FORCe: Fully Online and Automated Artifact Removal for Brain-Computer Interfacing

Ian Daly, Reinhold Scherer, Martin Billinger, and Gernot Müller-Putz

***Abstract*—A fully automated and online artifact removal method for the electroencephalogram (EEG) is developed for use in brain-computer interfacing (BCI). The method (FORCe) is based upon a novel combination of wavelet decomposition, independent component analysis, and thresholding. FORCe is able to operate on a small channel set during online EEG acquisition and does not require additional signals (e.g., electrooculogram signals). Evaluation of FORCe is performed offline on EEG recorded from 13 BCI particpants with cerebral palsy (CP) and online with three healthy participants. The method outperforms the state-of the-art automated artifact removal methods Lagged Auto-Mutual Information Clustering (LAMIC) and Fully Automated Statistical Thresholding for EEG artifact Rejection (FASTER), and is able to remove a wide range of artifact types including blink, electromyogram (EMG), and electrooculogram (EOG) artifacts.**

***Index Terms*—Automated online artifact removal, brain-computer interface (BCI), electroencephalogram (EEG), independent component analysis, wavelets.**

I. Introduction

BRAIN-COMPUTER interfaces (BCIs) allow control of a computer, or other device, via the modulation of neurological activity in the participants' brain and without requiring any activation of the efferent nervous system [1]. Therefore, BCIs have been proposed as potential assistive devices for participants who experience difficulties exerting control via their efferent nervous system. Proposed user groups include participants with spinal cord injury (SCI) [2], amyotrophic lateral sclerosis (ALS) [3], minimally conscious state participants [4], and participants with cerebral palsy (CP) [5].

Arguably, the most widely used mechanism for acquiring BCI control signals from the brain is the electroencephalogram (EEG) [2]. The EEG records summed electrophysiological activity generated from cortical neuronal activity and projected through the skull and scalp [6]. It has the advantage of providing a very high temporal resolution while being relatively cheap and portable, allowing for bedside and home use [7].

However, these advantages come at a cost. The EEG has a very poor spatial resolution due to the effects of volume conduction over the cortical surface [8]. More importantly, the amplitude of the EEG is often contaminated with other electrical activity, which may be observed in the signal but is not related to brain activity. Such additional signals are referred to as artifacts and may arise from external sources of noise, such as power lines and electrical equipment, or internally from the participant using the BCI. Artifacts arising from the participant may be generated by a number of sources including blinks, muscle movement, and head movement [9].

There is a need to remove each of these artifact types prior to analysis of the EEG and its use in BCI control, to ensure that any control achieved may be genuinely attributed to the participants brain activity. However, this is a nontrivial task. Artifacts, particularly participant generated artifacts, occupy overlapping spectral bands with the neurological activity of interest, may occur on many or all channels, and often have a larger amplitude than the EEG signal components of interest. Thus, simple frequency band or spatial filtering will not adequately remove them [9].

This problem is particularly important for online BCI operation. During offline EEG analysis it is possible to visually identify and remove epochs containing artifacts post-measurement. However, during online BCI operation this is not possible and instead some automated method for identifying, and ideally for removing, artifacts from the EEG is needed.

Many different automated artifact removal methods have been proposed for EEG denoising. Common methods for automated artifact removal include wavelet based denoising such as [10] and blind source separation methods such as [11].

However, many of these methods are not suitable for use in online BCI applications due to long runtimes or low accuracy. Some online methods have been proposed. For example, in [12] a method is proposed for artifact removal based upon ICA and support vector machines (SVMs) to classify artifactual components. However, the method is only designed to work with electrooculogram (EOG) and electromyogram (EMG) artifact types and is not highly accurate.

We propose a fully automated online artifact removal method for brain-computer interfacing (FORCe) based upon a combination of ICA, wavelet decomposition, and both hard and soft thresholding a set of key statistical, spectral, temporal, and spatial properties of the EEG. The method is primarily intended for use on the removal of participant generated artifacts. It is designed to be able to remove a wide range of such artifact types

Manuscript received April 22, 2013; revised January 16, 2014; accepted August 02, 2014. Date of publication August 13, 2014; date of current version September 03, 2015. This work was supported by the FP7 Framework EU Research Project ABC under Grant 287774.

I. Daly is with the Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology, Inffeldgasse 13/IV, 8010 Graz, Austria, and also with the Brain Embodiment Lab, School of Systems Engineering, University of Reading, Reading, Berkshire, RG6 6AY, U.K.

R. Scherer, M. Billinger, and G. Müller-Putz are with the Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology, Inffeldgasse 13/IV, 8010 Graz, Austria (e-mail: reinhold.scherer@tugraz.at).

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/TNSRE.2014.2346621

from the EEG accurately, while minimizing perturbations to artifact contaminated EEG epochs. Additionally, the method is able to operate without needing additional simultaneous EOG or EMG recordings.

Wavelet decomposition is first applied to EEG recorded on each channel within a 1 s time window. ICA is then used to translate the resulting approximation coefficients into independent components which are thresholded to remove artifact contaminated ICs. Soft thresholding is then applied to both the detail and approximation coefficients to reduce/remove the artifact contamination arising from spiking activity (e.g., EMG). Finally, the cleaned EEG signals are reconstructed.

We compare FORCe to the online automated state-of-the-art artifact removal method Lagged Auto-Mutual Information Clustering (LAMIC). LAMIC is a blind source separation (BSS) and clustering-based method which has been shown to outperform other state-of-the-art wavelet and spectrum analysis based artifact removal methods [13].

We also compare FORCe to the state-of-the-art offline artifact removal method Fully Automated Statistical Thresholding for EEG artifact Rejection (FASTER). FASTER is also based upon BSS methods and has been shown to be effective at removing/reducing a wide range of EEG artifacts [11].

FORCe is trained and tested in a simulated online BCI environment with EEG recorded from participants with CP [14]. It is also run online with healthy participants to demonstrate the methods efficacy during online BCI operation.

II. Methods

The proposed method (FORCe: Fully Online and automated artifact Removal for brain-Computer interfacing) is first described. Then the artifact removal methods LAMIC and FASTER, against which it is to be compared, are described. Finally, the tests applied to compare the methods are detailed.

A. Proposed Method: FORCe

1) Overview: Our proposed method FORCe (Fully Online and automated artifact Removal for brain-Computer interfacing) attempts to remove artifact components from 1 s windows of the EEG via the following steps.

1) Decompose the EEG on each channel into a set of approximation and detail coefficients via a wavelet decomposition. Denote $c_j^i \in C$ the jth coefficient set from the set of all coefficients C, from channel i.
2) Group all coefficients at the same decomposition level from each channel into sets of coefficients, $A_n = c_j^i \in C | \forall i \in K, j = n$, where K is the set of channels and n denotes the decomposition level.
3) For the set of approximation coefficients (A_1) estimate an ICA demixing matrix to separate the coefficients into maximally statistically independent components (ICs).
4) Multiply the set of approximation coefficients by the demixing matrix.
5) Identify ICs which contain artifacts and remove them.
6) Invert the ICA decomposition to obtain an estimate of the cleaned approximation coefficient set $\bar{A}_1$.
7) Identify spike zones in both the approximation and detail coefficient sets and apply soft thresholding to reduce their magnitude.
8) Reconstruct the cleaned EEG from the wavelet approximation and detail coefficient sets.

Each step of the FORCe method is detailed below.

2) Wavelet Decomposition: Wavelets attempt to decompose a signal by convolving it with a mother wavelet function at a range of different time and frequency locations and measuring the strength of the signal as a coefficient of the wavelet function [15]. For practical purposes the discrete wavelet transform (DWT) is used; this scales to the signal at a discrete set of times and frequencies.

The wavelet transform may be defined as

$$\omega(t, f) = \int_{-\infty}^{\infty} x(t) * \psi_{s,\tau}(t) dt \tag{1}$$

with

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t - \tau}{s} \right) \tag{2}$$

where $x(t)$ is the original signal and $*$ denotes the complex conjugation. $\omega(t, f)$ shows how the signal $x(t)$ is translated into a set of wavelet basis functions $\psi_{s,\tau}(t)$ at scale and translation dimensions s and τ. ψ is the mother wavelet function with which the signal is convolved.

The Symlet 4 "Sym4" mother wavelet is used in this work to decompose the signals into approximation and detail coefficients down to two decomposition levels. This is chosen based upon work in [16]. Approximation and detail coefficients are then used as the basis for the remaining steps in the artifact removal process.

3) Independent Component Analysis: Independent component analysis (ICA) attempts to separate multivariate signals into subcomponents which are maximally statistically independent from one another. The EEG is assumed to arise from the summed electrical activity generated from multiple independent sources. ICA attempts to estimate the mixing process which gave rise to the EEG from these sources and then, by inverting the mixing matrix, to attempt to reconstruct the sources [17].

