



Citation/Reference	Ben Somers, Tom Francart, Alexander Bertrand (2018), A generic EEG artifact removal algorithm based on the multi-channel Wiener filter Journal of Neural Engineering, vol. 15, 036007, 2018
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	https://doi.org/10.1088/1741-2552/aaac92
Journal homepage	http://iopscience.iop.org/journal/1741-2552
Author contact	alexander.bertrand@esat.kuleuven.be + 32 (0)16 321899
IR	

(article begins on next page)



A generic EEG artifact removal algorithm based on the multi-channel Wiener filter

Ben Somers¹, Tom Francart¹, and Alexander Bertrand²

¹KU Leuven, Department of Neurosciences, Research Group Experimental Oto-Rhino-Laryngology
(ben.somers@med.kuleuven.be, tom.francart@med.kuleuven.be)

²KU Leuven, Department of Electrical Engineering (ESAT), Stadius Center for Dynamical Systems, Signal Processing and Data Analytics (alexander.bertrand@esat.kuleuven.be)

Abstract—Objective: The electroencephalogram (EEG) is an essential neuro-monitoring tool for both clinical and research purposes, but is susceptible to a wide variety of undesired artifacts. Removal of these artifacts is often done using blind source separation techniques, relying on a purely data-driven transformation, which may sometimes fail to sufficiently isolate artifacts in only one or a few components. Furthermore, some algorithms perform well for specific artifacts, but not for others. In this paper, we aim to develop a generic EEG artifact removal algorithm, which allows the user to annotate a few artifact segments in the EEG recordings to inform the algorithm.

Approach: We propose an algorithm based on the multi-channel Wiener filter (MWF), in which the artifact covariance matrix is replaced by a low-rank approximation based on the generalized eigenvalue decomposition. The algorithm is validated using both hybrid and real EEG data, and is compared to other algorithms frequently used for artifact removal.

Main results: The MWF-based algorithm successfully removes a wide variety of artifacts with better performance than current state-of-the-art methods.

Significance: Current EEG artifact removal techniques often have limited applicability due to their specificity to one kind of artifact, their complexity, or simply because they are too “blind”. This paper demonstrates a fast, robust and generic algorithm for removal of EEG artifacts of various types, i.e. those that were annotated as unwanted by the user.

I. INTRODUCTION

The electroencephalogram (EEG) measures electric potentials generated by the brain using electrodes on the scalp. Due to its relatively low cost, high temporal resolution, and portability compared to other neuro-monitoring techniques, it is an important tool in both clinical and research settings.

Despite its wide range of applications, there are some drawbacks to the use of EEG as a research tool. Firstly, the neural responses of interest typically have a low signal-to-noise ratio (SNR), as the electrodes pick up the neural activity from many different brain regions, which interfere with the target response. Secondly, the measurement is very susceptible to artifacts, which are generally classified based on whether they are of biological (i.e. patient-related) or of technical (i.e. environment- or equipment-related) origin [1]. The most common biological artifacts are caused by eye blinking, moving the eyes, cardiac activity, head movements, or muscle contractions associated with speaking, clenching the jaws or swallowing. Technical artifacts originate from

the environment or the equipment, such as interference from power lines and electrical devices, or impedance changes in the recording electrodes.

Technical artifacts can be prevented or largely suppressed by measures related to subject preparation and experimental equipment. For example, the skin-electrode impedance can be reduced by cleaning of the skin before electrode placement, and by application of conductive paste or gel on the skin-electrode contact. This is important, as a high impedance makes the electrode wires more susceptible to environmental noise [1]. This susceptibility to interference can also be reduced by using recording amplifiers with a high input impedance [2] or electrodes with small pre-amplifiers in them (i.e. “active electrodes”) [3]. There are various other circuit techniques to reduce the effect of interference on the measurement, such as shielding, guarding, or using a driven right leg circuit [3], [4]. Finally, interference such as power line noise can also be reduced after the measurement by digital filtering of the recording.

Biological artifacts are more difficult to prevent, if possible at all with the desired experiment paradigm. In studies where short trials are measured intermittently, the subject can be asked not to blink and sit still, however this may distract the subject or become uncomfortable. In most studies, the EEG is measured continuously for several minutes or hours, making artifacts such as eye blinks unavoidable. Measuring while the subject’s eyes are closed is a possible solution, however this may undesirably change EEG dynamics [5], and is obviously not possible if the experiment involves visual tasks. If the EEG measurement consists of several repeated trials, for example corresponding to multiple repetitions of the same stimulus, these trials can be averaged over time to enhance the neural response. While such processing partly suppresses noise and artifacts, it is still recommended to remove the trials with strong artifacts from the analysis. This results in a loss of data and may require longer measurement times to compensate. In single-trial EEG measurements where averaging over trials is not possible, artifacts are even more problematic as there is no redundancy of the EEG segments they corrupt.

Biological EEG artifacts can have different spatial, temporal and spectral characteristics. Many of the artifact removal methods developed and used over the years depend on assumptions that hold better for some artifact types than for others

(for a review, see [6], [7]). While there is no single golden standard in the EEG artifact removal literature, Independent Component Analysis (ICA) is very popular to remove eye blink and eye movement artifacts [6], [8]. On the other hand, Canonical Correlation Analysis (CCA) has become popular for its capability to remove muscle artifacts [9], [10]. A joint approach that targets both ocular and muscle artifacts has recently been proposed in [11]. These Blind Source Separation (BSS) methods are intrinsically semi-automatic, as the artifact components need to be selected for removal after source separation. However, numerous techniques for automating the artifact component selection based on features, statistics or similarities to a reference signal exist [12]–[14].

In this paper, we propose and validate a generic EEG artifact removal tool based on Multi-channel Wiener Filtering (MWF)¹. The MWF is a linear statistical filtering technique for estimating a signal from noisy measurements, and has been widely used for problems related to audio, speech, and recently also in biomedical data processing [15]–[19]. While this method has attracted some attention for EEG signal processing due to its low complexity and extendibility towards a distributed method [19], its use for EEG artifact removal has not been explored or validated in literature. Our MWF-based algorithm can simultaneously remove various types of artifacts from the EEG. The method is semi-automatic and semi-supervised, as it requires the user to mark a few examples of artifact segments in the time domain EEG data or in any other transform domain where the artifacts are more visible. The algorithm will use these examples to train an MWF that is then applied to the full EEG recording to automatically clean the EEG signals by filtering out the artifacts in each channel, i.e., also in those parts that were not manually annotated.

The possibility to annotate artifact segments gives the user *a-priori* control over which artifact types the algorithm should target, allowing the MWF to “learn” which signal components should be treated as artifacts. This is different for blind (BSS-based) algorithms, where, after a data-driven transformation, the user selects *post-hoc* a subset of output components that are believed to correspond to artifacts, which are then directly subtracted from the data. These output components are blindly generated by the algorithm, and their artifact separation capabilities strongly depend on how well the data matches the underlying source separation conditions on which the algorithm is based. Furthermore, as the number of EEG channels is generally smaller than the number of neural sources, the relevant EEG data is often distributed over these output components in a seemingly arbitrary fashion, such that artifact-dominated components may sometimes still contain a rather large EEG residue, which is then also subtracted from the data. In addition, the artifact(s) can be split over multiple output components (depending on the artifact type and the applied algorithm), in which case their deletion removes an even more significant portion of the clean EEG signal. The MWF-based algorithm on the other hand uses all its degrees of freedom to optimally remove the pre-defined artifact types

¹We provide an open-source implementation of the algorithm, with a graphical user interface supported by EEGLAB [8]. The implementation is available for download at www.github.com/exporl/mwf-artifact-removal.

while at the same time minimizing the removal of actual EEG components.

It is noted that the MWF-based method can be extended towards a fully automatic and unsupervised one using a proper artifact detection algorithm to identify artifact-dominated segments in the data, similar to how semi-automatic BSS methods can be automated. However, the development of such an algorithm falls outside the scope of this paper.

This paper is structured as follows. In Section II, the MWF-based artifact removal algorithm is described in detail. In Section III, the methodology for validating the algorithm is described. In Section IV, the results are provided, including a benchmark with traditional ICA and CCA-based methods. In Section V, a discussion of the achieved results is given. Finally, in Section VI, the conclusions are drawn.

