accuracies in percent for subjects S1–S10 using the *short* time window as a function of the number of training trials per condition for SF by CSPs, SBF, BBF, and LP.

determined by the percentage of artifactual trials. While subjects S3 and S4 with the lowest percentage of artifactual trials (cf., Table I) also perform best, subject S2 achieves very high classification accuracies using unsupervised methods in spite of more than 50% of all trials containing substantial amounts of artifacts. Conversely, subject S7 does not achieve classification accuracies substantially above chance in spite of only a moderate amount of artifactual trials (14.6%). Regarding the different time windows used in the computation of the spatial filters and the classification procedure, it is noteworthy that, on average, all

TABLE IV
MEAN CLASSIFICATION ACCURACY IN PERCENT FOR DIFFERENT SUBJECT TYPES (USING THE LONG /SHORT TIME WINDOW) ACROSS ALL TRAINING SET SIZES

Spatial filter	Subjects			Mean across subjects
	Noisy	Moderate	Clean	
CSP (long/short)	69.7 / 69.5	65.3 / 64.2	95.9 / 90.1	72.3 / 71.5
SBF (long/short)	**81.2** / 77.1	**69.9** / 68.5	96.5 / 92.8	**79.8** / 76.8
BBF (long/short)	79.2 / 75.6	68.9 / 65.8	**96.9** / 92.5	78.6 / 75.1
LP (long/short)	75.8 / 73.6	67.6 / 64.1	93.8 / 89.3	76.1 / 72.9

methods for SF perform better when using the long time window (cf., Table IV). However, note that substantial subject-specific differences can be observed for SF by CSP. Specifically, the short time window performs better than the long time window on subjects S2 and S9 when using CSP. On average, however, this is counterbalanced by superior performance of the other subjects using the long time window. Hence, if not stated differently, classification results will subsequently refer to the long time window.

Regarding a comparison of the algorithms used for SF, different algorithms perform best for different subjects. Averaging across all training set sizes, best performance across both time windows is achieved for three subjects by using CSP (S1, S6, and S9), for four subjects by using SBF (S3, S5, S7, and S10), for two subjects by using BBF (S4 and S8), and for one subject by using LP (S2). If only the maximum training set size is considered, five subjects perform best when using CSP (S1, S3, S4, S6, and S9), three subjects perform best by using SBF (S2, S5, and S7), and two subjects perform best by using BBF (S8 and S10). However, it should be noted that differences due to different SF within subjects range from minor (e.g., in subjects S3 and S4) to quite substantial (e.g., in subject S10).

Regarding the performance of SF for different types of subjects (cf., Section III-B and Table I), considerable differences can be observed (see Table IV). For noisy subjects, SBF outperforms CSP by, on average, 11.5% and LP by, on average, 5.5%. SBF and BBF perform similar with a slight advantage for SBF of 2%. This observation is also valid but less pronounced for moderate subjects, for which SBF outperforms CSP and LP by 5.5% and 2.2%, respectively. For clean subjects, no substantial differences between the different algorithms used for SF can be observed, i.e., differences are below 2%. If only the maximum training set size is considered, SF by CSP slightly outperforms the other methods on the clean subjects, but differences again remain below 2%. Finally, averaged across all subjects and training set sizes, SBF outperforms LP by 3.6%, which, in turn, outperforms CSP by 3%. However, note that the average performance across all subjects is biased by a larger number of noisy than clean subjects.

In summary, the performance of different spatial filters is strongly subject-specific and varies with the percentage of trials containing substantial amounts of artifacts. While CSP performs slightly better than the other approaches for virtually artifact-free EEG recordings, beamforming performs substantially better than CSP and LP on EEG recordings heavily contaminated by artifacts. On average, no substantial differences can be observed between SBF and BBF.

E. Spatial Filters and Spectral Band Weighting

Fig. 3 illustrates how the subject-specific performance of the SF algorithms is reflected in the weights assigned to each spectral band in the classification procedure.

As can be seen in this figure, the classification procedure generally concentrates on the μ- and β-band (roughly around 12 Hz and from 25 to 35 Hz, respectively), which is in agreement with previous reports on motor imagery [19]. However, there are spatial filter-specific differences. While for SF by CSP, the classification procedure assigns strongest weights to the μ-band, both beamforming and LP appear to favor the β-band. This observation is particularly pronounced for subject S3. However, note that in spite of a focus on different spectral bands, all spatial filters achieve excellent classification results for subject S3. As of now, we cannot provide an explanation for this observation.