Formally, this may be defined as

$$x = Ws \tag{3}$$

where x denotes the EEG signals recorded from the scalp, s the original dipole sources from which the EEG originates, and W the linear mixing matrix. Reconstructing the sources from the EEG may, therefore, be performed by inverting the mixing matrix.

Multiple methods have been proposed to estimate the mixing matrix W from the EEG, with the majority focussed on finding an estimate of W that provides maximally statistically independent sources [18]. In this work the second order blind identification (SOBI) method for estimating the ICA mixing matrix W is employed. This method is based upon the joint diagonalization of intersignal correlation matrices over time. It is chosen based upon its observed success in separating artifacts from the EEG [19]. For further details on the method please refer to [20].

4) Identification of Artifact Contaminated ICs: The ICs which contain artifacts need to be accurately identified and removed. A number of approaches are taken to do this. The FORCe method aims to remove eye blinks, EOG activity related to other (non-blink) eye movements, EMG activity related to muscle movements, and electrocardiogram (ECG) artifacts.

ICs containing each of these artifacts may be identified in a variety of ways. Artifacts are known to differ from clean EEG in the following properties:

1) amount of temporal dependency within the signal [21];
2) amount of spiking activity observed within the signal [22];
3) kurtosis value of the signal measuring peakedness of the signal amplitudes over time [23];
4) similarity of the power spectral density distribution of the observed EEG to a $1/F$ distribution (with F denoting frequency) [24];
5) power spectral density in the Gamma frequency band and above ($>$ 30 Hz) [25];
6) standard deviation and topographic distribution of the amplitude values of the EEG time series [26];
7) peak amplitudes of the EEG time series [6], [27].

ICs are identified as likely to contain an artifact when they exceed thresholds for one or more of these criteria. However, it is possible for a period of clean EEG to exceed a threshold (a false positive identification). Therefore, to attempt to minimize the influence of such false positive artifact detections the number of thresholds exceeded by each IC is counted. ICs which exceed more than three thresholds are removed.

All thresholds chosen for use in this method are set based upon manual adjustment of threshold values and subsequent visual inspection of the resulting cleaned EEG time series. This process is performed on two EEG datasets, EEG recorded from a BCI participant with CP and EEG recorded from a healthy participant. Subsequently, EEG from these two participants is excluded from use in validation of the FORCe method. Each threshold is now detailed.

The amount of temporal dependency within a signal can indicate the presence of an artifact [21] and may be measured by the auto-mutual information (AMI). AMI may be considered to be a nonlinear analog of auto-correlation, i.e., a measure of dependence between a time series and a shifted versions of itself at some time lags τ

$$\mathrm{AMI}_\tau(X) = \mathrm{MI}\big(x(t), x(t+\tau)\big) \tag{4}$$

where MI denotes the mutual information

$$\mathrm{MI}(X) = \sum_{i}^{N} H(X_i) - H(X_1, \ldots, X_N) \tag{5}$$

and where $H(X_j) = -\sum_{i=1}^{N} p(x_i) \log p(x_i)$ and $H(X_i, \ldots, X_N) = \sum_{x_1} \cdots \sum_{x_N} p(x_1, \ldots, x_n)$, $\log\ (p(x_1, \ldots, x_n)\ )$ denote the entropy and joint entropy respectively of the random variable X, and $p(x_i)$ denotes the probability of X, estimated at x_i, $i = 1, \ldots, T_m$, where T_m denotes the number of samples in each realization of X.

Thus, AMI carries information that characterizes the temporal dynamics of the random variable X. Therefore, it can be used as a feature for distinguishing between signals that evolve differently in time.

In this work the lag offset (the number of samples between consecutive lags) is set to 60 (for example, with a sampling rate of 512 Hz, decomposed to two levels via wavelet decomposition, this would results in two lags). This choice is based upon observed temporal dynamics of artifacts reported in [13]. AMI is calculated on the scalp projections of each individual IC and a maximum and minimum threshold are set to 3.0 and 2.0 respectively.

Spiking activity in the signal may indicate the presence of EMG artifacts in an IC [22]. ICs with spiking activity are identified in two ways, by amplitude values much higher than other samples in the signal, and by spike zone coefficients.

High amplitude samples are identified as samples with amplitudes greater than $\mu(A) + (3 \times \sigma(A))$, where $\mu(A)$ denotes the mean amplitude over the signal and $\sigma(A)$ the standard deviation of amplitude values. ICs which contain amplitudes exceeding this threshold are marked for removal.

Spike zone coefficients are identified as samples for which

$$x_{i-1} < x_i > x_{i+1} \tag{6}$$

where x_i denotes the amplitude of the signal at sample i. Spikes are then counted by identifying all spike zone coefficients for which

$$C_o > (0.1 \times (\mu(c) + \sigma(x))) \tag{7}$$

where C_o denotes the coefficient of variation of each spike zone, $\mu(c)$ denotes the mean coefficient of variation over all spike zones, and $\sigma(c)$ the standard deviation. C_o is defined as

$$C_o = \frac{\mu(x_{i-1} : x_{i+1})}{\sigma(x_{i-1} : x_{i+1})} \tag{8}$$

where $\mu(\ldots)$ denotes the mean function and $\sigma(\ldots)$ the standard deviation. ICs for which the number of identified spikes divided by the number of spike zone coefficients is greater than 0.25 are marked for removal.

Kurtosis is calculated on the scalp projections of each IC, ICs for which $k > (\mu(k) + (0.5 \times \sigma(k)))$ are removed, where k denotes the kurtosis of the scalp projection of a single IC, $\mu(k)$ the mean of the kurtosis values over the scalp projections of all ICs, and $\sigma(k)$ the standard deviation.

The power spectra of the clean EEG is known to approximately follow a $1/F$ distribution [24]. However, this is only an approximate rule as particular cognitive tasks are known to produce deviations from this distribution in specific frequency bands (for example, in the alpha and beta frequency bands during motor imagery tasks [28]). Thus, a very wide threshold is applied to attempt to detect power spectra which differ by very large amounts from this $1/F$ distribution. Therefore, ICs for which the Euclidean distance between their PSD and an idealized $1/F$ distribution differs by more than 3.5 are marked for removal.

High power spectral density (PSD) in the gamma frequency band and above ($>$ 30 Hz) could indicate the presence of EMG artifact contamination in the EEG [25]. The mean PSD of the

scalp projection of each IC in frequencies above 30 Hz is, therefore, thresholded above 1.7.

High standard deviations in the EEG have also been reported to indicate the presence of EMG and other artifacts [26]. Standard deviation of the projection of the ICs is, therefore, thresholded via $\theta_p > \mu(\theta_p) + (2 \times \sigma(\theta_p))$, where θ_p denotes the standard deviation of the scalp projection of a single IC, $\mu(\theta_p)$ denotes the mean of all θ_p values, and $\sigma(\theta_p)$ the standard deviation of all θ_p values.

Additionally, the topographic distribution of standard deviations on each channel may indicate the presence of an artifact. Specifically, larger standard deviations of amplitude on frontal channels may indicate the presence of EOG activity [29]. Thus, the average standard deviation projected onto frontal channels by each IC is divided by the average standard deviation projected onto the remaining channels. ICs are removed when $R > (\mu(R) + \sigma(R))$, where R denotes the ratio of standard deviations between frontal channels and other channels on a single IC, $\mu(R)$ denotes the mean of ratios over all ICs, and $\sigma(R)$ denotes the standard deviation of ratios over all ICs.

Finally, the peak amplitude values in the projection of each IC are thresholded to attempt to remove ICs with large amplitudes, which could indicate the presence of artifacts [6]. This is done in two ways. First, the projections of each IC are thresholded to $\pm$100 uV, with any ICs which exceed this threshold marked for removal. Second, the peak-to-peak differences between the maximum and minimum amplitudes in the IC projections are thresholded to 60 uV, with any ICs exceeding this threshold also marked for removal.

5) Spike Zone Thresholding: Within both the approximation and detail coefficients derived from the wavelet decomposition, a soft thresholding approach is applied to attempt to reduce the influence of EMG artifact contamination in the EEG. This is derived from a successful approach described in [16]. First, spike zones are identified. These are then soft thresholded to reduce their amplitude.

Spike zones are identified via (6). For each identified spike the coefficient of variation (C_o) is calculated via (7).

The following soft threshold is then applied to all coefficients of variation

$$c(n) = \begin{cases} A \times c(n), & \text{if } c(n) > (G \times (\mu(c) + \sigma(c))) \\ c(n), & \text{else} \end{cases} \quad (9)$$

where $\mu(c)$ denotes the mean over all coefficients of variation, $\sigma(c)$ the standard deviation of all coefficients of variation, A denotes the weight applied to the threshold level, and G the gain applied to adjust the spike amplitude by when its corresponding coefficient of variation exceeds the threshold.