II. MULTI-CHANNEL WIENER FILTER-BASED ARTIFACT REMOVAL

A. Data Model

The M -channel EEG signal $\mathbf{y}[t] \in \mathbb{R}^M$ at sample time t is modelled as

$$\mathbf{y}[t] = \mathbf{n}[t] + \mathbf{d}[t], \quad (1)$$

where $\mathbf{n}[t] \in \mathbb{R}^M$ represents the true neural signals, and $\mathbf{d}[t] \in \mathbb{R}^M$ represents artifacts of other origin superimposed on the neural signals. Note that $\mathbf{d}[t]$ may simultaneously contain different types of artifacts, e.g., it can include eye blink artifacts in combination with eye movement and muscle artifacts. The signals $\mathbf{y}[t]$, $\mathbf{n}[t]$ and $\mathbf{d}[t]$ are treated as M -dimensional stochastic vectors, which are observed at specific sample times $t \in \mathbb{N}$. For the sake of conciseness, the sample index t will be omitted, unless the explicit indication of the sample time is necessary.

The signals \mathbf{y} , \mathbf{n} and \mathbf{d} are assumed to be zero-mean. This assumption is often satisfied for EEG signals, as mean subtraction is a common first preprocessing step. The signal covariance matrices \mathbf{R}_{yy} , \mathbf{R}_{nn} , \mathbf{R}_{dd} are defined as resp. $E\{\mathbf{y}\mathbf{y}^T\}$, $E\{\mathbf{n}\mathbf{n}^T\}$ and $E\{\mathbf{d}\mathbf{d}^T\}$, where $E\{\cdot\}$ is the expected value operator. If \mathbf{n} and \mathbf{d} are assumed to be uncorrelated, then

$$\mathbf{R}_{yy} = \mathbf{R}_{nn} + \mathbf{R}_{dd}. \quad (2)$$

For the sake of an easy exposition, we first present the MWF as a purely spatial filter, i.e., an instantaneous linear combiner, and later generalize it to a convolutive filter-and-sum operator using finite impulse responses (see Subsection II-C).

B. Multi-channel Wiener filter (MWF)

The MWF produces an estimate $\hat{\mathbf{d}}$ of the multi-channel artifact signal \mathbf{d} by linearly combining the channels of \mathbf{y} , i.e. $\hat{\mathbf{d}} = \mathbf{W}^T \mathbf{y}$. This linear combination is optimized in a minimum mean squared error (MMSE) sense, i.e., it optimizes the objective

$$\min_{\mathbf{W}} E\{\|\mathbf{d} - \mathbf{W}^T \mathbf{y}\|^2\}. \quad (3)$$

The i^{th} column of \mathbf{W} is the linear combiner used to estimate the i^{th} channel of \mathbf{d} . The solution to this minimization problem is [15], [20]

$$\mathbf{W} = \mathbf{R}_{yy}^{-1} \mathbf{R}_{yd}, \quad (4)$$

where $\mathbf{R}_{yd} = E\{\mathbf{y}\mathbf{d}^T\}$. Using (1) and the assumed uncorrelatedness between \mathbf{d} and \mathbf{n} , \mathbf{R}_{yd} can be shown to equal \mathbf{R}_{dd} , i.e.,

$$\mathbf{R}_{yd} = \mathbf{R}_{(n+d)d} = \mathbf{R}_{nd} + \mathbf{R}_{dd} = \mathbf{R}_{dd}, \quad (5)$$

so that the filter solution from (4) can be written as

$$\mathbf{W} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{dd}. \quad (6)$$

In practice, the covariance matrices can be estimated from an $M \times T$ observation matrix \mathbf{Y} in which the t^{th} column contains the observation $\mathbf{y}[t]$ at time t . Using an artifact detection method (see Section III-C), the T observations of \mathbf{Y} can be segmented into two mutually exclusive sets: \mathbf{Y}_a containing the T_a samples with artifacts, and \mathbf{Y}_c containing the T_c clean samples without artifacts. This segmentation allows \mathbf{R}_{yy} to be estimated as

$$\hat{\mathbf{R}}_{yy} = \frac{1}{T_a} \mathbf{Y}_a \mathbf{Y}_a^T, \quad (7)$$

where the hat denotes that it is an estimation. Similarly, \mathbf{R}_{nn} can be estimated from the $M \times T_c$ observation matrix \mathbf{Y}_c containing only samples without artifacts, as

$$\hat{\mathbf{R}}_{nn} = \frac{1}{T_c} \mathbf{Y}_c \mathbf{Y}_c^T. \quad (8)$$

Based on (2), an estimate for \mathbf{R}_{dd} can then be computed as

$$\hat{\mathbf{R}}_{dd} = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{nn}, \quad (9)$$

such that the optimal filter solution \mathbf{W} from (6) can be estimated using only covariance matrix estimates from the data as

$$\hat{\mathbf{W}} = \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{dd}. \quad (10)$$

Finally, using the additive model of (1), the neural responses can be obtained by subtracting the estimated artifact as

$$\hat{\mathbf{n}} = \mathbf{y} - \hat{\mathbf{W}}^T \mathbf{y}. \quad (11)$$

Note that it is also possible to estimate the clean EEG signal \mathbf{n} directly by replacing \mathbf{d} by \mathbf{n} in the MMSE problem in (3). This results in the filter $\mathbf{W} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{nn}$, which will yield the same clean signal estimate $\hat{\mathbf{n}}$ as in (11). However, we purposely estimate the artifact to exploit the typical low-rank structure of the artifact covariance matrix \mathbf{R}_{dd} (see Section II-D).

C. Including temporal/spectral information

Thus far, the MWF was defined as a purely spatial filter, i.e., it only makes a linear combination of the different channels to estimate the artifact. This exploits the different spatial distributions of the underlying sources (similar to ICA). However, in general the spectral properties of the artifacts also differ from the clean EEG spectrum, and their spectrum may also differ across channels. For example, artifacts that vary slowly (relative to background EEG) such as eye blinks or eye movement still have large autocorrelation coefficients after tens of milliseconds [21]. On the other hand, random, high-frequency artifacts such as muscle activity result in very low autocorrelations at short time lags [9]. To also exploit the

different spectral properties of the sources, the MWF can be extended by allowing it to design and perform a finite impulse response (FIR) filtering in each channel, which acts as a per-channel spectral filter. To this end, we include time-lagged versions of each channel, using both positive and negative time shifts up to lag τ , in the vector $\mathbf{y}[t]$ in (1), such that it becomes

$$\mathbf{y}[t] = \begin{bmatrix} y_1[t + \tau], \dots, y_1[t], \dots, y_1[t - \tau], \\ y_2[t + \tau], \dots, y_2[t], \dots, y_2[t - \tau], \dots \\ y_M[t + \tau], \dots, y_M[t], \dots, y_M[t - \tau] \end{bmatrix}^T. \quad (12)$$

This can be easily realized by stacking time-shifted versions of the observation matrix \mathbf{Y} along with the original one, i.e.

$$\tilde{\mathbf{Y}} = [\mathbf{Y}_{-\tau}^T, \dots, \mathbf{Y}_0^T, \dots, \mathbf{Y}_\tau^T]^T, \quad (13)$$

where \mathbf{Y}_τ represents \mathbf{Y} shifted over τ time lags (with proper zero-padding to match the original dimensions of \mathbf{Y}). An artifact detection method can be used to segment $\tilde{\mathbf{Y}}$ into segments with and without artifacts as in Subsection II-B. The MWF can then be computed using (10) and applied to the EEG data in $\tilde{\mathbf{Y}}$ (which includes the time-lagged versions).

The temporal information complements the spatial information captured in the covariance matrices, which also characterizes the artifact sources in some sense. For example, eye blink artifacts show higher amplitudes in frontal channels, and muscle artifacts appear in all channels but most prominently on the sides of the head. The artifact estimation performance of the MWF increases by including temporal correlations in addition to the spatial correlations, effectively making it a spatio-temporal filter.

D. GEVD-based MWF

As the number of underlying artifact sources (say Q) is usually much smaller than the number of channels M , the covariance matrix \mathbf{R}_{dd} typically has low rank. For the sake of an easy exposition, but without loss of generality, we focus on the case where no time lags are included in the MWF², in which case \mathbf{R}_{dd} will have exactly rank Q . However, since in practice the covariance matrices are estimated from noisy observations, the rank of the estimated matrix $\hat{\mathbf{R}}_{dd}$ is generally equal to M instead of Q . Additionally, because of the subtraction $\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{nn}$, the positive semi-definiteness of $\hat{\mathbf{R}}_{dd}$ is not guaranteed [22].

