

RESEARCH ARTICLE

Suppression of overlearning in independent component analysis used for removal of muscular artifacts from electroencephalographic records

Jan Sebek^{1*}, Radoslav Bortel¹, Pavel Sovka

Dept. of Circuit Theory, Czech Technical University, Faculty of Electrical Engineering, Prague, Czech Republic

¹ These authors contributed equally to this work.

* sebekja4@fel.cvut.cz



OPEN ACCESS

Citation: Sebek J, Bortel R, Sovka P (2018) Suppression of overlearning in independent component analysis used for removal of muscular artifacts from electroencephalographic records. PLoS ONE 13(8): e0201900. <https://doi.org/10.1371/journal.pone.0201900>

Editor: Lawrence M Ward, University of British Columbia, CANADA

Received: February 12, 2018

Accepted: July 24, 2018

Published: August 14, 2018

Copyright: © 2018 Sebek et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available from the Dryad database and can be accessed at the following URL: <https://doi.org/10.5061/dryad.r4b69rg>.

Funding: This work was supported by Czech Ministry of Education, Youth and Sports OP VVV funded project "CZ.02.1.01/0.0/0.0/16_019/0000765 Research Center for Informatics" and by the Internal CTU grant SGS17/183/OHK3/3T/13 "Special Applications of Signal Processing".

Abstract

This paper addresses the overlearning problem in the independent component analysis (ICA) used for the removal of muscular artifacts from electroencephalographic (EEG) records. We note that for short EEG records with high number of channels the ICA fails to separate artifact-free EEG and muscular artifacts, which has been previously attributed to the phenomenon called overlearning. We address this problem by projecting an EEG record into several subspaces with a lower dimension, and perform the ICA on each subspace separately. Due to a reduced dimension of the subspaces, the overlearning is suppressed, and muscular artifacts are better separated. Once the muscular artifacts are removed, the signals in the individual subspaces are combined to provide an artifact free EEG record. We show that for short signals and high number of EEG channels our approach outperforms the currently available ICA based algorithms for muscular artifact removal. The proposed technique can efficiently suppress ICA overlearning for short signal segments of high density EEG signals.

Introduction

During the measurement of head surface potentials with electroencephalography (EEG), weak brain signals are often corrupted by various sources of interference. One of the most common interference sources are potentials generated by muscles contracted by a measured subject. These are so called muscular or electromyographic (EMG) artifacts. The EMG artifacts are commonly occurring nuisance in many EEG records; therefore, good methods for their removal are valuable tools for EEG processing.

In the past it was suggested to use the independent component analysis (ICA) for the EMG artifact removal. Under suitable circumstances the ICA was demonstrated as a good approach for this task, and its use was developed in many works [1–11].

Competing interests: The authors have declared that no competing interests exist.

There are however limitations to what the ICA can achieve. If the signal length is not sufficient with respect to the number of EEG channels, the ICA suffers from so called overlearning [12–14], which prevents the separation and subsequent removal of EMG artifacts.

In the remainder of this section we introduce the problem of artifact removal with ICA more formally, specify the problem of overlearning, and review the current state-of-the-art methods in which this problem is addressed.

Let $\mathbf{x}[n] = [x_1[n], \dots, x_M[n]]^T$, $n = 1, \dots, N$ denote EEG signals measured using M EEG electrodes. We assume that these signals are created by a linear mixture of source signals $\mathbf{s}[n] = [s_1[n], \dots, s_{M'}[n]]^T$

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n], \tag{1}$$

where \mathbf{A} denotes an $M \times M'$ mixing matrix. Note that we do not expect the number of sources M' to be less or equal to the number of signals M . This is a common assumption in the derivation of the ICA, but cannot be realistically expected in the addressed problem. We further assume that the source signals $s_m[n]$ are either generated by a brain or by other interfering sources (e.g. contracted muscles).

Our task is to remove the artifacts from the brain signals. As mentioned above, for this purpose it is possible to use the ICA, which allows to estimate a separation matrix \mathbf{B} that transforms signals $\mathbf{x}[n]$ into source components $\hat{\mathbf{s}}[n] = [\hat{s}_1[n], \dots, \hat{s}_M[n]]^T$ that have their mutual dependence minimized

$$\hat{\mathbf{s}}[n] = \mathbf{B}\mathbf{x}[n]. \tag{2}$$

If the measured signals $\mathbf{x}[n]$ are sufficiently long, this approach was reported to achieve a fairly good separation of brain signals and artifacts [2, 6]—this means that the most of the components $\hat{s}_m[n]$ contain either brain signals or artifacts, while the mutual intermixing of these signals is noticeably reduced. Thus, it is possible to either manually [10, 11], or automatically [11] classify each $\hat{s}_m[n]$ as either a brain signal or an artifact, and retain only the brain signals, which we will denote as $\tilde{\mathbf{s}}[n] = [\tilde{s}_1[n], \dots, \tilde{s}_M[n]]^T$

$$\tilde{s}_m[n] = \begin{cases} \hat{s}_m[n], & \text{if } \hat{s}_m[n] \text{ is a brain signal,} \\ 0, & \text{if } \hat{s}_m[n] \text{ is an artifact.} \end{cases} \tag{3}$$

With brain signal components $\tilde{\mathbf{s}}[n]$ we can try to reconstruct the measured signals without the artifacts

$$\tilde{\mathbf{x}}[n] = \mathbf{B}^{-1}\tilde{\mathbf{s}}[n], \tag{4}$$

where $\tilde{\mathbf{x}}[n] = [\tilde{x}_1[n], \dots, \tilde{x}_M[n]]^T$ are the estimates of the measured signals with artifacts suppressed.

The above-mentioned methodology was used before [6, 11, 15, 16] with a certain success; however, it will work only if the measured signals are sufficiently long compared to the number of measured channels. This limitation is illustrated on the following two examples.