For the approximation coefficients the weight and gain are set to $A = 0.7$ and $G = 0.8$, respectively. For the detail coefficients they are set to $A = 0.2$ and $G = 0.07$, respectively.

B. Threshold Sensitivity

To evaluate the sensitivity of the FORCe method to the specific values chosen, a one-at-a-time sensitivity analysis is applied. Noise is added to each threshold individually and the performance of the method during simulated online operation is evaluated via measuring the event related (de)synchronization (ERD/S) strength after application of the method (a measure of decrease/increase in band-power from baseline during the motor imagery tasks [28]). The thresholds, step sizes, and the range of values used are listed in Table I.

TABLE I
THRESHOLDS USED IN THE PROPOSED FORCE ARTIFACT REMOVAL METHOD AND THE RANGE OF VALUES THEY CAN TAKE

Threshold	Min.	Step	Max.
Lag offset	1	1	120
Max AMI	1.0	0.1	6.0
Min AMI	1.0	0.1	6.0
IC amplitude	1.0	0.1	6.0
Number of spikes	0.1	0.01	0.5
Kurtosis	0.1	0.01	2.0
PSD energy distribution	1.0	0.1	7.0
30 Hz PSD	1.0	0.1	7.0
Std. scalp projections	1.0	0.01	4.0
Frontal / all channel ratio	0.5	0.01	2.0
Max amplitude	50	1	150
Peak to peak amplitude	30	1	90
Soft threshold: A (approx)	0.01	0.01	1.5
Soft threshold: G (approx)	0.01	0.01	1.5
Soft threshold: A (detail)	0.01	0.01	1.5
Soft threshold: G (detail)	0.01	0.01	1.5

C. LAMIC

The FORCe method is compared against the state-of-the-art automated artifact removal method LAMIC. LAMIC is a clustering algorithm developed for performing automatic artifact removal from the EEG [21] and has been shown to outperform other state-of-the-art methods [13].

EEG is first decomposed via the BSS algorithm temporal decorrelation source separation (TDSEP) [30]. Components are then clustered based upon the similarity of their AMI. The clustering procedure follows the general steps of hierarchical clustering methods, whereby at each step the pair of clusters that is closest, based on some proximity measure, is merged into a single cluster.

LAMIC performs the following steps to correct for artifacts.

1) After performing TDSEP, place the ICs for a single trial into an empty data set C.
2) Add into C some additional components that have previously been manually identified as containing artifacts. These additional components are referred to as templates.
3) Estimate the lagged AMI for each member of the set C with appropriate lag values. Lags of ten samples are used in this work based upon experimentation with a subset of the data.
4) Begin clustering nearest neighbor components in the set C based on their lagged AMI.
5) Continue clustering until only one cluster remains.
6) Select an appropriate level from the cluster tree at which the clusters are separated into artifact and nonartifact containing segments. This is done using the fusion levels approach [31].

The data is partitioned into clusters, depending on their distance from each other as estimated via the proximity matrix, using a hierarchical clustering method. Data is iteratively merged via the method described in [21] until a partition containing all data members is obtained.

Once the data is partitioned into groups and the final hierarchical tree is obtained, the partition is assessed in terms of how well it reflects the true data structure and the number of valid clusters ("correct" groupings) is determined. This is done via examination of the "fusion level" against each stage of grouping [31].

Fusion levels are calculated as per [21] and a cut in the hierarchical tree is made at the stage where the plot of the fusion levels flattens, as this shows that little improvement in the description of the data structure is to be gained above this level. In this work the fusion level parameter γ (described in [21]) is set to -0.25 based upon experimentation with a subset of the data (one BCI participant with CP and one healthy participant).

When the clustering has completed the level in the cluster tree denoted by the fusion levels will ideally contain separate clusters containing the artifact components and templates and the nonartifact components. Clusters that contain the artifact templates are removed and the remaining clusters are retained.

D. Faster

The FORCe method is also compared to a state-of-the-art offline artifact removal method called FASTER [11].

This method is based upon the translation of the observed EEG signals into component space via an ICA algorithms followed by statistical threshold based rejection of artifact contaminated components. The implementation of the algorithm for the EEGLAB toolbox was used in this study [32].

E. Comparison

The efficacy and efficiency of the FORCe method is compared with LAMIC and FASTER on two datasets. The methods are first evaluated in a pseudo-online test using prerecorded EEG from 13 participants with CP. The FORCe method is then tested during online BCI operation by three healthy participants to verify that it works online without timing problems.

The methods are compared via the following criteria:

1) visual inspection of the EEG before and after application of each method;
2) signal quality index (SQI) measure of suitability of the signal for BCI control [33];
3) inspection of power spectral densities before and after application of the methods;
4) inspection and comparison of the event-related de/synchronization (ERD/S) spectra and strengths from an MI task performed by participants with CP before and after application of each method;
5) offline classification accuracies for attempted MI BCI control by participants with CP;
6) computational runtime of the methods, when run on a standard laptop computer, during online EEG measurement from 2 healthy participants.

Note, although the artifact removal method is run both offline and online, the criteria by which its efficacy is evaluated are all calculated offline in subsequent analysis.

1) EEG Measurements: Offline: Fourteen participants with CP were measured as part of a study into their ability to control a BCI [14], [34], [35] (seven male, age range 20 to 58 with a median age of 36, SD = 10.97). Institutional review board (IRB) ethical approval was obtained for all measurements. Further participant details are reported elsewhere in [14].

EEG was recorded, at 512 Hz, from 16 electrode channels via the GAMMAsys active electrode system (g.tec, Austria). The following channels were used: AFz, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4, PO3, POz, PO4, O1, Oz, and O2.

EEG recorded from participant 1 was used for training and adjusting of the thresholds used by both the FORCe method and LAMIC. Therefore, EEG from the remaining 13 participants was used to evaluate the efficacy of the methods.

The participants were provided with an MI BCI, as described in [14]. This consisted of an initial calibration phase followed by an online feedback phase.

During calibration the participant was cued to perform kinaesthetically imagined movement of either hand or feet, mental arithmetic, or mental word-letter association. This was followed by classifier training and feedback using the two best classes.

Individual trial timing was as follows:

Second 0: a fixation cross appeared in the centre of the screen and remained there for the duration of the trial.
Second 1.5: A cue appeared on screen indicating the task to perform. This cue remained until second 3.5.
Remaining time: the participant was instructed to perform the cued task and halt when the cross disappeared.

After sufficient trials were recorded in the calibration phase for accurate estimation of the class boundaries the BCI automatically proceeded to the feedback phase in which the two most discriminative classes were selected for use. Note, because different participants required different numbers of trials before accurate classifier boundaries could be found each participant has a different number of corresponding trials. Further details on the feedback may be found in [14] and [36].

2) EEG Measurements: Online: It is important to verify that the FORCe method is able to remove artifacts from the EEG online. This means the method must be able to remove artifacts from an epoch of EEG in less time than the length of the epoch. An epoch of length 1 s is used; therefore, the method must clean the data in considerably less than 1 s.

To verify this and that the method is able to operate on a period of continuously recorded EEG without introducing timing delays or other related problems, the method is applied during online EEG recording, and cleaned online, in a paradigm designed to evoke artifact generation by the participants.

Three healthy individuals participated in an online experiment. All were male with ages of 30, 29, and 29. The first two were right handed and the third left handed.

EEG was recorded, at 512 Hz, from the same electrode positions as for the participants with CP using the same electrode system.

Participants were seated in a comfortable chair approximately 1 m in front of a monitor which cued them to blink, look around the room without moving their head, move their head, clench their jaw, move their arms, or sit still and relaxed.

Each cued trial lasted 4 s and was separated by an intertrial interval of 2 s. Each type of trial was repeated four times per session resulting in 24 trials of a total length of 144 s. In addition, auditory tone cues were played at both the beginning and end

of the trials. The auditory cue at the end was to ensure participants ceased movement at the correct time (for example, when looking around the measurement room the participant may not notice a visual cue to cease movement). Participants completed 4, 5, and 6 runs, respectively.

F. Signal Quality Index

The signal quality metric (SQI) [33] was employed to quantify the EEG signal quality. SQI is measured from 0 to 1, with 0 denoting clean artifact free EEG and 1 denoting heavily artifact contaminated EEG. Parameters were extracted from the EEG from four distinct cortical regions—frontal, central, temporal, and parietal/occipital located channels—and included maximum amplitude values, standard deviations, kurtosis values, and skewness values. These were then thresholded. Further details are provided in [33].

When applied to continuous EEG the SQI was applied in a moving window of length 1 s, step size 0.25 s [37]. In each window the percentage of thresholds exceeded was calculated and the mean and distribution of SQI values over the signal indicated how clean the EEG was.