Performing the subtraction based on a Generalized Eigenvalue Decomposition (GEVD) allows to force $\hat{\mathbf{R}}_{dd}$ to be positive semi-definite and of low rank, which has been shown to improve the estimation performance of the MWF [17]. Since \mathbf{R}_{yy} and \mathbf{R}_{nn} are both symmetric and positive definite, an invertible matrix \mathbf{V} can be found such that

$$\begin{aligned} \mathbf{V}^T \mathbf{R}_{yy} \mathbf{V} &= \text{diag}(\sigma_{y1}, \dots, \sigma_{yM}) = \Sigma_y \\ \mathbf{V}^T \mathbf{R}_{nn} \mathbf{V} &= \text{diag}(\sigma_{n1}, \dots, \sigma_{nM}) = \Sigma_n, \end{aligned} \quad (14)$$

²The reasoning can be generalized to the case where time lags are included, in which case \mathbf{R}_{dd} still has low-rank properties. Its rank will then depend on the autocorrelation function of each source, as well as the spectral filtering introduced by volume conduction.

while also satisfying

$$\mathbf{R}_{yy}\mathbf{V} = \mathbf{R}_{nn}\mathbf{V}\mathbf{\Lambda}, \quad (15)$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_M)$ and $\lambda_i = \sigma_{yi}/\sigma_{ni}$ for $i = 1, \dots, M$ [23]. Equation (15) is known as the GEVD, in which the columns of \mathbf{V} correspond to the Generalized Eigenvectors (GEVC), and the diagonal elements of $\mathbf{\Lambda}$ correspond to the Generalized Eigenvalues (GEVL) of the ordered matrix pair $(\mathbf{R}_{yy}, \mathbf{R}_{nn})$. Computing the GEVD is equivalent to finding the joint diagonalization of \mathbf{R}_{yy} and \mathbf{R}_{nn} . We assume, without loss of generality, that the GEVLs in $\mathbf{\Lambda}$ are sorted in descending order, i.e. $\lambda_1 > \dots > \lambda_M$.

The diagonalization in (14) allows us to write \mathbf{R}_{dd} as

$$\begin{aligned} \mathbf{R}_{dd} &= \mathbf{R}_{yy} - \mathbf{R}_{nn} \\ &= \mathbf{V}^{-T}\mathbf{\Sigma}_y\mathbf{V}^{-1} - \mathbf{V}^{-T}\mathbf{\Sigma}_n\mathbf{V}^{-1} \\ &= \mathbf{V}^{-T}(\mathbf{\Sigma}_y - \mathbf{\Sigma}_n)\mathbf{V}^{-1} \\ &= \mathbf{V}^{-T}\mathbf{\Sigma}_d\mathbf{V}^{-1}, \end{aligned} \quad (16)$$

where $\mathbf{\Sigma}_d = \text{diag}(\sigma_{d1}, \dots, \sigma_{dM})$ and $\sigma_{di} = \sigma_{yi} - \sigma_{ni}$. Since \mathbf{R}_{dd} has rank Q , only the first Q diagonal elements of $\hat{\mathbf{\Sigma}}_d$ will be non-zero. However, as $\hat{\mathbf{R}}_{dd}$ generally doesn't have rank Q , the GEVD can be used to compute a rank- Q approximation of \mathbf{R}_{dd} . To this end, we compute the GEVD of $(\hat{\mathbf{R}}_{yy}, \hat{\mathbf{R}}_{nn})$, and compute (16) where the last $M - Q$ diagonal elements of $\mathbf{\Sigma}_d$ are replaced by zeros [17].

In practice, Q is generally unknown, in which case there are several approaches to set the rank of $\hat{\mathbf{R}}_{dd}$. One can make an a-priori estimate based on prior knowledge of the artifact type. For example, if only eye blink artifacts are annotated, Q can be chosen equal to 1 (multiplied by the number of time lags used in the MWF), as they are well approximated by a single source point. Another approach is to observe the diagonal entries of $\hat{\mathbf{\Sigma}}_d$ and to define a threshold under which the diagonal entries should be set to zero. The most conservative approach is to only set the negative entries in $\hat{\mathbf{\Sigma}}_d$ to zero, in order to avoid using a covariance matrix estimate that is not positive semi-definite.

It is noted that there exist other approaches to make rank- Q approximations of $\hat{\mathbf{R}}_{dd}$ which are not based on a GEVD, e.g. by computing a truncated singular value or eigenvalue decomposition on $\hat{\mathbf{R}}_{dd} = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{nn}$ [17]. However, the GEVD-based approach typically yields a better performance, as it explicitly takes the underlying theoretical relationships between \mathbf{R}_{yy} , \mathbf{R}_{nn} and \mathbf{R}_{dd} into account. Furthermore, it can be shown that a GEVD-based low-rank approximation of \mathbf{R}_{dd} constrains the MWF to the subspace that captures the Q highest SNR-modes in the data [17], [22]. This is a desirable property, which is also exploited in several other GEVD-based algorithms, such as denoising source separation³ [24]–[26], the xDAWN algorithm⁴ [27], common spatial pattern analysis [29], etc.

³The original DSS algorithm in [24] was not formulated based on a GEVD, although its link with the GEVD was established in [25].

⁴The original xDAWN algorithm in [27] was based on a QR decomposition and a singular value decomposition, although it was later shown to be an instance of a GEVD filter in [28].

III. EXPERIMENT

A. EEG data acquisition

An EEG dataset with multiple recordings per subject was obtained, in which each recording contains a different artifact type, e.g. eye blink artifacts, muscle artifacts, movement artifacts, etc. These artifacts were induced purposefully by the subject by e.g. blinking the eyes, clenching the jaws or tilting the head after a visual cue, which occurred after random time intervals with a mean of 5 seconds. The artifact recordings, in addition to the induced artifacts, contain spontaneous eye blink artifacts, since these are impractical to prevent.

A BioSemi ActiveTwo amplifier with 64 electrodes in the international 10-20 electrode placement configuration was used for all recordings, with electrode Cz on top of the head serving as EEG reference. Data was recorded at 8192 Hz, and resampled to 200 Hz for offline processing. A high-pass filter with cut-off frequency at 1 Hz was applied to remove the DC component and electrode drift from the signals.

B. Generation of hybrid EEG data

EEG artifact removal performance is generally difficult to evaluate because the real artifact signal and clean EEG are unknown. Therefore, to complement the data set described in Section III-A, an additional data set is designed in which (known) eye blink artifacts are superimposed onto real EEG data with various Signal-to-Noise Ratios (SNRs). Because the ground truth signals are known, the artifact estimate can be directly compared to the actual artifact that was added.

To this end, a single-channel artifact signal d is generated by repeating an eye blink artifact template at random time intervals with an average of 5 seconds between consecutive eye blinks. The constructed single-channel signal d is then multiplied with a pre-defined mixing vector \mathbf{a} with realistic mixing weights to create the ground truth multi-channel artifact signal \mathbf{d} , which is added to \mathbf{n} according to (1), where \mathbf{n} consists of artifact-free EEG data from the same 10 subjects as in Section III-A. Note that the covariance matrix \mathbf{R}_{dd} of the artifacts added to the hybrid EEG data has a rank equal to exactly 1 if no time lags are used.

The generation of hybrid EEG data allows to investigate the effect of the artifact amplitude relative to the background EEG on algorithm performance. A factor γ is used to scale the artifacts \mathbf{d} to generate EEG data with different artifact SNR values. We define the artifact SNR as the ratio of artifact power and the clean EEG power [21], i.e.

$$\text{SNR} = 10 \log_{10} \frac{E\{(\gamma d_1)^2\}}{E\{(n_1)^2\}}, \quad (17)$$

where channel 1 (Fp1) is chosen as the reference. This is an arbitrary choice in order to define an SNR measure, but preferably a channel with large artifact amplitude is chosen. Because the clean EEG power is recording- and subject-dependent, the artifact signal \mathbf{d} is scaled such that the SNR is 0 dB for $\gamma = 1$. This normalizes the SNRs of the hybrid data so that comparisons across subjects can be made. An SNR increase or decrease by 6 dB indicates a doubling or halving of the artifact amplitudes.