Example 1: We have used a 2s long record from 9 EEG electrodes located in the frontal area. This record contains a clearly visible muscular artifact present in each of the 9 channels. We applied the above-mentioned method to the data (the ICA was implemented using the popular FastICA algorithm [17], Eq (14, 21, 24)). Fig 1A shows the original signals $\mathbf{x}[n]$, Fig 1B shows the separated components $\hat{\mathbf{s}}[n]$, and Fig 1C shows the reconstructed signals $\tilde{\mathbf{x}}[n]$. In this example the muscular artifacts were suppressed, and we have obtained a nice looking reconstruction of the original EEG.

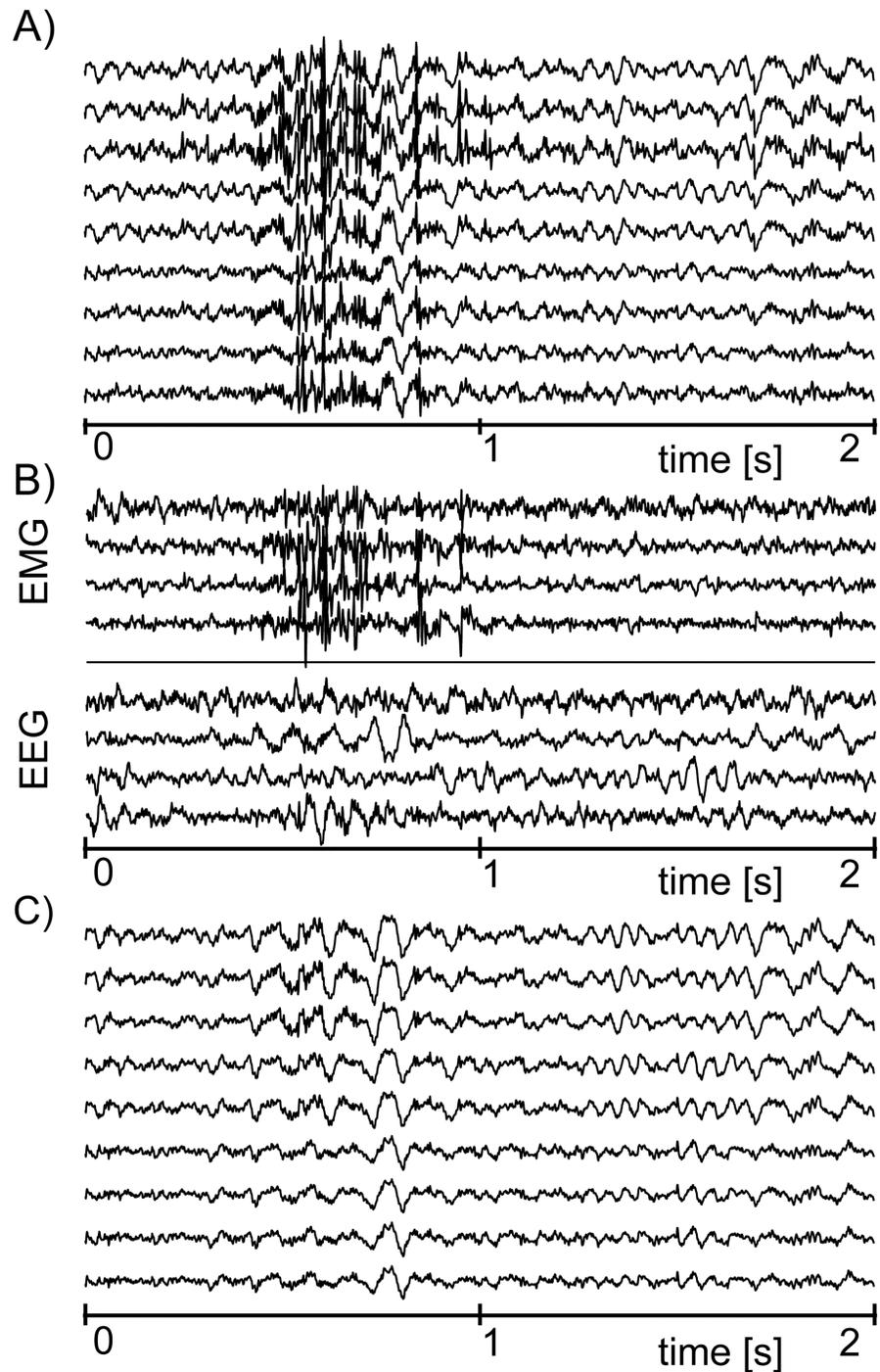


Fig 1. An example of muscular artifact suppression using the ICA. (A) The original EEG signals $x[n]$ corrupted by EMG artifacts. (B) The separated source components $\hat{s}[n]$, which were classified as brain signals (denoted as EEG) and muscular artifacts (denoted as EMG). (C) The reconstructed EEG signals $\hat{x}[n]$ with suppressed muscular artifacts.

<https://doi.org/10.1371/journal.pone.0201900.g001>

Example 2: We have used another 2s long record, but this time we used a 111 electrode array covering the frontal, parietal, temporal and occipital regions of subject's head. As in the previous example, the record contains a clearly visible muscular artifact. We again applied the above-mentioned methodology; however, the outcome was less successful.