Typical spatial filters focusing on the left motor cortex computed for subjects S3, S6, and S8 (using the long time window) are shown in Fig. 4. These subjects are chosen since S3 shows excellent performance for all spatial filters, S6 performs best when using CSP, and S8 performs best when using beamforming. Interestingly, spatial filters computed for subject S3 differ notably in spite of similar classification performance. While a typical spatial filter obtained by CSP shows a dipolar pattern with an additional focus on the contralateral hemisphere, spatial filters obtained by beamforming display a center-surround pattern that is similar to the Laplacian spatial filter. This is in agreement with the previous observation that in subject S3, beamforming and LP favor similar spectral weighting in the β-band, while CSP appears to lead to a focus on the μ-band (cf., Fig. 3). However, note that spatial filters obtained by beamforming appear more complex than the corresponding Laplacian spatial filter, which is also reflected in a higher classification accuracy. Further note that for subject S3, all spatial filters focus on an area directly underneath electrode C3. This is different for subject S6, for which CSP has a slightly more central and parietal focus than the other spatial filters. This illustrates an advantage of the supervised CSP approach, i.e., the capability to automatically determine the region of the brain most relevant for inferring the subject's intention. In subject S6, beamforming and LP probably slightly miss the most relevant ROI, resulting in a slight decrease in classification accuracy in comparison to CSP. However, the last row of Fig. 4, showing spatial filters computed for subject S8, illustrates that in some subjects, CSP fails to construct sensible spatial filters. Here, CSP does not compute spatial filters with a strong focus, resulting in a classification accuracy only slightly above chance. In comparison, both beamforming approaches compute spatial filters with a strong focus on and close to electrode C3 that appear similar in shape to the spatial filter computed by CSP for subject S3. Accordingly, the beamforming approaches achieve a classification accuracy close to 80% in subject S8. Finally, note that subject S8 displayed only a moderate amount of artifactual trials, indicating that contamination by artifacts is not the only cause for the failure of CSP to compute sensible spatial filters in some subjects.

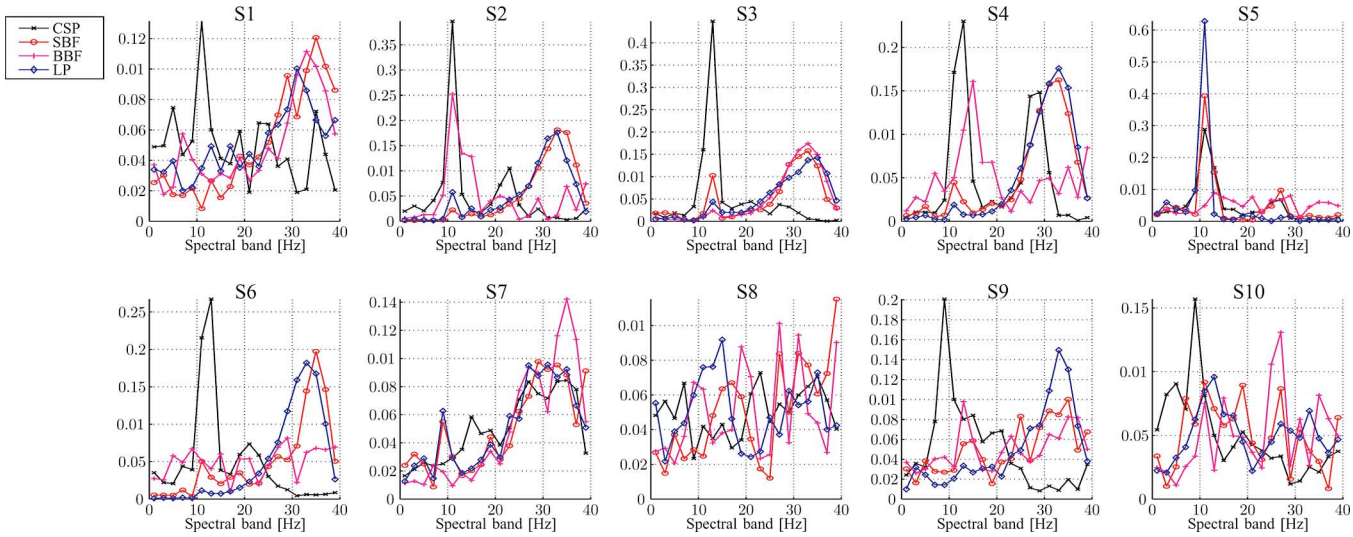


Fig. 3. Normalized mean absolute weights of different spectral bands across all training set sizes and both time windows as determined by the l_1-regularized logistic regression classifier for CSPs, SBF, BBF, and LP.