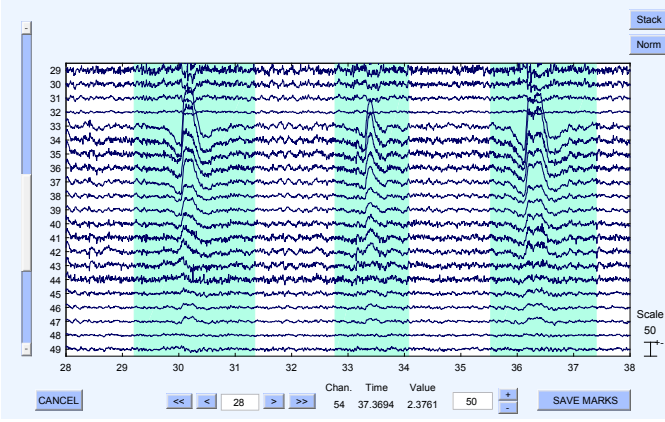


Fig. 1. Example of eye blink artifact marking in the GUI. The user can change display options for easier assessment of the EEG data. After marking, clicking the “Save Marks” button will run the MWF for artifact removal.

C. Marking of EEG artifacts

We implemented the MWF in a user-friendly graphical user interface (GUI), in which the EEG channels are visualized using EEGLAB [8], allowing the user to assess spatial and temporal characteristics of features in the EEG. An example of use of the GUI to mark eye blink artifacts is shown in Fig. 1. Once the users has marked some examples of segments containing undesired artifacts, a button click exports the time indices of the markings to MATLAB for further processing by the MWF algorithm. The markings are also saved to disk along with the recordings so they can be reused for future experiments without the user having to re-assess the data.

While marking the artifacts in the EEG, it is important to avoid starting or ending an artifact segment in the middle of an artifact. Failing to capture the entire artifact leads to presence of (parts of) artifacts in the clean EEG segments, which causes errors in the estimation of $\hat{\mathbf{R}}_{nn}$ in (8). Conversely, overestimating the length of the artifacts does not adversely affect MWF performance, because clean EEG still adheres to the model in (1), with the artifact term \mathbf{d} (temporarily) being equal to zero. Additionally, the low power of clean EEG relative to artifacts has a negligible effect on the estimation of $\hat{\mathbf{R}}_{yy}$ in (7). Because of these reasons, we chose to encompass the artifact with broad margins while marking the EEG (see Fig. 1).

D. Performance measures

The quality of the artifact removal is assessed using two complementary performance measures. The first measures the distortion of the clean EEG, and the second measures the degree of removal of the artifact. The first measure is computed over artifact-free segments. Ideally, applying the filter to these segments does not change them, since the artifact to be estimated and subtracted is zero. In fact, the estimated artifact signal in the artifact-free segments of channel i , given by \hat{d}_i , should be as close to zero as possible. To assess this, we use the Signal-to-Error Ratio (SER) [19], [21] in a single

channel i , computed as

$$\text{SER}_i = 10 \log_{10} \frac{E\{(y_i)^2\}}{E\{(\hat{d}_i)^2\}} \Big|_{\text{clean segments}}. \quad (18)$$

To obtain a single measure, the SERs of individual channels are combined by weighted averaging over channels as

$$\text{SER} = \sum_{i=1}^M p_i \cdot \text{SER}_i, \quad (19)$$

where the normalized weights p_i are proportional to the artifact power in every channel, which can be estimated by subtracting the power in the clean segments from the power in the artifact segments:

$$p_i = \frac{E\{(y_i)^2\} \Big|_{\text{artifact segments}} - E\{(y_i)^2\} \Big|_{\text{clean segments}}}{\sum_{i=1}^M \left(E\{(y_i)^2\} \Big|_{\text{artifact segments}} - E\{(y_i)^2\} \Big|_{\text{clean segments}} \right)}. \quad (20)$$

This weighting ensures that the channels with the largest artifacts have the greatest contribution to the overall performance measure. Note that, while the SER is computed only over artifact-free segments, it is also a measure for the distortion of the clean EEG in the artifact segments, but it cannot be measured there.

A second performance measure is computed over the artifact segments. The artifact estimate \hat{d}_i should closely resemble the real, underlying artifact signal, i.e. the residue $d_i - \hat{d}_i$ should be as small as possible. This leads to the Artifact-to-Residue Ratio (ARR) [21] in a single channel i , computed as

$$\text{ARR}_i = 10 \log_{10} \frac{E\{(d_i)^2\}}{E\{(d_i - \hat{d}_i)^2\}} \Big|_{\text{artifact segments}}. \quad (21)$$

Only for the hybrid EEG data, d_i is known. For real data, the ARR is computed by substituting d_i by y_i in (21). This approximation is valid in artifact segments for high amplitude artifacts (such as eye blinks, muscle activity,...). A weighted average across channels to obtain the ARR as a single measure is employed in a similar fashion as in (19). The weights p_i are identical as those defined in (20).

SER and ARR are both expressed in dB. Good performance of the artifact removal algorithm is indicated by both high SER and high ARR. The two performance measures reflect two complementary aspects of artifact removal and should always be evaluated simultaneously. Indeed, note that the SER can easily be made infinitely high by using the trivial all-zero filter for \mathbf{W} in (11), but then the ARR will reduce to 0 dB.

E. Effect of temporal information

As discussed in Section II-C, including several lagged versions of the EEG data in the filter can increase estimation performance. For an EEG sample rate of 200 Hz, channels can be shifted with a resolution of 5 ms. To investigate the effect of the temporal information, the number of lags included in the observation matrix in (13) is gradually increased. Note that an increase of τ by 1 increases the number of entries in \mathbf{y} by $2M$ due to both positive and negative time lags being

included. This increase in channels for increasing τ affects the computational cost of the algorithm.

F. Effect of the rank approximation of the GEVD-MWF

As discussed in Section II-D, a GEVD-based low-rank positive semi-definite estimate of \mathbf{R}_{dd} may improve the performance of the MWF. However, choosing an appropriate rank Q is not trivial. First, the conservative strategy, where only the negative entries of $\hat{\Sigma}_d$ are set to zero, is applied and compared to the (full-rank) MWF performance.

For low-dimensional artifacts (e.g. eye blinks), we hypothesize better artifact estimation when the rank of $\hat{\mathbf{R}}_{dd}$ is reduced further, so that it matches the low-dimensional structure of the artifact. We will investigate this effect by lowering the amount of retained eigenvalues for eye blink artifacts. The performance is also compared with the same strategy applied to muscle artifacts, which typically have a higher dimension due to multiple myogenic sources.

G. Types of artifacts

The MWF is a generic method and can also be applied to remove other artifacts than eye blink or muscle artifacts. EEG recordings are often affected by glitches or spikes originating from a sudden change in skin-electrode contact impedance. This phenomenon is usually caused by the subject moving the head, causing electrodes to shift around. We will refer to this as a “movement” artifact, which manifests as a very high-amplitude peak that typically affects only a single channel or a few neighboring channels. The morphology of these artifacts can differ greatly (unlike e.g. eye blinks, which are repetitive for a subject). The MWF will be applied to remove movement artifacts by marking a single movement artifact in the EEG recording, demonstrating the capability to remove artifacts with focused spatial and sparse temporal structure (i.e. appearing only once in a few channels in the recording).

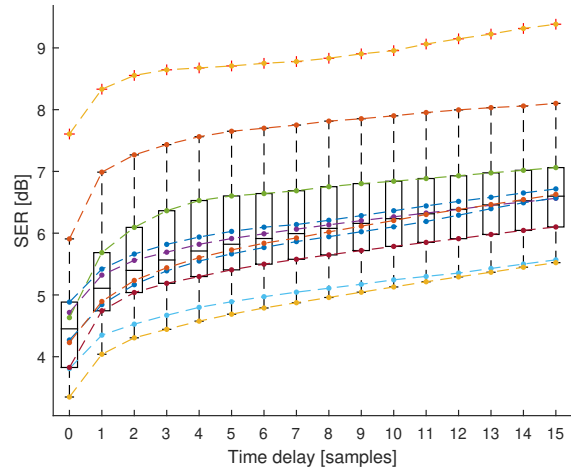
H. Comparison with other artifact removal algorithms

The MWF is compared to several other methods that are popular for EEG artifact removal. The MWF was applied to the EEG recordings with eye blink artifacts, and to the EEG recordings containing both eye blink and muscle artifacts. The same recordings are processed by three component-based methods: infomax ICA [8], [30], which is implemented in EEGLAB, the FastICA algorithm [31] from the FastICA MATLAB toolbox, and CCA [9], [21]. Artifact removal is performed semi-automatically by selecting artifact components to reject after visualization. After rejection, the signals are projected back to the EEG sensor space, resulting in the clean EEG signals. The same (manual) EEG segmentation as used for the MWF is used to compute SER and ARR performance parameters for each method.