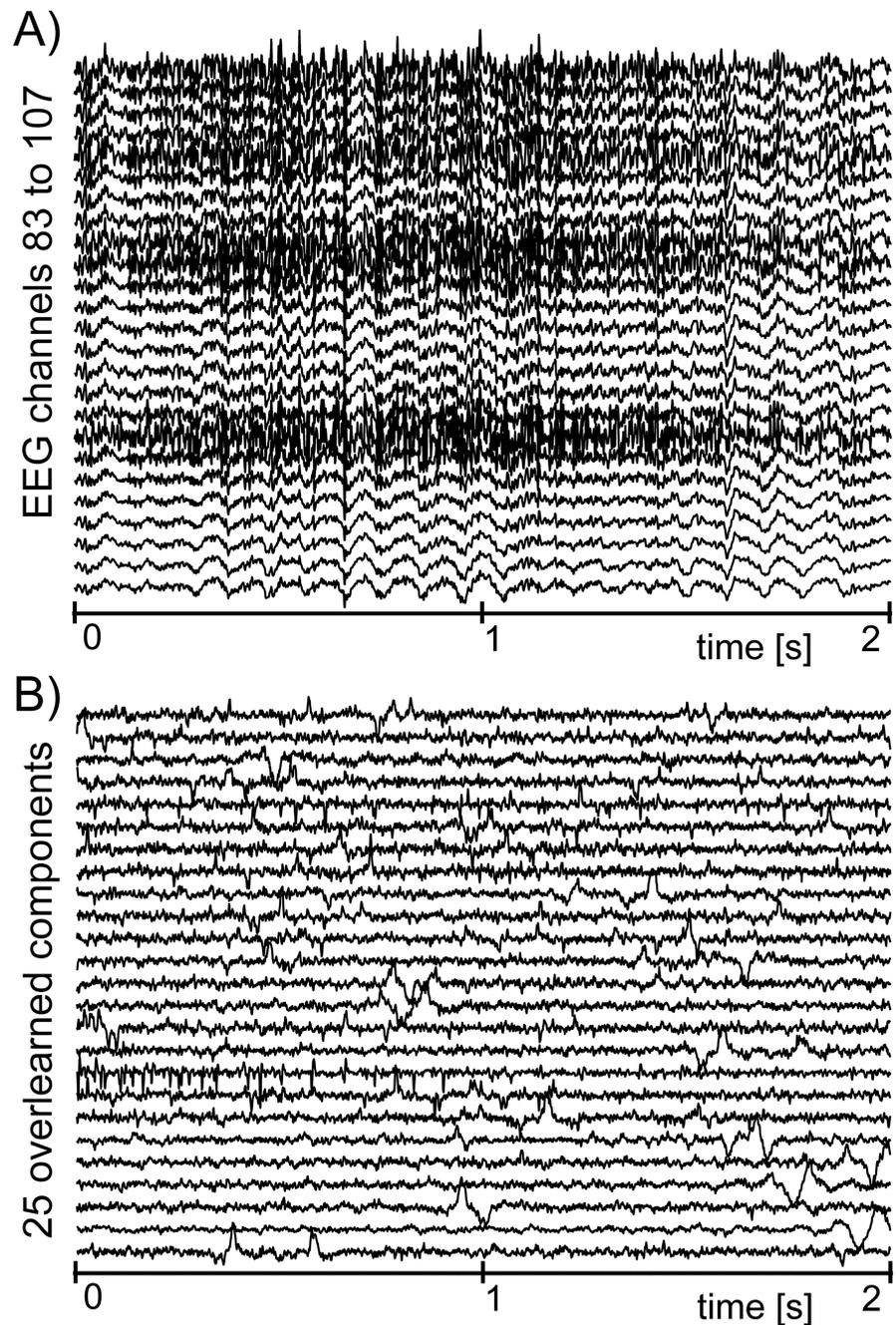


Fig 2. An example of a failed attempt to separate muscular artifacts from 111 channel EEG record. (A) EEG signals $x[n]$ corrupted by EMG artifacts (for spatial reasons only 25 out of 111 EEG channels are shown). (B) The source components $\hat{s}[n]$ provided by ICA (again for spatial reasons only 25 out of 111 components are shown; the remaining components look very similar to those that are presented—they are composed of occasional bumps, and no clear separation of EEG and muscular artifacts can be recognized).

<https://doi.org/10.1371/journal.pone.0201900.g002>

The original signals $x[n]$ (shown in Fig 2A) were separated into source signals $\hat{s}[n]$ (Fig 2B); however, this time we obtained no clear separation between EEG and EMG. In fact, the resulting waveforms resemble neither EEG nor EMG, and are merely composed of occasionally repeated ‘bumps’.

What we illustrated in Example 2 is a phenomenon known as overlearning [12–14]. The overlearning occurs when the ICA algorithm has too many degrees of freedom, which allows it to find components $\hat{s}[n]$ that appear to be even more independent than if $\hat{s}[n]$ contained separated EEG and EMG signals.

There are several ways how to tackle the problem of overlearning.

First, we could use longer EEG records. This, however, is not desirable. While the above-mentioned algorithm might sometimes be able to suppress muscular artifacts, it will inevitably distort the reconstructed EEG. For example, by removing 4 of the 9 components $\hat{s}_m[n]$ in the Example 1 we have reduced the rank of the reconstructed data $\tilde{x}[n]$ from the original 9 to 5 (note that the rank of an M dimensional signal $x[n]$ with length N we understand the rank of the matrix $[x_m[n]]_{m=1 \dots M, n=1 \dots N}$). Thus even though the result may be ‘nice looking’ some loss of information in the reconstructed signals must have occurred. Therefore, we would like to limit the application of muscular artifact removal and the corresponding rank reduction only to the time intervals, where the muscular artifacts have really occurred, avoiding unnecessary distortion of the artifact free EEG. Better yet, the interval length should not exceed the time during which the EMG artifact origin position (and the corresponding mixing matrix \mathbf{A}) is close to stationary. These time intervals can be fairly short, often less than a few seconds [18].

Next, in the past there were several published works that suggested how to suppress overlearning.

Works [12–14, 19] suggested that the proper choice of contrast functions in the ICA algorithm can decrease the level of overlearning. However, the improvement provided is often not sufficient to completely solve the problem. In our Example 2 we have used the FastICA algorithm with robust approximation of negentropy as the contrast function (which was suggested as the best choice in [12]); however, the overlearning is still distinctly present.