G. ERD/S

ERD/S refers to a reduction in oscillatory activity in the α (8–13 Hz) and β (13–30 Hz) frequency bands during motor activity or imagery and followed by an increase over baseline values after movement cessation [28]. Detection of ERD/S may be used to identify when a BCI participant is attempting motor imagery (MI) and, therefore, forms the basis of MI-BCI control. Ideally, after application of an artifact removal method the ERD/S effect should either be more prominent (if artifacts were present) or unaltered (if the EEG were clean).

ERD/S strengths were calculated as changes in bandpower between 1.5–8 s against a baseline period of 0 to 1.5 s relative to the start of the trial. The ERD/S spectra were also calculated from trials of 0 to 8 s relative to the start of the trial, via the method described in [38], during the MI tasks the participants with CP were asked to perform.

Mean ERD/S strength was calculated as the sum of the relative bandpowers between 1.5–8 s relative to the start of the trial from the α band (8–13 Hz).

H. Classification

To determine if there is an effect on BCI classification results, for the participants with CP, induced by application of either of the investigated artifact removal methods, accuracies are calculated offline for both the original EEG and the EEG cleaned by each of the methods (the FORCe method, LAMIC, and FASTER). This ensures that comparisons made are meaningful, as with both datasets the accuracy reflects the best accuracy achievable via offline analysis. For a given trial of length 8 s (0–8 s relative to the cue presentation time), band power features were extracted from the frequency bands 8–13 Hz and 13–30 Hz via a sliding window of length 1 s, step size 1 sample. A linear discriminant analysis (LDA) classifier was then applied in a sliding window of length 0.25 s. Training and validation was performed within a leave-one-out train and validation scheme for trials recorded during the training phase (four classes), independently for each participant.

TABLE II
THRESHOLD SENSITIVITY RESULTS. ASTERISKS (*) INDICATE THRESHOLDS TO WHICH THE FORCE METHOD IS SIGNIFICANTLY SENSITIVE ($p < 0.01$)

Threshold	Senstivity	
Lag offset	-0.649	*
Max AMI	0.556	*
Min AMI	-0.288	*
IC amplitude	0.000	*
Number of spikes	0.684	*
Kurtosis	0.899	-
PSD energy distribution	0.955	-
30 Hz PSD	0.275	-
Std. scalp projections	-0.965	*
Frontal / all channel ratio	0.654	*
Max amplitude	0.706	*
Peak to peak amplitude	-0.419	-
Soft threshold: A (approx)	-0.942	*
Soft threshold: G (approx)	-0.191	*
Soft threshold: A (detail)	0.859	*
Soft threshold: G (detail)	0.999	*

I. Visual Inspection

To verify the efficacy of the artifact removal methods visual inspection of the signals, before and after artifact removal, was performed by two blinded reviewers. The reviewers had two (HH) and five (ID) years of experience at labelling EEG for artifacts and the dataset was split evenly between them. Comparisons were then made between the methods and the original EEG based upon the percentage contamination of the signals by each type of artifact.

J. Statistics

Comparisons of the efficacy of each of the methods, as measured by each of the criteria described here, are made using nonparametric Friedman tests. Multiple comparisons corrections are performed via Tukeys honestly significant difference routine.

III. RESULTS

The sensitivity of the FORCe method to different thresholds is first reported. The efficacy of each of the methods is then reported from offline application to the EEG recorded during attempted BCI control by participants with CP. Finally, the results from the online measurements are reported.

A. Threshold Sensitivity

The sensitivity of each threshold to random variations is reported in Table II. Sensitivity is reported in terms of correlation between final success of the method (as measured via ERD/S strength) and randomly selected threshold values. Thus, the greater the sensitivity value the higher the sensitivity of the method to the threshold.

Note, in all cases the manually selected thresholds produce higher mean ERD/S strengths than the mean ERD/S over all randomly selected thresholds. The asterisks indicate thresholds

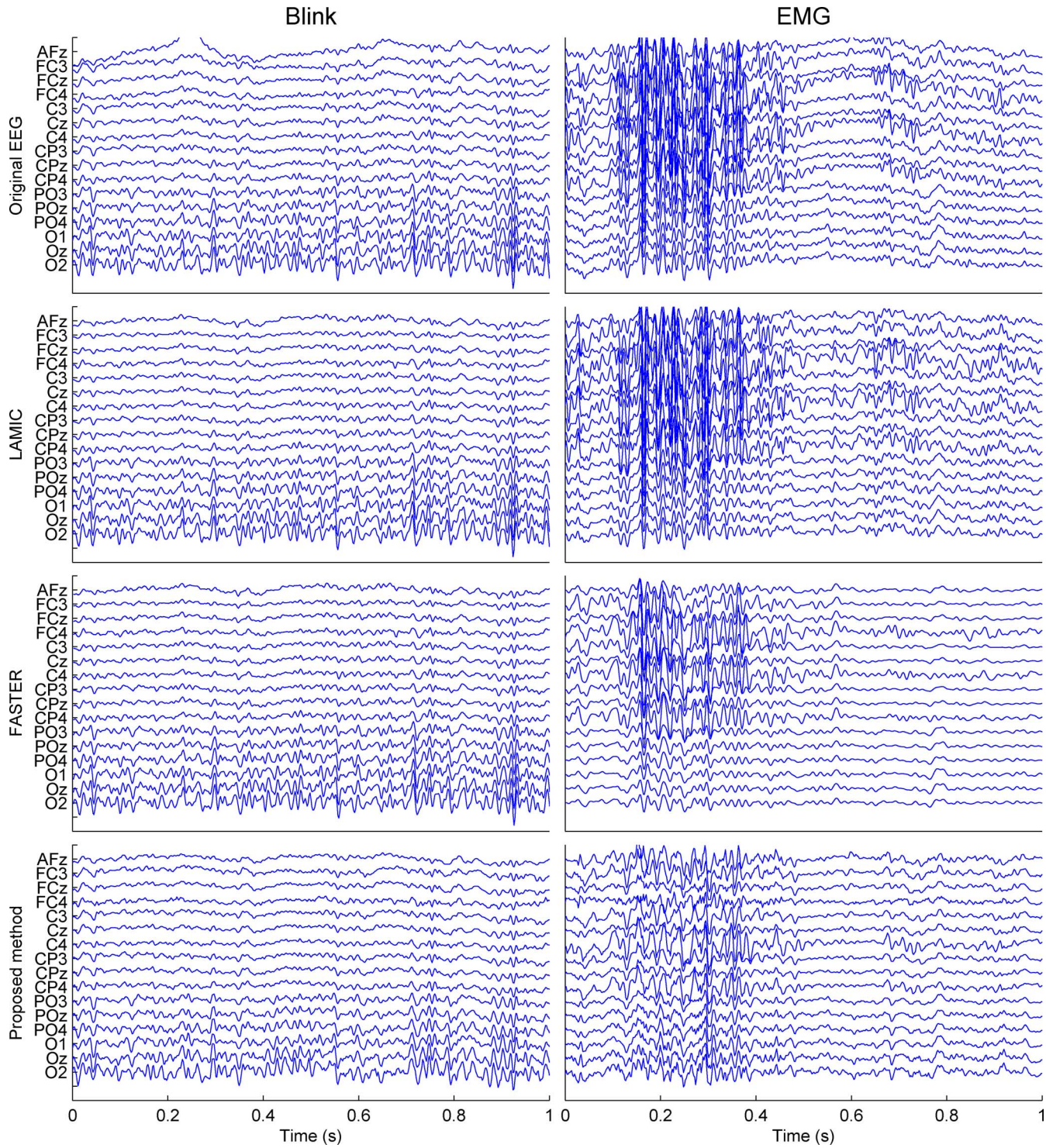

Fig. 1. Example signals before and after artifact removal by each of the methods from a participant with CP (participant 9). Figures in the top row illustrate EEG epochs prior to artifact removal containing a blink (left column) and EMG (right column). Second row figures illustrate the EEG signal after cleaning by LAMIC and the third after cleaning via FASTER. Figures on the lowest row illustrate the EEG after cleaning by the proposed (FORCe) method.

for which the manually selected thresholds are observed to perform significantly better than randomly selected thresholds ($p < 0.05$) as assessed via a two sample Kolmogorov-Smirnov test.

B. Offline Data

Examples of EEG epochs cleaned by LAMIC, FASTER, and the FORCe artifact removal method are shown in Fig. 1.

Note, all methods remove the blink artifact. However, when applied to the EMG artifact the FORCe method produces a clear reduction in the EMG contamination, while LAMIC and FASTER, by contrast, produce only small changes in the signal.