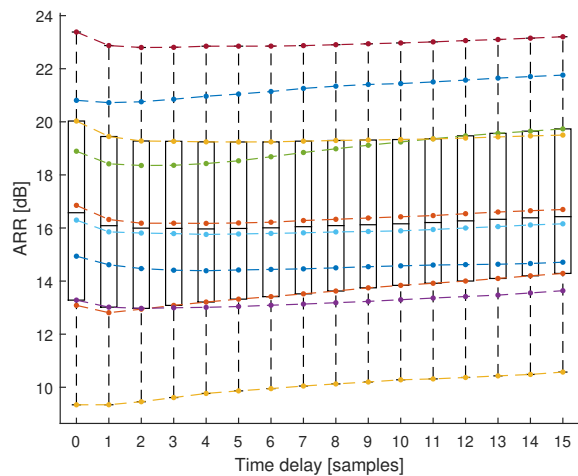
IV. RESULTS

A. Choice of time lags

The effect of increasing the amount of delayed signal versions included in the artifact estimation is depicted in



(a) The SER increases as more time delays are used in the MWF. The increase is the largest for the first few delays.



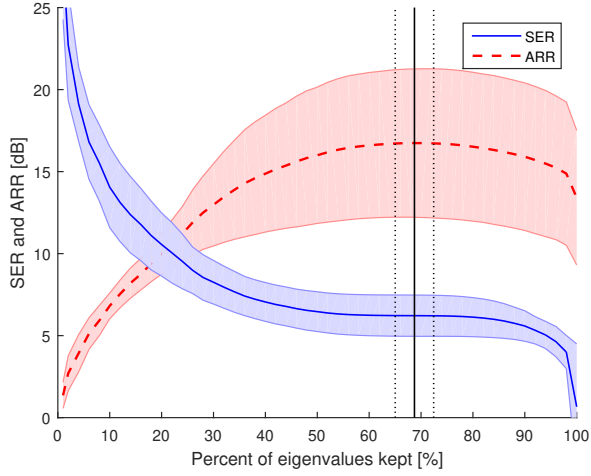
(b) The ARR is invariant to the number of delays used in the MWF.

Fig. 2. SER and ARR for combined eye blink and muscle artifact removal. Both performance measures are plotted as a function of the number of time delays used in the MWF. Dashed lines show trends for individual subjects.

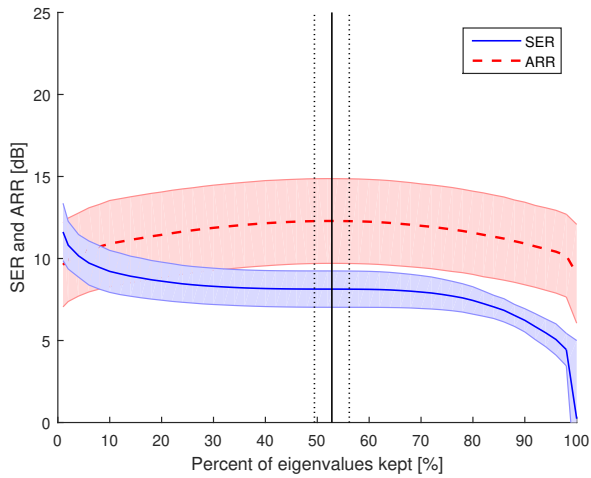
Fig. 2. For $\tau = 0$, no delayed versions are included. For $\tau > 0$, all delayed versions with delays $-\tau, \dots, 0, \dots, \tau$ are included. In this experiment, the MWF is computed by setting only the negative eigenvalues to zero in $\hat{\Sigma}_d$ in (16). The filter is trained on and applied to EEG data of 10 subjects containing both muscle and eye blink artifacts. Both performance measures SER and ARR are depicted on a group level by boxplots in Fig. 2(a) and 2(b), respectively. The dashed lines show the performance measure trends for individual subjects.

B. Choice of rank

The effect of replacing the estimated artifact covariance matrix \mathbf{R}_{dd} with a GEVD-based lower rank approximation is depicted in Fig. 3(a) for the dataset of 10 subjects containing both muscle and eye blink artifacts. The performance measures SER and ARR are shown as a function of the percentage of retained eigenvalues, i.e. the percentage of diagonal entries in $\hat{\Sigma}_d$ not replaced by zero. If 100% of eigenvalues are retained,



(a) Combined muscle and eye blink artifact removal.



(b) Eye blink artifact removal.

Fig. 3. Performance measures as a function of the rank of the artifact covariance matrix approximation. The 100% point corresponds to the regular MWF where no eigenvalues are rejected. For both figures, 5 time delays ($\tau = 5$) were included in the computation of the MWF. The optimal percent of eigenvalues to keep corresponds well with the percentage that would be obtained by only keeping positive eigenvalues (indicated by the black vertical lines).

the GEVD-based MWF reduces to the regular MWF. The solid lines depict the mean across subjects, the shaded area width indicates the standard deviation.

As discussed in Section III-F, a conservative approach for deciding which eigenvalues to keep, is to put only the negative ones to zero. The vertical black lines in Fig. 3(a) show what percentage of retained eigenvalues that approach results in. The solid black line depicts the mean percentage across subjects, and the dotted lines depict the standard deviation. Fig. 3(b) shows results of the same analysis performed on the dataset containing only eye blink artifacts.

Fig. 4 shows a comparison of the distribution of GEVLs for the data containing only eye blink artifacts, and the data containing eye blink and muscle artifacts, for one subject.

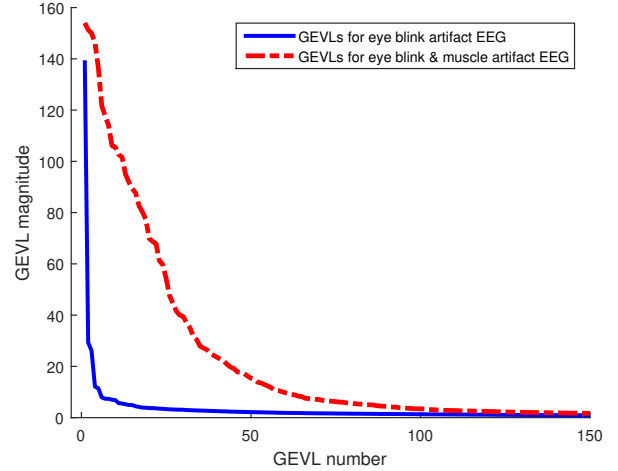


Fig. 4. Comparison of GEVL distributions for two different artifact types from a representative subject. Only the first 150 GEVLs are shown (out of 693 GEVLs, 63 channels with $\tau = 5$). The GEVLs are ordered by descending magnitude. Eye blink artifacts exhibit a narrower distribution than muscle artifacts, which results in filters of lower rank if only eye blink artifacts are targeted.

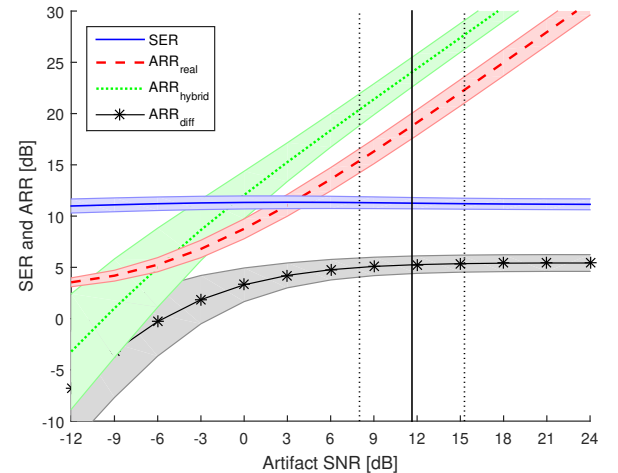
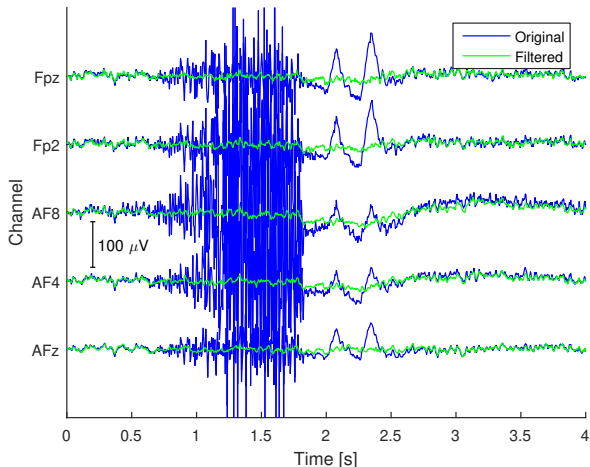