Works [13, 14] suggested to use the PCA to decrease the dimensionality of $x[n]$ prior to the ICA computation. When the dimension reduction is sufficient, this approach will undoubtedly suppress overlearning; however, this will be at the cost of EEG distortion. The brain signals may have smaller power than the EMG artifacts, thus the dimensionality reduction will inevitably remove some energy from EEG, possibly removing useful information that may be missing in a subsequent EEG analysis. Thus, the application of the PCA for the dimension reduction is not always desirable, especially, if there is a way to avoid it.

Works [13, 14] also noted that the ‘bumps’ observable in the separated components produced by an overlearned ICA may have their energy predominantly at low frequencies, and therefore suggested to high-pass filter the analyzed signals prior to the ICA application. It was suggested to use either a 1Hz fixed cut-off frequency high-pass filter or an AR-process based high pass filter. It was further claimed that this high-pass filtering lessens the overlearning. In our experience this approach is not sufficient. In Example 2, the signals were already high pass filtered with a 1Hz cut-off filter (to suppress the baseline wonder), but the overlearning has still occurred. The ‘bumps’ in the separated components are simply faster (with duration of about 50ms); therefore, their occurrence cannot be prevented by prior high pass filtering without damaging useful brain signals.

Work [3], which concentrated on the suppression of ocular artifacts, suggested to use a PCA based dimension reduction to extract 3 principal components with the highest power, separate these components using the ICA, and then subtract the separated components from the original signals using a linear regression. This method is peculiar in that it uses no artifact classification. It assumes that the ocular artifacts will have the highest power, and thus will dominate the strongest principal components. Due to a distinct dimension reduction by the PCA, this methodology was claimed to be resilient to overlearning; however, we do not find it useful for other

than ocular artifacts—not all EMG (or other) artifacts have power higher than brain signals, and consequently they may not get included into a few strongest principal components. In fact, with weaker artifacts, the strongest principal components may contain mostly brain signals, the removal of which would lead to the corruption of artifact free EEG signals. Still, for the sake of completeness, we include this method in our evaluation of the current state-of-the-art.

Last, there are some methods that try to alter the basic approach (2)–(4). The wavelet enhanced ICA [7, 8] extends the brain signal selection by adding wavelet filtering. Surface Laplacian [10] was also proposed to filter the EEG channels after the basic approach (2)–(4) was applied. These techniques, however, have no effect on the overlearning, which already occurs in the separation step (2). Another technique, the ensemble empirical mode decomposition and subsequent ICA [20], first decomposes each of the measured signals $\mathbf{x}[n]$ into the intrinsic mode functions, and then the ICA is applied. This does affect overlearning, but in a negative way, because the number of signals separated by ICA increases, which makes the effects of overlearning more severe.

All in all, in our opinion, none of the above-mentioned methods provides an efficient way to suppress the ICA overlearning without problematic side effects. We have therefore developed our own approach to the ICA overlearning suppression, which seems to outperform the above-mentioned methods, and provides better EEG reconstruction of short EEG records with EMG artifacts.

Methods

In this section we will first describe the suggested algorithm, then point out some of its properties, and suggest how to choose its parameters.

Suggested algorithm

We propose to separate the M -dimensional signal $\mathbf{x}[n]$ into K subspaces with dimension L

$$\mathcal{X}_k[n] = \mathbf{P}_k \mathbf{x}[n], \quad k = 1, \dots, K, \tag{5}$$

where \mathbf{P}_k are $L \times M$ subspace projection matrices (we assume that their rank is L , and their choice will be discussed in subsection Choice of subspaces), and $\mathcal{X}_k[n]$ are L dimensional signals.

Next, we apply the ICA separately to each signal $\mathcal{X}_k[n]$

$$\hat{\mathcal{S}}_k[n] = \mathbf{B}_k \mathcal{X}_k[n], \tag{6}$$

where \mathbf{B}_k is a separation matrix estimated for $\mathcal{X}_k[n]$ (note that this means that the ICA will be computed K times for each $\mathcal{X}_k[n]$ separately).

After, we classify $\hat{\mathcal{S}}_k[n] = [\hat{s}_{k,1}[n], \dots, \hat{s}_{k,L}[n]]^T$ as either brain signals or artifacts, and retain only brain signals $\tilde{\mathcal{S}}_k[n] = [\tilde{s}_{k,1}[n], \dots, \tilde{s}_{k,L}[n]]^T$

$$\tilde{s}_{k,i}[n] = \begin{cases} \hat{s}_{k,i}[n] & \text{if } \hat{s}_{k,i}[n] \text{ is a brain signal,} \\ 0 & \text{if } \hat{s}_{k,i}[n] \text{ is an artifact.} \end{cases} \tag{7}$$

This operation can be expressed in a matrix form as

$$\tilde{\mathcal{S}}_k[n] = \mathbf{Q}_k \hat{\mathcal{S}}_k[n], \tag{8}$$

where \mathbf{Q}_k is a diagonal matrix with its diagonal comprised of zeros and ones, positioned so that the signals $\hat{s}_{k,i}[n]$ classified as artifacts are eliminated according to (7).

Now, we reconstruct signals $\tilde{\mathcal{X}}_k[n]$, in which the muscular artifacts are suppressed

$$\tilde{\mathcal{X}}_k[n] = \mathbf{B}_k^{-1} \tilde{\mathcal{S}}_k[n]. \tag{9}$$

The above-mentioned sequence of steps can be expressed as

$$\tilde{\mathcal{X}}_k[n] = \mathbf{B}_k^{-1} \mathbf{Q}_k \mathbf{B}_k \mathbf{P}_k \mathbf{x}[n] = \mathbf{C}_k \mathbf{P}_k \mathbf{x}[n], \tag{10}$$

where we denoted $\mathbf{C}_k = \mathbf{B}_k^{-1} \mathbf{Q}_k \mathbf{B}_k$.