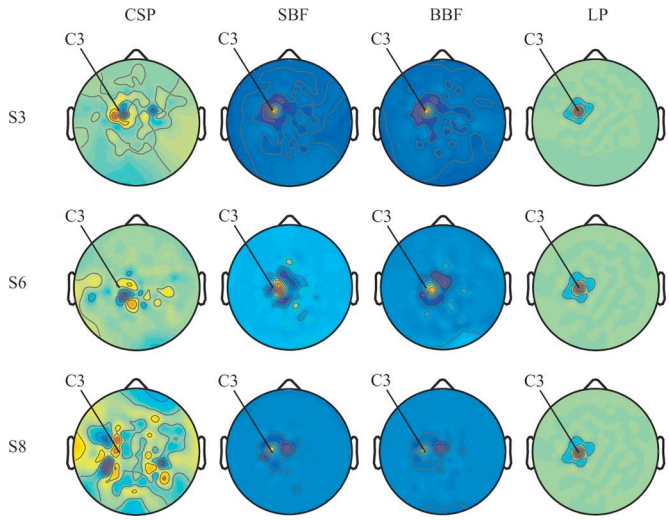


Fig. 4. Set of typical spatial filters focusing on left motor cortex obtained by CSPs, SBF, BBF, and LP for training set size of 100 trials per condition and the long time window. Plotted with [6]. Colors indicate normalized electrode weights.

F. Beamforming With Real-Time Feedback

To establish the feasibility of beamforming for BCIs with real-time feedback, the experimental setup was adapted in the following way. First, a certain number of training trials were recorded using the same experimental paradigm as described in Section III-A. This training dataset was then used to compute two static beamformers and train a logistic regression classifier with l_1 regularization, as described earlier. After training, real-time feedback was provided to the BCI user. This was carried out by sending the recorded EEG data via transmission control protocol (TCP)/IP to MATLAB/Simulink running at 500 Hz. The two static beamformers were then applied to every new data sample, and the resulting two extracted EEG components were bandpass-filtered as in the offline evaluation procedure. The variances of the temporally and spatially filtered time series within a trial were then calculated recursively at every sample step (i.e., not using a sliding window but in an accumulative manner), and the current estimate was fed into the previously trained logistic regression model. The output of the model at each sample point, ranging from 0 to 1, was then presented to the subject by drawing a white filled square on the screen. The output of the model was linearly mapped to the horizontal position of the square, with an output of 0 mapped to the left border and an output of 1 mapped to the right border of the screen. The horizontal position of the square thus informed the BCI user of the certainty of the classifier about her/his intention (with the left/right border of the screen indicating 100% certainty of an imaginary movement of the left/right hand). To further motivate the subject, two white boxes were drawn at the left and right borders of the screen into which the subject had to move the white square. Also, the color of the centrally displayed arrow was set to green or red, depending on whether the output of the classifier lead to a correct decision or an error. Each trial ended after a preset time, or if a certain threshold of certainty of the classifier was achieved. This threshold criterion was only checked after a certain minimum time into each trial to ensure sensible estimates of the variances of the EEG components. Finally, each trial began with a pause of 3 s.

Due to the excellent performance in the offline experiment, subject S4 was asked to perform again in the online experiment. Twenty five trials per condition were recorded as training data, corresponding to a training time of 8 min and 20 s. For the online experiment, the regularization constant used to train the logistic regression model was not determined by cross validation, but rather heuristically tuned on the offline datasets to a generic value found to work well across subjects. Then, five blocks of 20 trials per condition were carried out with feedback provided, with a break of approximately 2 min between each block. The obtained classification results are shown in Table V, along with the minimum and maximum trial lengths and the thresholds for

TABLE V
RESULTS OF THE ONLINE EXPERIMENT FOR SUBJECT S4

Block #	Min trial length	Max trial length	Thresholds	Classification accuracy
1	9.99 s	10 s	[0.1 0.9]	92.5%
2	9.99 s	10 s	[0.1 0.9]	87.5%
3	6 s	10 s	[0.1 0.9]	87.5%
4	6 s	30 s	[0.1 0.9]	92.5%
5	6 s	30 s	[0.1 0.9]	90.0%

termination of a trial. The mean classification accuracy across all blocks was 90.0%, which is within the range expected due to the classification accuracy obtained by subject S4 in the offline experiment using the SBF approach with the same amount of training data (cf., Table II). A video of this experiment can be watched at http://www.lsr.ei.tum.de/research/videos/biomedical-engineering/one-dimensional-cursor-control-by-a-non-invasive-brain-computer-interface/.