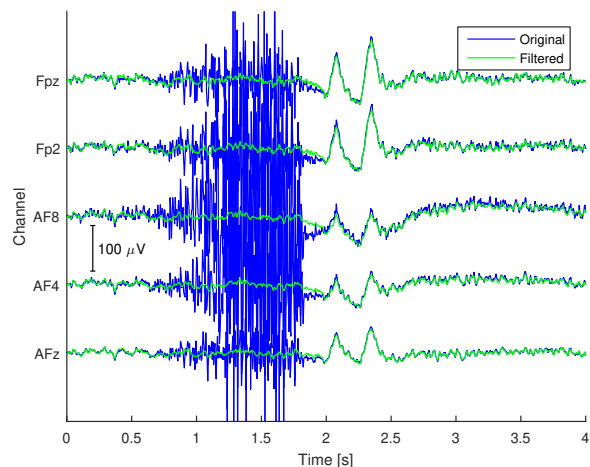
Fig. 5. Effect of artifact SNR on MWF performance for hybrid EEG data with eye blink artifacts. The solid and dashed vertical black lines depict resp. the mean and standard deviation of the estimated artifact SNR in real EEG data. Whereas the SER is invariant to artifact SNR, the ARR increases with increasing SNR. Because the difference between ARR_{hybrid} and ARR_{real} is constant above 3 dB, ARR_{real} can be used as an approximation of ARR_{hybrid} if the ground truth artifact is unknown.

C. Hybrid EEG data with eye blink artifacts

Fig. 5 shows the result of applying the MWF for eye blink artifact removal from the hybrid EEG data as a function of artifact SNR as defined in Section III-B. The data is processed with the MWF using 5 time delays ($\tau = 5$) and only positive eigenvalues were retained in the low-rank approximation. The solid lines depict the mean across subjects, the shaded area width equals the standard deviation. The ARR is shown computed in two ways, as for the hybrid EEG data the ground truth artifact \mathbf{d} is known. ARR_{hybrid} is computed using (21), while ARR_{real} is computed using (21) but replacing \mathbf{d} by \mathbf{y} .



(a) Combined muscle and eye blink artifact removal by marking both muscle and eye blink artifacts.



(b) Muscle artifacts removal by marking only muscle artifacts and not marking eye blink artifacts.

Fig. 6. Original and MWF-filtered data are shown for 5 channels on the frontal right side of the head. A strong muscle artifact is present from 0.5 to 2 s, followed by two consecutive eye blinks. Only the artifacts marked by the user are removed.

The difference $ARR_{\text{diff}} = ARR_{\text{hybrid}} - ARR_{\text{real}}$ is shown as well.

The artifact SNR for the real eye blink data of all subjects is estimated and shown on Fig. 5. The solid black line depicts the mean artifact SNR across subjects, and the dotted lines depict the standard deviation across subjects.

D. Performance for different artifact types

Fig. 6(a) shows an example of simultaneous muscle and eye blink artifact removal with the MWF ($\tau = 5$, only positive eigenvalues) after marking all eye blink and muscle artifacts. Fig. 6(b) shows the same EEG segment, but processed after marking only the segments containing muscle artifacts. As a result, the eye blink artifacts remain unaffected by the MWF, whereas the muscle artifacts are removed. Fig. 7 shows an example of a movement artifact affecting only a few channels

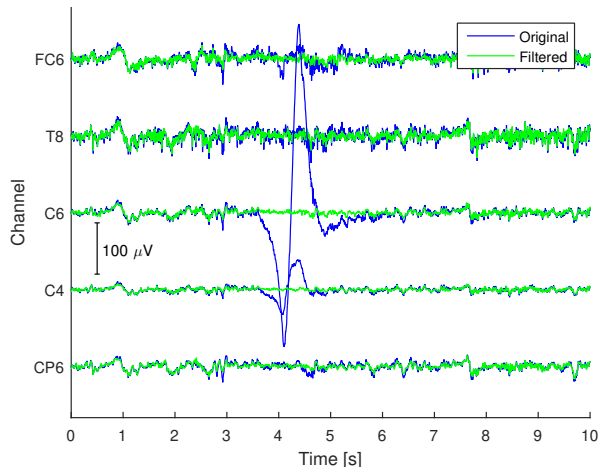


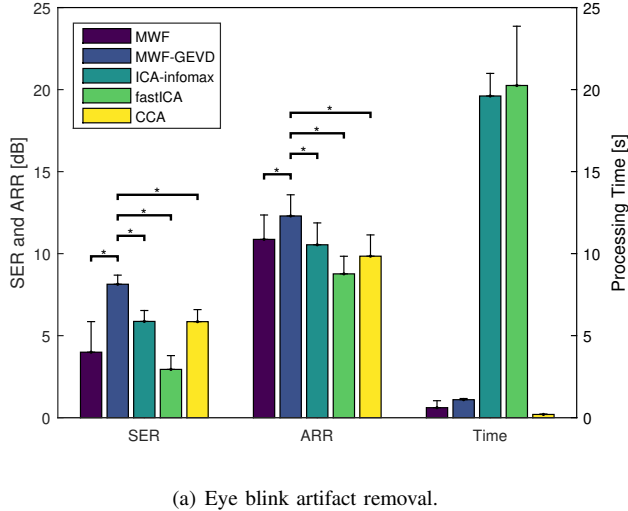
Fig. 7. Removal of a movement artifact. Original and MWF-filtered data are shown for 5 neighboring channels on the central right side of the head. The artifact is most strongly present in electrode C6 and lasts multiple seconds.

and lasting several seconds. The EEG processed with the MWF ($\tau = 5$, only positive eigenvalues), after marking only this specific instance of artifact, is shown on top of the original data.

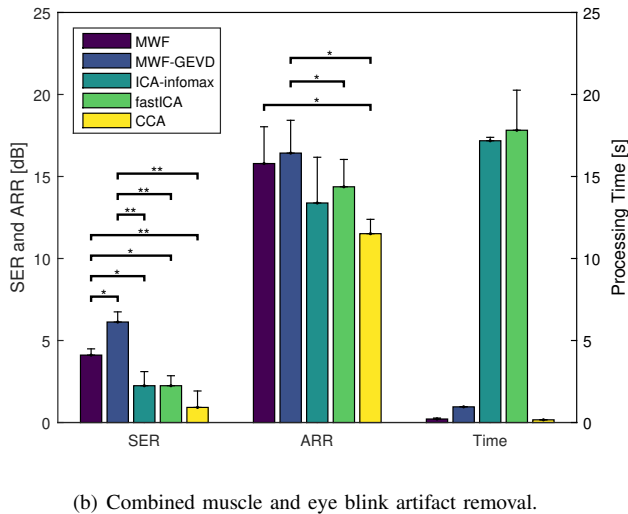
E. Comparison with other artifact removal techniques

Fig. 8(a) and 8(b) show the results of applying the different artifact removal algorithms to resp. the eye blink data and the eye blink + muscle artifact data. The mean and standard deviation of the performance measures across subjects are shown, as well as the computation time required for artifact removal (not including time taken by manual artifact marking or component selection). MWF corresponds to the basic MWF as described in Section II-B. MWF-GEVD is the improved version using both temporal information ($\tau = 5$) and GEVD-based low-rank approximation, retaining only positive eigenvalues. For infomax ICA, the number of ICA training steps was limited to 256, which was generally enough to converge and reduced computation time. For FastICA, the maximum allowed number of iterations per component was limited to 250 for the same reasons. For CCA, the canonical components were computed between $\mathbf{y}[t]$ and $\mathbf{y}[t - 1]$ (i.e. shifted over one time lag of 5 ms). When processing the eye blink artifact EEG data, only segments containing eye blinks (and other ocular artifacts) where marked for the MWF, and only independent or canonical components containing the same artifacts were selected for removal. For the data containing both muscle artifacts and eye blink artifacts, all artifact segments containing either eye blinks (and other ocular artifacts) and muscle activity were selected.

Group median differences between all methods and the MWF-based methods are compared using a two-tailed Wilcoxon signed rank test. The p-values are corrected for multiple comparisons using the Bonferroni-Holm method per dataset and performance measure. Significant differences with a p-value exceeding the significance level $\alpha = 0.05$ are indi-



(a) Eye blink artifact removal.