Last, we combine the individual signals $\tilde{\mathcal{X}}_k[n]$, reconstructing the artifact free EEG. For this purpose we denote

$$\mathcal{P} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \vdots \\ \mathbf{P}_K \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} \mathbf{C}_1 & 0 & \dots & 0 \\ 0 & \mathbf{C}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{C}_K \end{bmatrix}, \tag{11}$$

$$\tilde{\mathcal{X}}[n] = \begin{bmatrix} \tilde{\mathcal{X}}_1[n] \\ \tilde{\mathcal{X}}_2[n] \\ \vdots \\ \tilde{\mathcal{X}}_K[n] \end{bmatrix}. \tag{12}$$

Using notations (11) and (12), all the above-mentioned steps can be expressed as

$$\tilde{\mathcal{X}}[n] = \mathcal{C} \mathcal{P} \mathbf{x}[n]. \tag{13}$$

To combine the elements of $\tilde{\mathcal{X}}[n]$ into the reconstructed signals $\tilde{\mathbf{x}}[n]$, we reverse the projection into subspaces (matrix \mathcal{P}) by employing the Moore-Penrose pseudoinverse of \mathcal{P}

$$\tilde{\mathbf{x}}[n] = \mathcal{P}^\dagger \tilde{\mathcal{X}}[n], \text{ where } \mathcal{P}^\dagger = (\mathcal{P}^T \mathcal{P})^{-1} \mathcal{P}^T. \tag{14}$$

Thus, the entire transformation that suppresses the artifacts can be expressed as

$$\tilde{\mathbf{x}}[n] = \mathcal{P}^\dagger \mathcal{C} \mathcal{P} \mathbf{x}[n] = \mathcal{D} \mathbf{x}[n], \tag{15}$$

where we denoted $\mathcal{D} = \mathcal{P}^\dagger \mathcal{C} \mathcal{P}$.

Notes about properties of suggested algorithm

Since the dimension L of the subspaces is smaller than the dimension M of the original signals $\mathbf{x}[n]$, the ICA that is used to find the separation matrices \mathbf{B}_k may be less prone to overlearning. In essence by choosing L sufficiently small, the problems that we illustrated in our Example 2 disappear, and the ICA behaves as in low dimensional case shown in Example 1.

If the subspace projection matrices \mathbf{P}_k are chosen so that \mathcal{P} has full rank, and if no artifacts are found in the separated signals $\hat{\mathcal{S}}_k[n]$, then $\tilde{\mathcal{S}}_k[n] = \hat{\mathcal{S}}_k[n]$ (i.e. \mathbf{Q}_k are identity matrices), \mathbf{C}_k and \mathcal{C} become identity matrices, and \mathcal{D} becomes the identity matrix. Consequently, no distortion is introduced into the processed signals. This is a great advantage over the PCA based overlearning suppression approach, where the rank of measured signals is always reduced, irrespective to the number of components that are being removed (i.e. the rank is reduced

even if no artifacts are being removed), which causes unavoidable distortion of the resulting reconstructed signals.

If artifacts are detected in the source signal $\hat{\mathbf{S}}_k[n]$ the rank of $\tilde{\mathbf{S}}_k[n]$ and \mathbf{C} will be reduced. The rank of \mathbf{D} may then also be reduced; however, in the Results section we will show that the newly proposed method causes dimensionality reduction much less severe than the one caused by the PCA based overlearning suppression.

Choice of subspaces

Three points should be observed when choosing the subspaces defined by the projection matrix \mathcal{P} .

1. The dimension L should be small enough so that the overlearning does not pose a problem. It should however remain big enough so that there is still a sufficient number of signals so that the ICA can separate brain signals and artifacts.
2. The matrix \mathcal{P} should be full rank (otherwise \mathbf{D} would not be identity when no artifacts are being removed, and unnecessary EEG distortion would occur).
3. The orientation of subspaces should be chosen so that the individual ICAs can achieve good separation between brain signals and artifacts in separated signals $\hat{\mathbf{S}}_k[n]$ even with the reduced dimension L of $\mathcal{X}_k[n]$.

To address point (i), in this paper we identify the optimal range of values L by applying the algorithm to EEG data, and evaluating for which L we obtain the best reconstruction of brain signals. In the following sections we will introduce a methodology for the evaluation of the quality of brain signal reconstruction, and we will show that even for various EEG electrode systems we can identify a fixed range of values of L where the algorithm always performs well. The reader can then use the value of L suggested by our evaluation, and the optimization of L does not need to be repeated.

Point (ii) can be achieved by construction of \mathcal{P} —simply by choosing sufficient number of sufficiently diverse subspaces, \mathcal{P} can be full rank.

Point (iii) is somewhat more difficult to tackle. In this paper we would like to take a simple approach that is based on our experience with processing of EEG data, and suggest a choice that can be shown to produce good quality of reconstruction of brain signals. Specifically, we have observed that when we use a high density EEG array (e.g. with 111 electrodes), but limit ourselves to a small group of adjacent electrodes (e.g. 10-15 adjacent electrodes), the ICA usually provides a nice separation of brain signals and muscular artifacts. We have actually already illustrated this situation in our Example 1 which shows signals from a subgroup of 9 adjacent electrodes chosen from the 111 electrode array used in Example 2. In this paper we would therefore suggest the following steps to choose the matrices \mathbf{P}_k .

First, let us denote d_{ij} the Euclidean distance between i -th and j -th electrode. Next, we define vectors

$$\boldsymbol{\ell}_k = [\ell_{k,1}, \dots, \ell_{k,L}], \tag{16}$$

where $\ell_{k,i}$, $i = 1, \dots, L$ are indexes of L electrodes with the smallest distance $d_{\ell_{k,i},k}$ from the k -th electrode (with k -th electrode included). Now, we define the vectors

$$\mathbf{P}_{k,i} = [p_{k,i,1}, \dots, p_{k,i,M}], \quad p_{k,i,j} = \begin{cases} 1, & \text{if } \ell_{k,i} = j, \\ 0, & \text{otherwise,} \end{cases} \tag{17}$$

and construct the projection matrices as

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{p}_{k,1} \\ \vdots \\ \mathbf{p}_{k,L} \end{bmatrix}, \quad k = 1, \dots, M. \tag{18}$$

Thus, each projection matrix \mathbf{P}_k will extract a subspace composed of L signals from a group of L electrodes closest to the k -th electrode (k -th electrode included).