Fig. 2 illustrates power spectra from portions of the EEG labelled as clean and as containing artifacts before and after application of the FORCe method. Note that, in the case of artifact

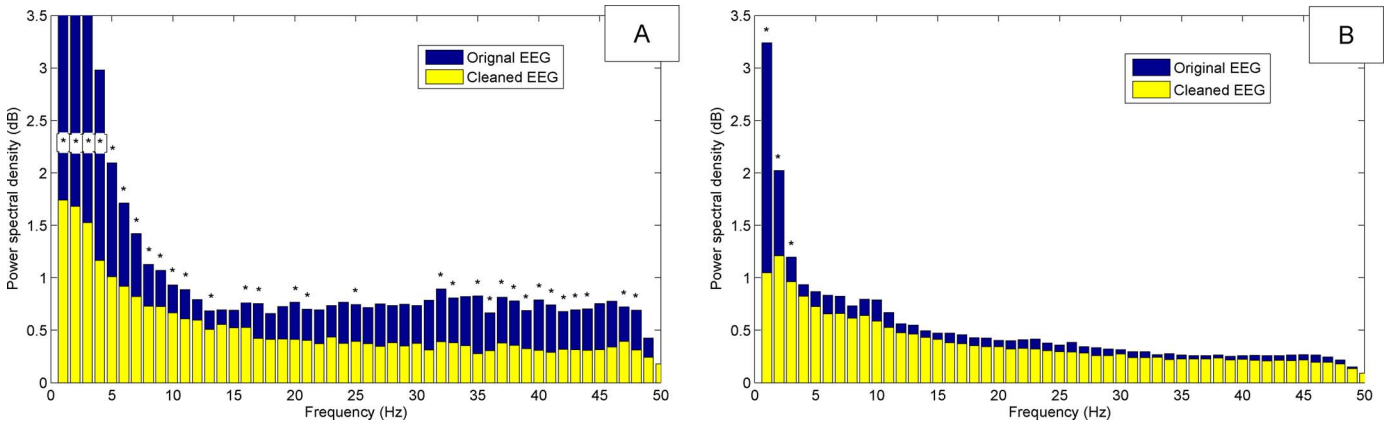

Fig. 2. Power spectra calculated via Welch's method with a Hamming window and a 512-point FFT; 32 point overlap. Power spectra are calculated from a typical participant with CP (participant 9) from the original EEG and the EEG cleaned by the FORCe method. Power spectra are calculated over portions of the signal visually identified as containing artifacts (plot A) and free of artifacts (plot B). Frequency bands of width 1 Hz that are significantly different between conditions (as determined by paired t-tests, corrected for multiple comparisons via the false discovery rate method, $p < 0.01$) are indicated via asterisks (*).

TABLE III
MEAN SQIS CALCULATED FROM ORIGINAL EEG SIGNALS, EEG CLEANED BY FORCE METHOD, EEG CLEANED BY FASTER, AND EEG CLEANED BY LAMIC. ASTERISKS (*) INDICATE HIGHLY SIGNIFICANT ($p < 0.001$) DIFFERENCES IN SQI FROM ORIGINAL EEG, AS ASSESSED BY NONPARAMETRIC PAIRWISE KRUSKAL-WALLIS ANOVAS BETWEEN GROUP PAIRS

Method	Mean	Std.	
Original EEG	0.133	0.032	
LAMIC	0.099	0.039	*
FASTER	0.091	0.030	*
FORCe	0.063	0.024	*

contaminated EEG, the artifact removal significantly reduces the power spectra at a broad range of frequencies ($p < 0.01$, false discovery rate corrected). In the case of portions of the EEG labelled as clean only very low frequencies ($\leq$ 3 Hz) are significantly reduced. This may be due to low frequency dc offset effects being removed via the FORCe method. In the case of LAMIC and FASTER similar observations are made. However, while the high frequency band activity ($>$ 30 Hz) is significantly reduced by our method, both LAMIC and FASTER do not significantly reduce this.

Table III lists the SQIs before and after application of the methods. Note, there are significant decreases in SQI induced by each of the methods but that the FORCe method induces greater reductions than LAMIC or FASTER. This is confirmed by applying a nonparametric Friedman test with column factor Group (original EEG, EEG cleaned via LAMIC, EEG cleaned via FASTER, and EEG cleaned via the FORCe method) and row factor participant number. A highly significant effect of the column factor Group is observed $\chi^2(3, N = 51) = 33.46$, $p < 0.001$. After correction for multiple comparisons (Tukeys honestly significant difference test) the original EEG and the EEG cleaned by the FORCe method are both observed to be significantly different from other groups, with the FORCe method performing significantly better than both FASTER and LAMIC.

Fig. 3 presents an example of ERD/S spectra calculated from EEG recorded from a participant with CP (participant 2) during hand and feet MI. Note, when only a small amount of artifact is present, both LAMIC, FASTER, and the FORCe method produce equivalent results. However, when there is a large amount of artifact present in the signal the FORCe method is much more successful at reducing the artifact contamination of the spectra.

Mean ERD/S values are compared between the methods. The mean ($\pm$standard deviations) ERD/S strengths from each dataset are listed in Table IV. Additionally, a nonparametric Friedman test is used to compare the ERD/S strengths between column factor "Conditions" with row factor participant number. Although no significant differences are found ($\chi^2(3, N = 111) = 2.7$, $p > 0.05$) there is a visibly apparent, small, increase in ERD/S strength after application of the FORCe method.

Peak classification accuracies are calculated from the motor imagery period (3.5–8 s relative to the start of the trial) and listed in Table V for each of the artifact removal methods. A nonparametric Friedman test, performed over the peak classification accuracies achieved with EEG cleaned by each method, with column factor method and row factor participant number, reveals a significant effect of the factor "Method" $\chi^2(3, N = 51) = 12.57$, $p = 0.005$. Post-hoc Tukey's honestly significant criterion tests reveal significant differences between the original EEG and the FORCe method ($p = 0.019$) and between the FORCe method and FASTER ($p = 0.007$). However, no significant differences are found between the original EEG and LAMIC ($p = 0.769$) or the original EEG and FASTER ($p = 0.988$).

Additionally, a measure of the central tendency of the accuracies may be used to compare the accuracies achieved with each of the datasets. The median accuracy achieved with the original EEG in the time period 3.5– 8 s relative to the start of the trial is 0.219 ± 0.019 ($\pm$ standard deviation), while the median accuracy achieved using data cleaned via the LAMIC method is 0.220 ± 0.032, the median accuracy achieved with FASTER is 0.233 ± 0.020, and the median accuracy achieved with the FORCe method is 0.255 ± 0.027.

Fig. 4 illustrates the mean and standard deviation of the peak accuracies achieved with the original EEG and the EEG after application the methods. Note, each method produces an increase

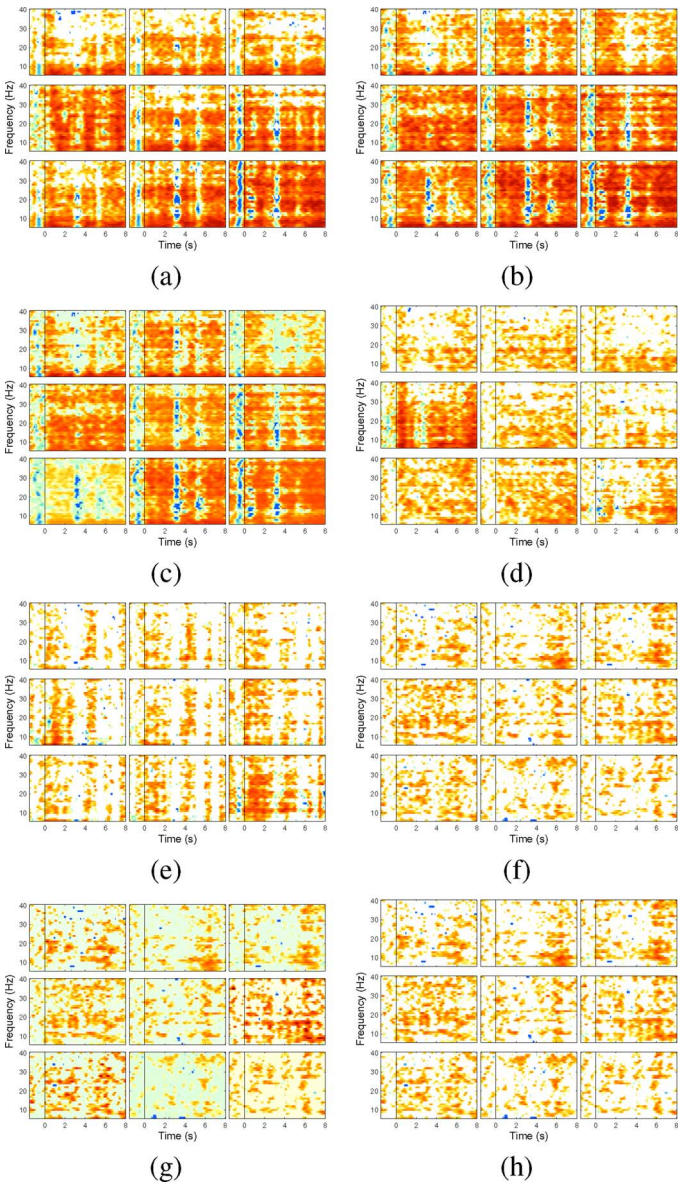

Fig. 3. ERD/S spectra during left hand and feet MI as measured from the original EEG, EEG after cleaning by LAMIC, EEG cleaned by FASTER, and EEG after cleaning by the FORCe method. Channels illustrated in each subfigure are (from top-left to bottom-right): FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, and CP4. (a) Hand MI, Original EEG. (b) Hand MI, LAMIC. (c) Hand MI, FASTER. (d) Hand MI, FORCe. (e) Feet MI, Original EEG. (f) Feet MI, LAMIC. (g) Feet MI, FASTER. (h) Feet MI, FORCe.