(b) Combined muscle and eye blink artifact removal.

Fig. 8. Performance comparison of different artifact removal methods on EEG datasets containing only eye blink artifacts or both eye blink and muscle artifacts. Mean and standard deviation of performance measures are indicated by bar height and errorbar, respectively. Good performance is indicated by high SER and ARR. Significant differences with the MWF-based methods are indicated single ($p < 0.05$) or double ($p < 0.01$) asterisks, with all p-values corrected for multiple comparisons using the Bonferroni-Holm method. In addition to SER and ARR, the computation times per subject are shown.

cated by a single asterisk, and p-values exceeding $\alpha = 0.01$ are indicated by double asterisks.

V. DISCUSSION

A. Selection of EEG artifacts

The MWF is integrated with a user-friendly GUI that allows manual marking of undesired artifacts in the EEG. The user can scroll through the EEG data and mark any temporal segment that is identified as “not clean”, and have it removed by the MWF. This is advantageous, since often EEG artifacts are encountered that are clearly not of neural origin, but whose origin is not known. Another advantage of this method is that it facilitates marking of artifacts in the raw EEG. Other methods like ICA or CCA require identification of artifact components

in the transformed component space, which needs additional topographic plotting in order to assess the spatial distribution of a component. More importantly, using component-based methods, the user fully depends on the algorithm to blindly isolate the artifacts from the clean EEG in a single component, which is not always the case, e.g. artifacts can be mixed with clean EEG in a single component.

The method is suited for removal of any kind of artifact that exhibits a temporal on-off pattern, which is demonstrated for common EEG artifacts such as eye blinks, muscle artifacts and movement artifacts. They can be removed independently by only marking artifacts of a single kind, despite the presence of other artifacts (see Fig. 6(b)), or simultaneously, as demonstrated for eye blink and muscle artifacts (see Fig. 6(a)). Even a single glitch or movement artifact can be marked as artifact and removed⁵ (see Fig. 7).

In the particular case where the artifact only appears in a small number of channels (as in Fig. 7), the MWF will mainly rely on information from neighboring channels to reconstruct the clean EEG, i.e. similar results could be obtained by training an MWF only on channels adjacent to the affected channels. However, making this channel selection is unnecessary, as the MWF automatically infers which channels are the relevant ones to remove the artifacts. The GUI therefore only allows the user to select time segments, but not channels.

B. Choice of parameters for GEVD-MWF

The number of time lags and the rank approximation of the GEVD-MWF are the most important parameters to consider. By increasing the number of time lags, more degrees of freedom are available in the MWF design. This does not change the artifact suppression performance (Fig. 2(b)), but it does reduce the distortion of the clean EEG, indicated by an increasing SER (Fig. 2(a)). The effect of including temporal information is largest for the first few delays included, and the gain decreases when more time lags are included. The gain of more temporal information must also be weighed against the increase in computational cost.

By replacing the artifact covariance matrix estimate $\hat{\mathbf{R}}_{dd}$ with a lower rank approximation, the effect of covariance estimation errors can be reduced (i.e. $\hat{\mathbf{R}}_{dd}$ not being positive semi-definite or having an increased rank). Fig. 3 shows the effect of the chosen rank on both performance measures. Rejecting the lowest 5-10% of GEVLs significantly increases both SER and ARR (see the right part of Fig. 3(a) and 3(b)). Including these modes makes the artifact estimate more noisy, resulting in poorer artifact estimation (ARR) and more EEG distortion (SER). After this initial increase in performance, there is a range of percentages (50-90%) for which the performance is relatively stable. The algorithm performance in case only positive eigenvalues are retained also falls within this range, seemingly at the maximal ARR. If more eigenvalues are

⁵It is noted that, if an artifact with a single occurrence is annotated together with a large number of other artifacts, the contribution of the former to the covariance matrix $\hat{\mathbf{R}}_{yy}$ will be minor and the resulting MWF may not effectively remove it. In this case, it could be beneficial to design two separate MWFs, i.e., one for the glitch and one for the more frequently appearing artifacts.

rejected, the trade-off effect between SER and ARR becomes apparent. Modes containing artifact components are rejected, which causes worse artifact estimation, but also less noise on the artifact estimate, resulting in lower clean EEG distortion. As mentioned in Section III-D, this trade-off results in an infinite SER and zero ARR when all eigenvalues are replaced by zeros, making the artifact-estimating GEVD-MWF the (trivial) all-zero filter.

The results are similar for EEG data containing only eye blinks, as shown in Fig. 3(b). However, there are two large differences compared to data containing muscle artifacts as well. First, the range of relatively stable performance is a lot larger (20-80%). Second, the trade-off between ARR and SER starts at a much lower percentage of retained eigenvalues. These differences can be explained from the fact that eye blink artifacts are low-dimensional and can be captured well in only a few modes. Muscle artifacts do not exhibit such a low-rank structure, and to estimate them, enough modes (and thus eigenvalues) are needed. The distribution of the eigenvalues in Fig. 4 indeed shows that the eigenvalue distribution for data with eye blinks is a lot narrower compared to muscle artifacts, which makes it possible to estimate eye blink artifacts with a GEVD-MWF of lower rank than is needed for muscle artifacts.

It is noted that the strategy to only retain the positive eigenvalues seems to lead to maximal ARR and an almost-optimal SER. This is good news, as this is an automatic strategy which does not require the manual selection of an eigenvalue cut-off, while still reaching a near-optimal performance.

In conclusion, the rank and the number of time lags that result in the best performance can be chosen using the same strategy for all subjects and artifact types. The rank can be set automatically by the algorithm by only retaining positive eigenvalues. The time lag can be set by the user, considering the trade-off between SER and computation time. The performance gain of the first few time lags is greatest, after which the gains become smaller and smaller for increasing number of lags. The computation time increases cubically with the number of lags because of the GEVD operation.

C. Effect of artifact SNR

In Fig. 5, the effect of artifact SNR on performance is shown for hybrid EEG data. The constant SER across SNRs indicates that the amount of EEG distortion in clean EEG segments does not depend on the artifact SNR, i.e. the EEG distortion caused by the MWF is low and does not depend on artifact amplitude. The ARR, which measures artifact suppression quality, increases for increasing SNR, which is expected from (21) as the numerator grows with the artifact amplitude. The difference between ARR_{hybrid} and ARR_{real} is constant for high SNRs (i.e. above 0 dB). This indicates that approximating \mathbf{d} by \mathbf{y} in 21 in order to compute the ARR for real data is valid for high-amplitude artifacts, as mentioned in Section III-D. For low SNRs (below 0 dB), the approximation is not valid: ARR_{hybrid} will keep decreasing as its numerator will tend towards zero (see 21), but the numerator of ARR_{real} will still contain the clean EEG power in \mathbf{y} , causing ARR_{real} to flatten out for decreasing SNRs.

The eye blink artifacts measured in the real data have a mean SNR around 12 dB. This means ARR_{real} is a valid measure to evaluate the performance of artifact removal for the real EEG data.

D. Comparison with other artifact removal techniques

The results for eye blink only data in Fig. 8(a) show the highest SER and ARR for the GEVD-based MWF. However, the basic MWF without the improvements described in Sections II-C and II-D achieves an SER worse than Infomax ICA and CCA, and an ARR of comparable magnitude. These findings are consistent with the observations in [21], where ICA and CCA were compared to the basic MWF for eye blink artifact removal in real and hybrid EEG data. This indicates that the MWF improvements investigated in this paper (i.e. including temporal information and incorporating the GEVD) strongly boost the performance of the MWF as an artifact removal algorithm, outperforming other current state-of-the-art algorithms.

Looking at complexity, the ICA methods have long computation times, while CCA and MWF are fast methods. The computation time of the basic MWF is comparable to CCA, but increases for more included time delays. Note that the performance of the GEVD-based MWF can be boosted even further at the cost of longer computation times (see Fig. 2 and Section V-B). For larger numbers of input EEG channels, for example in 128- or 256-channel EEG systems, the comparison of computation times for the different methods is still valid. Because computing a GEVD is an $\mathcal{O}(M^3)$ procedure [23] and CCA is efficiently computed based on a GEVD, the GEVD-MWF and CCA have similar complexities. ICA-based methods start with an initial pre-whitening step, which also has an $\mathcal{O}(M^3)$ complexity, even before the iterations of the algorithm start.