Before continuing we will make a few notes about the projection matrices \mathbf{P}_k given by (18).

Note that each $\mathbf{p}_{k,i}$ (i.e. each row of \mathbf{P}_k) is composed of zeros and a single value of 1. Thus, the columns of \mathbf{P}_k are orthogonal, and consequently the columns of \mathcal{P} are orthogonal as well. Further, because $k = 1, \dots, M$, each column of \mathcal{P} contains at least a single value of 1, and so has a nonzero norm. Consequently, \mathcal{P} is a full rank matrix, just as we have required in point (ii).

We do not claim that our choice of \mathbf{P}_k is optimal, but we will illustrate and claim that it provides better results than all the above-mentioned state-of-the-art methods available today. Further optimization of \mathbf{P}_k may be possible, but it will not be addressed in this paper.

Last, some additional rationale for this choice of subspaces can be provided. The energy of surface EEG is typically dominated by shallow brain sources and artifacts. The head surface regions affected by these sources are typically somewhat spread due to the volume conduction; however, the greatest surface potential changes are still somewhat localized to the proximity of the origin of the source [21]. By choosing each subspace as the signals from a group of adjacent electrodes, we try to minimize the number of sources that dominate the signals in a subspace (at least as opposed to a subspace that would be composed of signals from nonadjacent electrodes spread all around the subject’s head). Consequently, it becomes easier for the ICA to separate brain signals and artifacts as was requested in point (iii).

ICA algorithm

To find the separation matrices \mathbf{B}_k , we suggest to use the FastICA [17, Eq (14, 21, 24)] on data that were pre-whitened by PCA [22]. This is a popular approach, with fast convergence and previous successes in artifact separation [2, 3]. Note that the suggested methodology may work with some other ICA algorithm, but testing the performance with other ICA algorithms is out of the scope of this paper. It is also noteworthy that works [12–14] reported that the overlearning does not seem to be dependent on the choice of a specific ICA algorithm.

Artifact classification

To decide whether the separated components $\hat{\mathcal{S}}_k[n]$ are of a brain or a muscular origin, we used a simple classifier that compares the power of a signal at lower and higher frequencies. Compared to clean EEG, the EMG artifacts are known to have higher energy at higher frequency bands [11, 18, 23], which allows their easy identification.

We first filtered each signal $\hat{s}_{k,i}[n]$ into three frequency bands 3-30Hz, 60-90Hz and 110-140Hz. Then, we estimated the average power of the signals in each of these bands as $P_{k,i,1}$, $P_{k,i,2}$ and $P_{k,i,3}$, respectively. Next, we computed the minimum of power ratios

$$\alpha_{k,i} = \min(P_{k,i,1}/P_{k,i,2}, P_{k,i,1}/P_{k,i,3}). \tag{19}$$

Last, the values of $\alpha_{k,i}$ were compared with a chosen threshold T to determine whether $\hat{s}_{k,i}[n]$ is of a brain or a muscular origin.