TABLE IV
ERD/S STRENGTHS IN ORIGINAL EEG AND EEG AFTER APPLICATION OF EACH OF THE ARTIFACT REMOVAL METHODS. MEAN AND STANDARD DEVIATION (STD.) ARE LISTED OVER ALL PARTICIPANTS

Method	Mean	Std.
Original EEG	0.316	0.171
LAMIC	0.331	0.224
FASTER	0.322	0.179
FORCe	0.359	0.155

in classification accuracy. However, this increase is only significant after application of the FORCe method.

Mean blinded reviewer scores for the percentage of the signals identified as containing artifacts for each online method

TABLE V
PEAK CLASSIFICATION ACCURACIES ACHIEVED ON TRAINING TRIALS (4 CLASS) BY EACH PARTICIPANT WITH ORIGINAL EEG AND EEG CLEANED BY EACH METHOD. FOR EACH METHOD PEAK ACCURACY, AND SIGNIFICANCE OF THE ACHIEVED ACCURACY AGAINST THE NULL HYPOTHESIS OF RANDOM CLASSIFIER RESULTS, IS LISTED. SIGNIFICANCE LEVELS ARE ADJUSTED VIA FALSE DISCOVERY RATE METHOD [39] AND ADJUSTED p-VALUES ARE REPORTED. P. DENOTES PARTICIPANT NUMBER

P.	Original EEG		LAMIC		FASTER		FORCe	
	Acc.	p_{adj}	Acc.	p_{adj}	Acc.	p_{adj}	Acc.	p_{adj}
2	0.35	0.12	0.35	0.12	0.35	0.08	0.42	0.07
3	0.37	0.04	0.39	<0.01	0.35	0.04	0.39	0.01
4	0.37	0.08	0.37	0.08	0.35	0.12	0.35	0.08
5	0.37	0.07	0.32	0.12	0.32	0.12	0.56	<0.01
6	0.28	0.38	0.40	0.07	0.36	0.12	0.40	0.07
7	0.37	0.12	0.42	0.07	0.42	0.07	0.40	0.07
8	0.48	<0.01	0.38	0.03	0.28	0.32	0.38	0.03
9	0.32	0.22	0.36	0.12	0.34	0.22	0.34	0.22
10	0.33	0.22	0.37	0.04	0.40	0.07	0.40	0.07
11	0.32	0.22	0.32	0.22	0.35	0.12	0.35	0.12
12	0.36	0.04	0.40	0.01	0.35	0.04	0.41	0.01
13	0.40	0.04	0.40	0.04	0.40	0.04	0.45	0.01
14	0.40	0.04	0.38	0.06	0.37	0.06	0.43	0.03

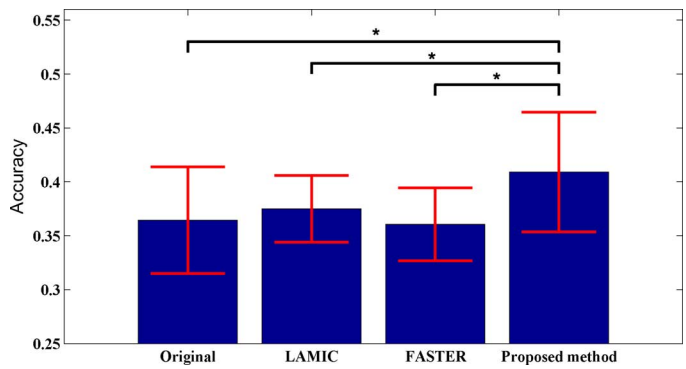

Fig. 4. Mean $\pm$ standard deviation accuracies achieved with the EEG before and after application of each artifact removal method. Asterisk (*) indicates significant difference in classification results between the original EEG and the EEG after application of the proposed (FORCe) method.

are listed in Table VI. Mean and standard deviation percentages of artifact contamination for each artifact (blinks, EMG, movement, failing electrodes, and slow EOG) are listed.

The percentage of artifact contamination between the original EEG and the EEG after application of each of the methods is compared via t-tests. The false discovery rate method is applied to correct for multiple comparisons. Significant changes ($p < 0.05$, corrected) are indicated in Table VI. Asterisks (*) indicate significant improvement over the original EEG and daggers (†) indicate significant differences between the FORCe method and LAMIC.

C. Online Test

Use of the method during online artifact removal is also observed to visibly reduce the presence of artifacts in the signal in a similar manner to that during offline operation.

Examples of power spectra before and after application of the online artifact removal method are shown in Fig. 5. Note that the artifact removal method reduces power at frequencies

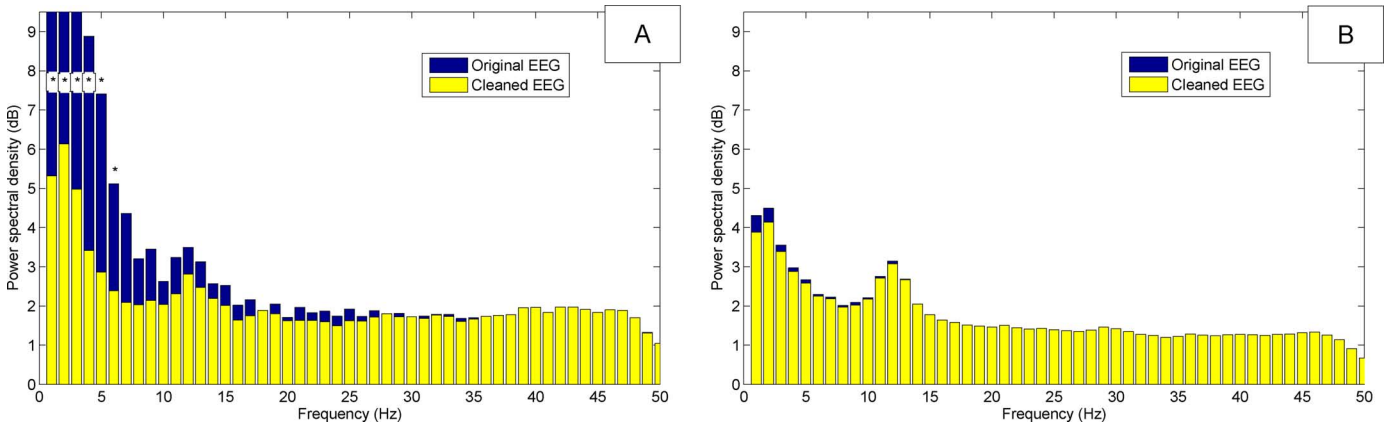

Fig. 5. Online artifact removal: example of EEG power spectra before and after application of the FORCe online artifact removal method. Plot A illustrates the power spectra for portions of EEG identified as containing artifacts before and after application of the FORCe method. Plot B illustrates the power spectra for portions of EEG labelled as clean. Asterisks (*) indicate significant differences ($p < 0.01$, false discovery rate correction applied). Note, cases where the original EEG is not visible have approximately the same magnitude as the original EEG (not significantly different $p > 0.01$).