Fig. 8(b) shows similar results for the case where both eye blinks and muscle artifacts are removed by the algorithms. Again the GEVD-based MWF outperforms the basic MWF and the component-based methods. In most cases for ICA and CCA, the muscle artifact components were very ill-separated from the clean EEG components, often being present in almost half of the components. The rejection of multiple ill-separated components inevitably leads to larger EEG distortion, which explains the low SERs for ICA and CCA. On the other hand, the eye blink artifacts were often still easily identifiable as a single component. The poor performance of CCA is surprising, as it is a popular method for muscle artifact removal. It should be noted that the muscle artifacts in this study were deliberately induced by clenching the jaw in the EEG, causing high artifact amplitudes with possibly many myogenic generators, making it difficult to separate in only a few components. This is in contrast with the ‘‘spontaneous’’ muscle artifacts to which CCA was applied in [9], which have lower amplitudes and are present in fewer channels.

It is also noteworthy that component-based methods reduce the rank of the EEG data, as mixing matrix elements corresponding to the artifact components are set to zero before back-projecting the components to the EEG sensor space. The

MWF subtracts a reduced-rank artifact estimate from the EEG, which generally does not reduce the rank of the processed EEG.

VI. CONCLUSIONS

In this paper, we have successfully implemented and validated the MWF as a method for semi-automatic artifact removal from EEG. The method is integrated with a graphical user interface, which allows the user to mark unwanted artifacts directly in the EEG signals. We demonstrated the algorithm performance using both hybrid and real EEG recordings from 10 subjects. We found that for best performance, the rank of the artifact covariance matrix approximation can be chosen automatically by the algorithm so that only positive generalized eigenvalues contribute to the artifact estimation. The second most important parameter, the number of time lags included in the filter, can be chosen by the user and involves a trade-off between performance and computation time.

The method was compared with several algorithms which are popular for EEG artifact removal, such as ICA and CCA. We found that the MWF outperformed the other methods regarding artifact removal performance. Using a low number of time lags in the MWF, the computation time is very short and similar to CCA. Both the MWF and CCA are at least an order of magnitude faster than ICA.

The MATLAB toolbox implementing the MWF and the GUI is made available online (see footnote in Section I) in the hopes that it will provide a useful and user-friendly tool for other researchers dealing with artifacts in EEG studies. Future work includes the development of temporal artifact detection methods based on signal statistics which enable the MWF to be used as a fully automatic artifact removal algorithm.

VII. ACKNOWLEDGEMENTS

The authors declare no conflict of interest. This work was carried out at the ExpORL research group of KU Leuven, and at ESAT Laboratory of KU Leuven, in the frame of KU Leuven Special Research Fund BOF/STG-14-005, C14/16/057, CoE PFV/10/002 (OPTEC) and OT/14/119, and Research Projects FWO G0D7516N, G0A4918N, and 1512316N. Additionally, this project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 637424, ERC starting grant to Tom Francart). The first author is supported by a PhD grant for Strategic Basic research by the Research Foundation Flanders (FWO) under project number 1S46117N. The scientific responsibility is assumed by its authors.

REFERENCES

- [1] M. Teplan, "Fundamentals of EEG measurement," *Measurement science review*, vol. 2, no. 2, pp. 1–11, 2002.
- [2] T. C. Ferree, P. Luu, G. S. Russell, and D. M. Tucker, "Scalp electrode impedance, infection risk, and EEG data quality," *Clinical Neurophysiology*, vol. 112, no. 3, pp. 536–544, 2001.
- [3] J. C. Huhta and J. G. Webster, "60-Hz interference in electrocardiography," *IEEE Transactions on Biomedical Engineering*, no. 2, pp. 91–101, 1973.
- [4] A. Metting van Rijn, A. Peper, and C. Grimbergen, "High-quality recording of bioelectric events," *Medical and Biological Engineering and Computing*, vol. 28, no. 5, pp. 389–397, 1990.
- [5] R. J. Barry, A. R. Clarke, S. J. Johnstone, C. A. Magee, and J. A. Rushby, "EEG differences between eyes-closed and eyes-open resting conditions," *Clinical Neurophysiology*, vol. 118, no. 12, pp. 2765–2773, 2007.
- [6] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal – state-of-the-art and guidelines," *Journal of neural engineering*, vol. 12, no. 3, p. 031001, 2015.
- [7] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals – practices and possibilities," *IEEE transactions on information technology in biomedicine*, vol. 16, no. 3, pp. 488–500, 2012.
- [8] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [9] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2583–2587, 2006.
- [10] M. De Vos, S. Riès, K. Vanderperren, B. Vanrumste, F.-X. Alario, V. S. Huffel, and B. Burle, "Removal of muscle artifacts from EEG recordings of spoken language production," *Neuroinformatics*, vol. 8, no. 2, pp. 135–150, 2010.
- [11] X. Chen, A. Liu, Q. Chen, Y. Liu, L. Zou, and M. J. McKeown, "Simultaneous ocular and muscle artifact removal from EEG data by exploiting diverse statistics," *Computers in Biology and Medicine*, vol. 88, pp. 1–10, 2017.
- [12] A. Delorme, S. Makeig, and T. Sejnowski, "Automatic artifact rejection for EEG data using high-order statistics and independent component analysis," in *Proceedings of the third international ICA conference*, 2001, pp. 9–12.
- [13] H. Nolan, R. Whelan, and R. Reilly, "FASTER: fully automated statistical thresholding for EEG artifact rejection," *Journal of neuroscience methods*, vol. 192, no. 1, pp. 152–162, 2010.
- [14] A. Mognon, J. Jovicich, L. Bruzzone, and M. Buiatti, "ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features," *Psychophysiology*, vol. 48, no. 2, pp. 229–240, 2011.
- [15] S. Doclo and M. Moonen, "GSVD-based optimal filtering for single and multimicrophone speech enhancement," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2230–2244, 2002.
- [16] B. Van Dun, J. Wouters, and M. Moonen, "Multi-channel Wiener filtering based auditory steady-state response detection," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 2. IEEE, 2007, pp. II-929.
- [17] R. Serizel, M. Moonen, B. Van Dijk, and J. Wouters, "Low-rank approximation based multichannel Wiener filter algorithms for noise reduction with application in cochlear implants," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 4, pp. 785–799, 2014.
- [18] H. Maki, T. Toda, S. Sakti, G. Neubig, and S. Nakamura, "EEG signal enhancement using multi-channel Wiener filter with a spatial correlation prior," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2639–2643.
- [19] A. Bertrand, "Distributed signal processing for wireless EEG sensor networks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 6, pp. 923–935, 2015.
- [20] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Prentice Hall, 1996.
- [21] B. Somers and A. Bertrand, "Removal of eye blink artifacts in wireless EEG sensor networks using reduced-bandwidth canonical correlation analysis," *Journal of Neural Engineering*, vol. 13, no. 6, p. 066008, 2016.
- [22] A. Hassani, A. Bertrand, and M. Moonen, "GEVD-based low-rank approximation for distributed adaptive node-specific signal estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2557–2572, 2016.
- [23] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore: The Johns Hopkins University Press, 1996.
- [24] J. Särelä and H. Valpola, "Denoising source separation," *Journal of Machine Learning Research*, vol. 6, no. Mar, pp. 233–272, 2005.
- [25] A. de Cheveigné and L. C. Parra, "Joint decorrelation, a versatile tool for multichannel data analysis," *Neuroimage*, vol. 98, pp. 487–505, 2014.
- [26] S. H. Sardouie, M. B. Shamsollahi, L. Albera, and I. Merlet, "Denoising of ictal EEG data using semi-blind source separation methods based

- on time-frequency priors,” *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 839–847, 2015.
- [27] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, “xDAWN algorithm to enhance evoked potentials: application to brain–computer interface,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 8, pp. 2035–2043, 2009.
- [28] B. Rivet and A. Souloumiac, “Optimal linear spatial filters for event-related potentials based on a spatio-temporal model: Asymptotical performance analysis,” *Signal Processing*, vol. 93, no. 2, pp. 387–398, 2013.
- [29] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, “Optimal spatial filtering of single trial EEG during imagined hand movement,” *IEEE transactions on rehabilitation engineering*, vol. 8, no. 4, pp. 441–446, 2000.
- [30] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, “Independent component analysis of electroencephalographic data,” in *Advances in neural information processing systems* 8, 1996, pp. 145–151.
- [31] A. Hyvärinen, “Fast and robust fixed-point algorithms for independent component analysis,” *IEEE transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.