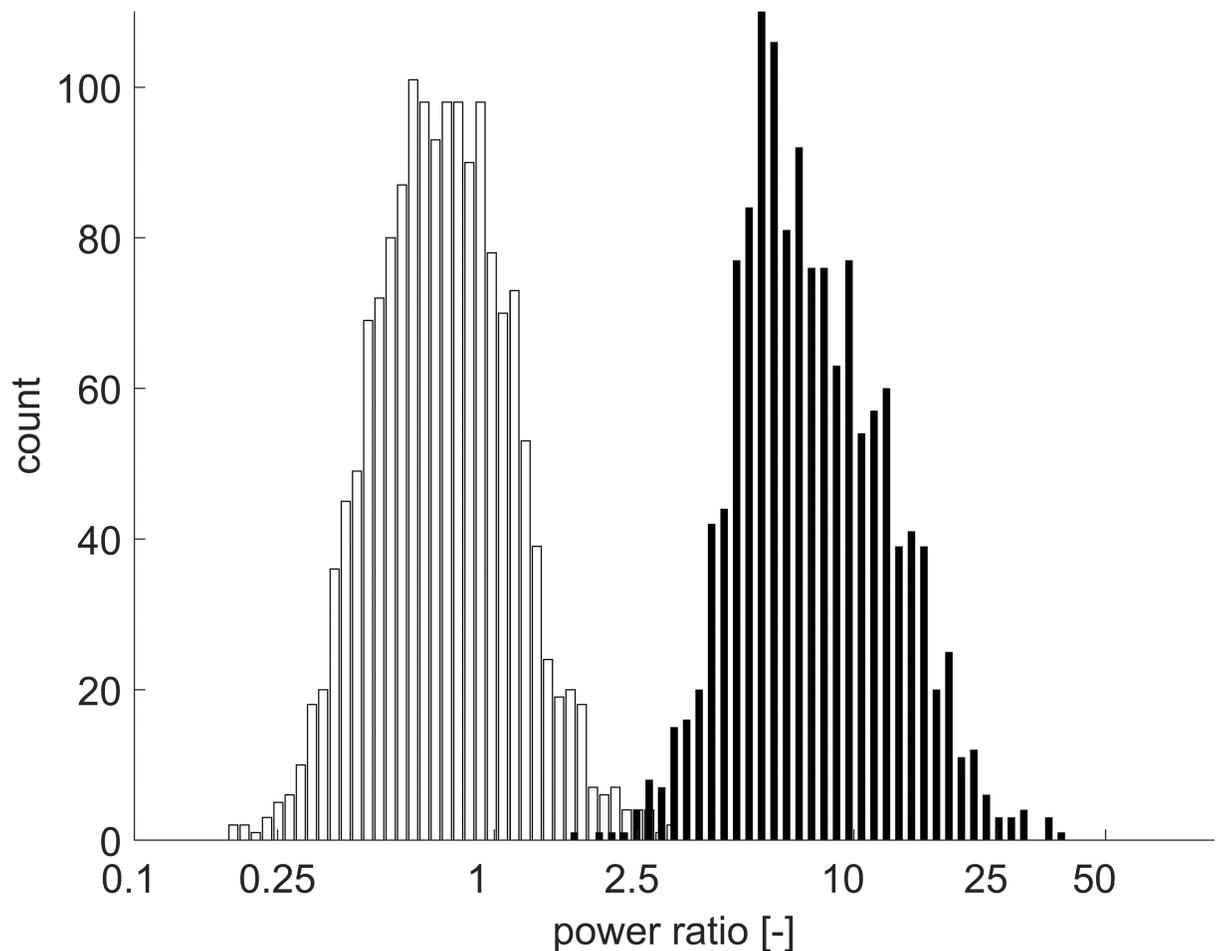


Fig 3. A histogram of power ratios α for source components classified as EMG (white) and source components classified as brain signals (black).

<https://doi.org/10.1371/journal.pone.0201900.g003>

To determine the optimal value of the threshold T , we computed $\alpha_{k,i}$ for 2992 separated signals $\hat{s}_{k,i}[n]$ that were manually classified as either brain signals or muscular artifacts. The computed $\alpha_{k,i}$ were then used to plot histograms shown in Fig 3, where a clear separation can be seen between the values of $\alpha_{k,i}$ for the individual classes. From Fig 3 we can read that the optimal value of T is about 2.5. Thus, a signal $\hat{s}_{k,i}[n]$ was classified as a muscular artifact if $\alpha_{k,i} < 2.5$.

Evaluation of algorithm performance

To evaluate the properties of the suggested method, we may apply it to EEG records containing EMG artifacts. This approach has however one serious drawback. While it is possible to observe that the artifacts are suppressed, it is not possible to judge how well the brain signals are reconstructed—with a common EEG record we have no way of knowing what exactly should the EEG look like, when the artifacts are removed. Therefore, this method cannot be objectively evaluated on real world EEG signals corrupted by EMG artifacts.

To circumvent this limitation while keeping the evaluation as realistic as possible, we have devised a testing procedure that uses real world EEG and EMG records, but simulates their

mixing using a realistic head model. Specifically, we used the following approach. We started with artifact free EEG signals, henceforth denoted as $\mathbf{x}_{clean}[n]$. These signals were recorded on 20 healthy subjects (10 male, 10 female) with age ranging from 19 to 37 years (25.4 ± 5.1 years). During the acquisition of EEG data, the participants were seated with their head supported by a headrest, and instructed to relax. From each subject we obtained a 10-minutes long record with their eyes open and 2-minute long records with their eyes closed. These measurements were approved by a local ethical committee and an informed consent in the written form was obtained from all participants. The data from 111 electrode system with a reference placed on a forehead were digitized with 16 bit resolution and sampled at the frequency of 1024Hz. The data were filtered by a notch filter removing any possible power noise interference at 50Hz and higher harmonics, and by a high pass filter with 1Hz cut-off frequency to suppress the baseline wander. The channels showing poor electrode connection were visually identified and rejected from further processing. After recording, the EEG was visually checked for artifacts. Special attention was paid to the fringe electrodes that are typically most affected by muscular activity. Only the parts of EEG records with no visually recognizable artifacts were used for testing.

Once, the artifact free EEG signals $\mathbf{x}_{clean}[n]$ were gathered, we added the muscular artifacts obtained by a simulation on boundary element method (BEM) based realistic head model with the same electrode arrangement as the one used for the EEG measurement. We used a BEM model composed of 19440 elements arranged in 3 layers representing air-skin, skin-skull and skull-brain boundaries. This model was previously used in [24, 25] (see Fig 1 in [25]). In this simulation we placed 5 sources representing cervical muscles and 6 sources representing mandibles into the head model. Temporal signals for these sources were obtained independently for each source from separate surface EMG records measured by electrodes placed on mandibles and across neck during jaws contraction and head movements. Using this combination of EMG signals and realistic head modeling, we created EMG artifacts that could be realistically observed in EEG electrodes. Once the EMG artifacts, henceforth denoted as $\mathbf{x}_{artifact}[n]$, were generated, they were added to the artifact-free EEG, creating a mixture $\mathbf{x}[n]$ that was subsequently processed by the newly suggested and other state-of-the-art algorithms

$$\mathbf{x}[n] = \mathbf{x}_{clean}[n] + \sqrt{\eta} \cdot \mathbf{x}_{artifact}[n]. \tag{20}$$

The constant η was chosen as

$$\eta = \frac{\sum_{n=1}^N \|\mathbf{x}_{clean}[n]\|^2}{\sum_{n=1}^N \|\mathbf{x}_{artifact}[n]\|^2} \xi, \tag{21}$$

where $\|\cdot\|$ is the Euclidean vector norm, and the ξ sets the ratio between the energy of signals $\mathbf{x}_{clean}[n]$ and the energy of signals $\mathbf{x}_{artifact}[n]$. For testing we choose $\xi = 0$, $\xi = 1$ and $\xi = 4$ (the first setting provides artifact free EEG, the second setting makes the energy equal, and the last setting makes the artifact energy four times stronger).

Illustrations of signals $\mathbf{x}[n]$ obtained through this approach for $\xi = 1$ are in Figs 4 and 5. Fig 4 shows signals in temporal domain, and presents topographic maps of average power spectral density (PSD) in various frequency bands. Fig 5 shows the PSDs of EEG with simulated EMG contamination at five different scalp locations. PSDs in Fig 5 were estimated by averaging periodograms computed from 700 segments, which were 1s long and were weighted by a Hamming window.