TABLE VI
PERCENTAGE ARTIFACT CONTAMINATION, BEFORE AND AFTER APPLICATION OF EACH METHOD, AS RATED BY BLINDED REVIEWERS. MEAN AND STANDARD DEVIATIONS ARE LISTED. STANDARD DEVIATION IS AGGREGATED OVER PARTICIPANTS. FOR EACH ROW, SIGNIFICANT DIFFERENCES BETWEEN NUMBERS OF OBSERVED INSTANCES OF EACH ARTIFACT TYPE, BETWEEN EITHER LAMIC AND ORIGINAL EEG OR BETWEEN FORCE AND ORIGINAL EEG ARE INDICATED VIA ASTERISKS (*), WHILE SIGNIFICANT DIFFERENCES BETWEEN LAMIC AND FORCE ARE INDICATED VIA DAGGERS (†). MOVEMENT ARTIFACTS REFER TO LOW FREQUENCY LARGE ARTIFACTS ASSOCIATED WITH HEAD MOVEMENTS, AND OFTEN OCCUR CONJOINTLY WITH EMG. ALL SIGNIFICANCES ARE CORRECTED FOR MULTIPLE COMPARISONS VIA FALSE DISCOVERY RATE METHOD

Blink	4.22	2.53	3.86	1.07
	(±3.15)	(±4.23)	(±3.47)	(±1.70) *
EMG	26.99	26.15	13.91	10.35
	(±16.14)	(±13.36)	(±13.15)	(±6.29) * †
Movement	2.90	2.46	2.09	0.81
	(±3.58)	(±4.00)	(±4.29)	(±0.73)
Failing	0.01	0.01	0.01	0.01
electrode	(±0.04)	(±0.04)	(±0.05)	(±0.03)
Slow EOG	0.93	0.83	0.52	0.36
	(±2.20)	(±1.86)	(±1.69)	(±0.73)
Clean EEG	65.03	67.99	79.61	85.40
	(±16.26)	(±15.47)	(±13.99)	(±10.77) * †

below the alpha frequency band by a large amount and at higher frequencies by a small amount.

During online operation a mean SQI of 0.066 is achieved over the participants with a standard deviation of 0.027. This is similar to that achieved by the FORCe method during offline operation and demonstrates that the method performs equivalently during online artifact removal.

The runtime of the online methods are also recorded. Our method has a mean runtime of 0.382 s (±0.076 s), while LAMIC has a mean runtime of 0.226 s (±0.023 s). Run times are measured over a total of 15 runs spread between three participants. Each run contained 24 trials and had an approximate duration 144 s.

IV. DISCUSSION

The FORCe method has been shown to effectively remove a wide range of participant generated artifacts from the EEG including blinks, EMG, and EOG. The method is able to run during online EEG acquisition and is verified both offline and online. Offline application of the method to EEG recorded from 13 participants with CP during attempted control of an MI-BCI revealed a significant reduction in visually identifiable artifacts, improvement in SQI, reduction in visible artifacts, increase in ERD/S strength, and increase in classification accuracy.

The method is also demonstrated to operate online during EEG acquisition. It may be used to clean the EEG signals and make them available for subsequent processing steps after a short delay. The method does not exhibit an increasing time delay (time leak). We expect future computer technology advancements will reduce this delay even further. The FORCe method is, therefore, suitable for use during online BCI operation when run on a "typical" laptop (2.8 GHz, 2 GB RAM).

It is important to consider the robustness of the method to differing datasets and requirements. Currently the method has been tested on datasets containing 16 EEG channels. It is expected that the method would scale approximately linearly with an increasing number of channels. Thus, with 32 channels the method could take approximately twice as long to run per epoch. This would prevent its use online with the current hardware configuration but does not present an insurmountable problem. Indeed it may be argued that a trend in practical BCI research is to reduce the number of channels used rather than increase them. Thus, this constraint is not overly limiting.

The sample rate at which the data is acquired presents a restriction to the method. A doubling of the sample rate would present the need to decompose the signal to additional levels via the wavelet transformation. This would result in a greater than linear increase in the amount of data to be thresholded and more severely constrains the runtime of the method. However, the sample rate at which the data was originally recorded (512 Hz) is already quite high when compared to a number of other BCI studies. Thus, this restriction on the method is not likely to prove overly onerous for the usability of the method.

It may be argued that, as thresholds were only trained on two participants, the method may not be robust. However, it was shown to successfully reduce the influence of artifacts in a

large number of both healthy participants and participants with a range of CP induced difficulties. Additionally, a sensitivity analysis on the selected threshold values reveals the chosen threshold values to, in every case, perform better than, and in many cases, significantly better than, randomly drawn values from appropriate ranges. Therefore, it may be concluded that the method is robust over a relatively diverse population.

The FORCe method is compared to state-of-the-art methods, LAMIC and FASTER, and is seen to produce significantly better performance as measured by all the metrics employed. As both LAMIC and FASTER have, in previous studies, been compared to other artifact removal methods (Blind source separation, Wavelet based methods etc.), and have been demonstrated to exhibit superior performances, we may conclude that the method proposed and described in this work is also able to perform better than these alternative artifact removal methods.

The only metric at which LAMIC performs best is computational runtime. However, given the small difference, and that our proposed FORCe method is still capable of running fast enough for online operation, this is not considered a significant disadvantage.

The method is unable to completely remove the influence of EMG in every case. This could be for a number of reasons. EMG is broad band and may contaminate multiple channels. Nonetheless, the method does reduce its influence on the signal considerably. Indeed, more electrodes could lead to better reduction of the EMG artifact as the resulting greater number of ICs could allow better separation of the EMG artifact component.

Indeed, existing artifact removal approaches which attempt to remove EMG and/or EOG artifacts often rely on either a much larger number of electrode channels [40] or other reference signals such as EMG or EOG [41]. By way of contrast our FORCe method does not require any additional reference channels and operates accurately on a relatively small set of EEG channels.

Alternatively, the method could be combined with an approach in [42] for the removal of head movement artifacts. This approach uses an accelerometer to record head movement and then attempts to separate ICs which correlate with the accelerometer signal. This is based upon the assumption that EMG induced by head movement would be statistically related, and hence grouped, with ICs which also contain low frequency signals correlating with the accelerometer signal.

The approach was able to remove head movements offline when applied to a long signal epoch. However, some adaptation of the method would be required to allow it to operate during online BCI use, as our FORCe method is capable of.

Future work will seek to validate the FORCe method during online BCI operation by both healthy participants and participants with specific disabilities for which a BCI could be beneficial. Additionally, it is important to further validate the efficacy of the method during other types of BCI operation, for example, during event-related potential (ERP) based BCI control.

V. Conclusion

A novel fully automated online artifact removal method (FORCe) is proposed and validated on an offline BCI dataset and during online EEG acquisition. The method is able to remove/reduce the influence of artifact types including blinks, EMG, ECG, and EOG while preserving EEG components derived from neurological activity. The method is able to operate on only 16 EEG channels and does not require additional signals. We, therefore, propose that the method could be highly beneficial during online BCI operation for a large variety of applications.

Acknowledgment

This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein. The authors would like to extend their thanks to H. Hiebel for assistance with EEG artifact labelling.

References

[1] 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, pp. 767–791, Jun. 2002.

[2] G. R. Müller-Putz, R. Scherer, G. Pfurtscheller, and R. Rupp, "EEG-based neuroprosthesis control: A step towards clinical practice," *Neurosci. Lett.*, vol. 382, no. 1-2, pp. 169–74, 2005.

[3] A. Kübler, F. Nijboer, J. Mellinger, T. M. Vaughan, H. Pawelzik, G. Schalk, D. J. McFarland, N. Birbaumer, and J. R. Wolpaw, "Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface," *Neurol.*, vol. 64, no. 10, pp. 1775–7, May 2005.

[4] C. Pokorny, D. Klobassa, G. Pichler, H. Erlbeck, R. Real, A. Kuebler, D. Lesenfants, D. Habbal, Q. Noirhomme, M. Risetti, D. Mattia, and G. Mueller-Putz, "The auditory-based single-switch BCI: Paradigm transition from healthy subjects to minimally conscious patients," *Artificial Intell. Medicine*, p. 300, 2013, to be published.

[5] C. Neuper, G. R. Müller, A. Kübler, N. Birbaumer, and G. Pfurtscheller, "Clinical application of an EEG-based brain-computer interface: A case study in a patient with severe motor impairment," *Clin. Neurophysiol.*, vol. 114, pp. 399–409, Mar. 2003.

[6] E. Niedermeyer, "The normal EEG of the waking adult," *Electroencephalography: Basic Principles, Clinical Applications and Related Fields 1*, pp. 149–173, 1999.

[7] E. Sellers, T. Vaughan, and J. Wolpaw, "A brain-computer interface for long-term independent home use," *Amyotroph. Lateral Scler.*, vol. 11, no. 5, pp. 449–455, 2010.

[8] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford, U.K.: Oxford Univ. Press, 2006.

[9] W. O. Tatum, B. A. Dworetzky, and D. L. Schomer, "Artifact and recording concepts in EEG," *J. Clin. Neurophysiol.*, vol. 28, no. 3, pp. 252–63, Jun. 2011.

[10] X. Yong, M. Fatourechi, R. K. Ward, and G. E. Birch, "Automatic artefact detection in a self-paced brain-computer interface system," in *Proc. 2011 IEEE Pacific Rim Conf. Communicat., Comput. and Signal Processing*, Aug. 2011, pp. 403–408.

[11] H. Nolan, R. Whelan, and R. B. Reilly, "FASTER: Fully automated statistical thresholding for EEG artifact rejection," *J. Neurosci. Methods*, vol. 192, no. 1, pp. 152–62, Sep. 2010.