IV. DISCUSSION

In this paper, an alternative to supervised algorithms for SF in the context of noninvasive BCIs was presented. Based on the principle of beamforming, spatial filters were constructed that extract EEG sources originating within specific ROIs within the brain. In this way, neurophysiological *a priori* knowledge can be utilized to optimally attenuate EEG sources not providing information on a subject's intention for a given experimental paradigm. The main advantage of beamforming is its unsupervised nature, rendering it robust against artifactual EEG components.

The proposed algorithm was experimentally validated in a two-class motor-imagery paradigm, and it was shown that SF by beamforming substantially outperforms SF by CSP if the recorded EEG is strongly contaminated by artifactual components. On virtually artifact-free datasets, CSP performed slightly better than beamforming. Furthermore, beamforming could be shown to consistently outperform LP. Finally, the feasibility of constructing BCIs with real-time feedback using beamformers was demonstrated.

It should be noted that in this study, only a two-class motor-imagery paradigm has been considered. In principle, SF via beamforming can be extended easily to multiclass paradigms. For example, in motor-imagery paradigms using multiple body parts, several beamformers can be constructed, with the ROIs focused on those regions of the motor cortex representing the specific parts of the body. However, it remains to be experimentally established if beamforming also displays the advantageous properties observed here if it is applied to ROIs buried deep within the cortex, e.g., if motor imagery of the feet is utilized.

Regarding a fair comparison of CSP and beamforming, it should be further noted that both algorithms can be improved in several ways. Instead of automatically selecting a set of spatial filters obtained by CSP, spatial filters can also be visually inspected and selected according to prior knowledge on meaningful filter topographies. While this can be expected to increase the robustness of CSP, it requires expert supervision, and thus, limits the routine applicability of BCIs. Further note that several heuristics exist to fine-tune CSP [3]. In the case of beamforming, the probably most relevant factor affecting classification accuracy is a misplaced ROI. In this study, ROIs were chosen rather arbitrarily at locations assumed to include the left and right motor cortex. It is expected that classification performance can be further improved by a subject-specific adaptation of the location and size of the ROIs using cross-validation on the training set. Furthermore, it would be interesting to determine optimal generic, i.e., subject-independent, ROIs for a given experimental paradigm. For other factors that affect beamformer performance and could be improved on, also confer Section II-C.

In summary, we do not wish to argue that either CSP or beamforming perform superior in general, but rather see both methods as complimentary approaches to SF in noninvasive BCIs. While CSP probably provides theoretically optimal spatial filters [9], beamforming can be particularly useful if CSP fails to compute sensible spatial filters—whether this is due to subjects not being able to induce strong modulations of their μ-rhythm, a strong contamination of the recorded EEG by artifactual components, or too few training trials being available. As such, we believe that beamforming might prove to be particularly useful when working with subjects in later stages of ALS since experiments here usually have to be conducted in clinical environments under nonoptimal conditions. Finally, we would like to point out that the observation of CSP and beamforming favoring different spectral bands might indicate that both approaches extract (at least partially) independent information on the subject's intention. As such, a combination of CSP and beamforming might prove to be useful.

REFERENCES

[1] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Process. Mag.*, vol. 18, no. 6, pp. 14–30, Nov. 2001.
[2] G. Blanchard and B. Blankertz, "BCI competition 2003—Data set IIa: Spatial patterns of self-controlled brain rythm modulations," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1062–1066, Jun. 2004.
[3] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Mueller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 41–56, Jan. 2008.
[4] C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller, "Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis," *Pattern Recognit. Lett.*, vol. 28, pp. 957–964, 2007.
[5] J. F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 9, no. 10, pp. 2009–2025, Oct. 1998.
[6] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, 2004.
[7] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 141–144, Jun. 2003.
[8] J. Gross and A. A. Ioannides, "Linear transformations of data space in MEG," *Phys. Med. Biol.*, vol. 44, no. 8, pp. 2081–2097, 1999.

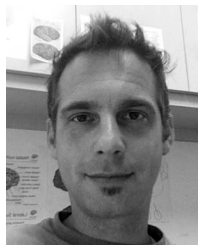
[9] M. Grosse-Wentrup and M. Buss, "Multi-class common spatial patterns and information theoretic feature extraction," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 1991–2000, Aug. 2008.
[10] M. Grosse-Wentrup, K. Gramann, E. Wascher, and M. Buss, "EEG source localization for brain–computer interfaces," in *Proc. 2nd Int. IEEE EMBS Conf. Neural Eng.*, Arlington, VA, Mar. 2005, pp. 128–131.
[11] M. Grosse-Wentrup, K. Gramann, and M. Buss, "Adaptive spatial filters with predefined region of interest for EEG based brain–computer-interfaces," in *Advances in Neural Information Processing Systems,* vol. 19, B. Schoelkopf, J. Platt, and T. Hoffman, Eds. Cambridge, MA: MIT Press, 2007, pp. 537–544.
[12] K. Koh, S.-J. Kim, and S. Boyd, "An interior point method for large-scale l_1-regularized logistic regression," *J. Mach. Learning Res.*, vol. 8, pp. 1519–1555, 2007.
[13] S. G. Mason, A. Bashashati, M. Fatourechi, K. F. Navarro, and G. E. Birch, "A comprehensive survey of brain interface technology designs," *Ann. Biomed. Eng.*, vol. 35, no. 2, pp. 137–169, 2007.
[14] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, pp. 386–394, 1997.
[15] K. R. Mueller, C. W. Anderson, and G. E. Birch, "Linear and nonlinear methods for brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 165–169, Jun. 2003.
[16] A. Y. Ng, "Feature selection, L_1 vs. L_2 regularization, and rotational invariance," in *Proc. 21st Int. Conf. Mach. Learning (ICML 2004)*, C. E. Brodley, Ed. Jul. 4–8, Banff, AB, Canada: ACM.
[17] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*. London, U.K.: Oxford Univ. Press, 2005.
[18] M. Palus, "Nonlinearity in normal human EEG: Cycles, temporal asymmetry, nonstationarity and randomness, not chaos," *Biol. Cybern.*, vol. 75, no. 5, pp. 389–396, 1996.
[19] G. Pfurtscheller and F. H. Lopes da Silva, "Even-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, pp. 1842–1857, 1999.
[20] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "EEG-based discrimination between imagination of right and left hand movement," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, pp. 642–651, 1997.
[21] H. Ramoser, J. Mueller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, Dec. 2000.
[22] S. Rush and D. A. Driscoll, "EEG electrode sensitivity—An application of reciprocity," *IEEE Trans. Biomed. Eng.*, vol. BME-16, no. 1, pp. 289–296, 1969.
[23] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
[24] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

Moritz Grosse-Wentrup (S'01–M'09) was born in Germany in 1978. He received the Diploma in electrical engineering and information technology and the Dr.Eng. degree from the Technische Universität München, Munich, Germany, in 2004 and 2008, respectively.

He is currently a Postdoctoral Researcher in the Department of Empirical Inference, Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany. His current research interests include brain–computer interfaces, causal inference, and information theory.

Christian Liefhold was born in Germany in 1982. He received the Diploma in electrical engineering and information technology from the Technische Universität München, Munich, Germany, in 2008.

He is currently a Research Assistant at the Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany. His current research interests include the optimization and the application of spatiotemporal preprocessing for EEG-based brain–computer interfaces.



Klaus Gramann received the B.S. degree in psychology from the University of Giessen, Giessen, Germany, in 1994, and the M.S. and Ph.D. degrees from the Technichal University of Aachen, Aachen, Germany, in 1998 and 2002, respectively.

During 2007, he was a Postdoctoral Fellow in the Department of Experimental Psychology, University of Munich, Munich, Germany. He is currently a Project Scientist at the Swartz Center for Computational Neuroscience, University of California, San Diego.

Dr. Gramann received the Borchers Honors for his Ph.D. thesis in 2002 and was awarded the Lienert stipend in 2006.

Martin Buss (S'94–M'95) received the Diploma in electrical engineering from the Technical University Darmstadt, Darmstadt, Germany, in 1990, the Dr.Eng. degree in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1994, and the Habilitation degree from the Department of Electrical Engineering and Information Technology, Technische Universität München, Munich, Germany.

In 1988, he was a Research Student at the Science University of Tokyo, Tokyo, Japan. From 1994 to 1995, he was a Postdoctoral Researcher in the Department of Systems Engineering, Australian National University, Canberra, Australia. From 1995 to 2000, he was a Senior Research Assistant and a Lecturer in the Department of Electrical Engineering and Information Technology, Institute of Automatic Control Engineering, Technische Universität München, where he has been a Full Professor since 2003 and has also been the Coordinator of the German Research Foundation (DFG) Excellence Research Cluster "Cognition for Technical Systems"—CoTeSys since 2006. From 2000 to 2003, he was a Full Professor, the Head of the Control Systems Group, and the Deputy Director of the Institute of Energy and Automation Technology, Faculty IV—Electrical Engineering and Computer Science, Technical University Berlin, Berlin, Germany. His current research interests include automatic control, mechatronics, multimodal human–system interfaces, optimization, nonlinear, and hybrid discrete-continuous systems.