[12] S. Halder, "Online artifact removal for brain-computer interfaces using support vector machines and blind source separation," *Computat. Intell. Neurosci.*, vol. 2007, pp. 10–16, 2007.

[13] I. Daly, N. Nicolaou, S. J. Nasuto, and K. Warwick, "Automated artifact removal from the electroencephalogram: A comparative study," *Clinical EEG Neurosci.*, vol. 44, no. 4, pp. 291–306, Oct. 2013.

[14] I. Daly, M. Billinger, J. Laparra-Hernández, F. Aloise, M. L. García, J. Faller, R. Scherer, and G. Müller-Putz, "On the control of brain-computer interfaces by users with cerebral palsy," *Clin. Neurophysiol.*, vol. 124, no. 9, pp. 1787–1797, Sep. 2013.

[15] K. Englehart, *Wavelet Methods in Biomedical Signal Processing*, B. H. Kevin Englehart Philip Parker, Ed. Boca Raton, FL, USA: CRC, 2003, vol. 1, Handbook of Neuroprosthetic Methods.

[16] P. S. Kumar, R. Arumuganathan, K. Sivakuma, and C. Vimal, "Removal of ocular artifacts in the EEG through wavelet transform without using an EOG reference channel," *Int. J. Open Problems Comptut. Math.*, vol. 1, no. 3, pp. 188–200, 2008.

[17] M. P. S. Chawla, H. K. Verma, and V. Kumar, "Artifacts and noise removal in electrocardiograms using independent component analysis," *Internat. J. Cardiol.*, vol. 129, no. 2, pp. 278–281, 2008.

[18] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, vol. 13, no. 4-5, pp. 411–430, Jun. 2000.
[19] A. C. Tang, M. T. Sutherland, and C. J. McKinney, "Validation of SOBI components from high-density EEG," *NeuroImage*, vol. 25, no. 2, pp. 539–53, Apr. 2005.
[20] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.
[21] N. Nicolaou and S. J. Nasuto, "Automatic artefact removal from event-related potentials via clustering," *J. VLSI Signal Process.*, vol. 48, pp. 173–183, 2007.
[22] D. Talsma, "Auto-adaptive averaging: Detecting artifacts in event-related potential data using a fully automated procedure," *Psychophysiology*, vol. 45, no. 2, pp. 216–28, 2008.
[23] D. Mantini, M. G. Perrucci, C. Del Gratta, G. L. Romani, and M. Corbetta, "Electrophysiological signatures of resting state networks in the human brain," in *Proc. Nat. Acad. Sci. USA*, Aug. 2007, vol. 104, pp. 13 170–5, 32.
[24] L. Ward, *Dynamical Cognitive Science*. Cambridge, MA, USA: MIT Press, 2002.
[25] I. Goncharova, "EMG contamination of EEG: Spectral and topographical characteristics," *Clin. Neurophysiol.*, vol. 114, no. 9, pp. 1580–1593, Sep. 2003.
[26] M. Junghofer, T. Elbert, D. M. Tucker, and B. Rockstroh, "Statistical control of artifacts in dense array EEG/MEG studies," *Psychophysiology*, vol. 37, no. 4, pp. 523–532, Jul. 2000.
[27] S. Fitzgibbon, D. Powers, K. Pope, and C. Clark, "Removal of EEG noise and artifact using blind source separation," *J. Clin. Neurophysiol.*, vol. 24, no. 3, pp. 232–243, Jun. 2007.
[28] G. Pfurtscheller and F. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, Nov. 1999.
[29] O. G. Lins, T. W. Picton, P. Berg, and M. Scherg, "Ocular artifacts in EEG and event-related potentials. I: Scalp topography," *Brain Topogr.*, vol. 6, no. 1, pp. 51–63, Jan. 1993.
[30] A. Ziehe and K. Müller, "TDSEP—An efficient algorithm for blind source separation using time structure," in *Proc. Int. Conf. Artificial Neural Networks*, 1998, pp. 675–680.
[31] J. Podani, "New combinatorial clustering methods," *Vegetatio*, vol. 81, no. 1-2, pp. 61–77, Jul. 1989.
[32] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, no. 134, pp. 9–21, 2004.
[33] I. Daly, F. Pichiorri, J. Faller, V. Kaiser, A. Kreilinger, R. Scherer, and G. Mueller-Putz, "What does clean EEG look like?," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2012.
[34] I. Daly, F. Aloise, P. Arico, J. Belda, M. Billinger, E. Bolinger, F. Cincotti, D. Hettich, M. Iosa, J. Laparra, R. Scherer, and G. Mueller-Putz, "Rapid prototyping for hBCI users with Cerebral Palsy," in *Proc. BCI Meet.*, 2013.
[35] I. Daly, M. Billinger, R. Scherer, and G. Müller-Putz, "Brain-computer interfacing for users with cerebral palsy, challenges and opportunities," in *Proc. Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion Lecture Notes in Computer Sci.*, 2013, pp. 623–632.
[36] J. Faller, C. Vidaurre, T. Solis-Escalante, C. Neurper, and R. Scherer, "Autocalibration and recurrent adaptation: Towards a plug and play online ERD-BCI," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 20, no. 3, pp. 313–319, May 2012.
[37] R. Scherer, G. Moitzi, I. Daly, and G. Mueller-Putz, "On the use of games for non-invasive EEG-based functional brain mapping," *Trans. Computat. Intelligence AI Games*, 2013.
[38] B. Graimann, J. E. Huggins, S. P. Levine, and G. Pfurtscheller, "Visualization of significant ERD/ERS patterns in multichannel EEG and ECoG data," *J. Clin. Neurophysiol.*, vol. 113, pp. 43–47, Jan. 2002.
[39] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Statistical Soc.. Ser. B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.
[40] J. T. Gwin, K. Gramann, S. Makeig, and D. P. Ferris, "Removal of movement artifact from high-density EEG recorded during walking and running," *J. Neurophysiol.*, vol. 103, no. 6, pp. 3526–3534, Jun. 2010.
[41] R. Croft and R. Barry, "Removal of ocular artifact from the EEG: A review," *Clin. Neurophysiol.*, no. 1, pp. 5–19, 2000.
[42] I. Daly, M. Billinger, R. Scherer, and G. Mueller-Putz, "On the automated removal of artifacts related to head movement from the EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 3, pp. 427–434, May 2013.

Ian Daly received the M.Eng. degree in computer science and the Ph.D. degree in cybernetics from the University of Reading, Reading, U.K.

Between May 2011 and 2013 he was a Postdoctoral Researcher in the Laboratory of Brain-Computer Interfaces, Graz University of Technology, Graz, Austria. He is currently a Postdoctoral Researcher at the University of Reading, Reading, U.K. His research interests focus on BCIs, nonlinear dynamics, machine learning, signal processing, and connectivity analysis in the EEG and fMRI. He is also interested in the neurophysiological correlates of motor control and stimuli perception and how they differ between healthy participants and individuals with neurological and physiological impairments.

Reinhold Scherer received the M.Sc. and Ph.D. degrees in computer science from Graz University of Technology, Graz Austria, in 2001 and 2008, respectively.

From 2008 to 2010 he was Postdoctoral Researcher and member of the Neural Systems and the Neurobotics Laboratories, University of Washington, Seattle, USA. Since 2011, he has been an Assistant Professor at the Institute for Knowledge Discovery, BCI Laboratory, Graz University of Technology, Graz, Austria, and member of the Institute for Neurological Rehabilitation and Research at the rehabilitation centre Judendorf-Strassengel, Austria. His research interests include BCIs based on EEG and ECoG signals, statistical and adaptive signal processing and robotics-mediated rehabilitation.

Martin Billinger received the M.Sc. degree in electrical engineering from the Graz University of Technology, Graz, Austria. He is working toward the Ph.D. degree in computer science in the Laboratory of Brain-Computer Interfaces at the same university.

His research interests include brain-computer interfaces, bio-signal processing, machine learning, mathematical modelling and effective connectivity analysis in the EEG.

Gernot Müller-Putz received the Ph.D. degree in electrical engineering ("New Concepts in Brain-Computer Communication Use of Steady-State Somatosensory Evoked Potentials, User Training by Telesupport and Control of Functional Electrical Stimulation") from Graz University of Technology, Graz, Austria, in 2004. In 2008 he received his "venia docendi" for medical informatics ('Towards EEG-based Control of Neuroprosthetic Devices') from the faculty of computer science, Graz University of Technology.

Currently, he is Head of the Institute for Knowledge Discovery, Graz University of Technology, Austria. He is also Head of the Laboratory for Brain-Computer Interfaces (BCI-Lab) at Graz University of Technology, where, beginning in 2000, he has worked on noninvasive electroencephalogram-based (EEG) brain-computer interfacing (BCI) for the control of neuroprosthetic devices. His research interests include EEG-based neuroprosthesis control, hybrid BCI systems, the human somatosensory system and assistive technology.