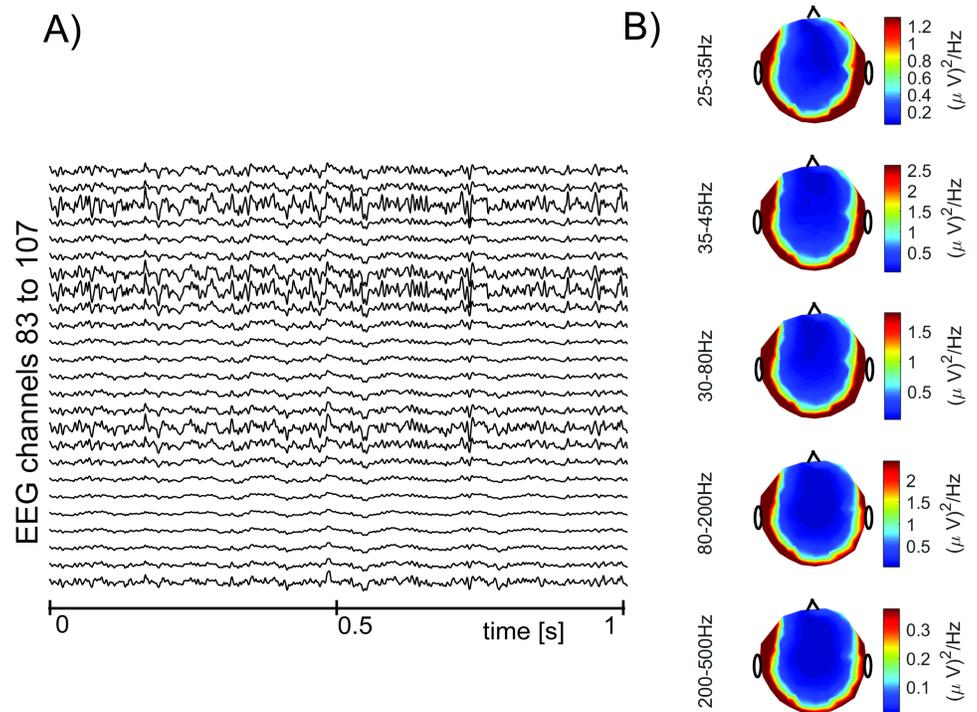


Fig 4. An example of EEG with the simulated EMG contamination. (A) 25 EEG channels with simulated EMG contamination. (B) Topographic maps of average power spectral density in various frequency bands.

<https://doi.org/10.1371/journal.pone.0201900.g004>

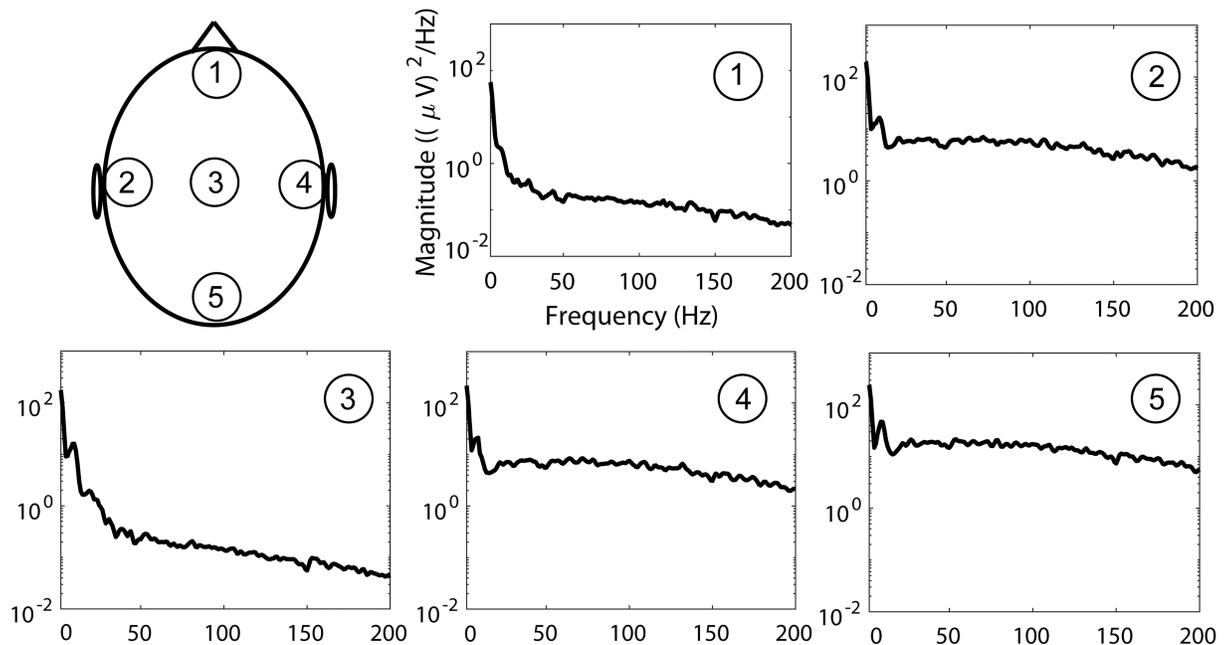


Fig 5. Examples of PSDs of EEG with simulated EMG contamination at several scalp locations.

<https://doi.org/10.1371/journal.pone.0201900.g005>

Once the reconstructed signals $\tilde{\mathbf{x}}[n]$ were obtained, the quality of their reconstruction was evaluated using an average correlation coefficient

$$r = \frac{1}{M} \sum_{m=1}^M \frac{\sum_{n=1}^N \tilde{x}_m[n] x_{clean,m}[n]}{\sqrt{\sum_{n=1}^N \tilde{x}_m^2[n] \sum_{n=1}^N x_{clean,m}^2[n]}}. \tag{22}$$

Because the values of r are random numbers, the comparison between the values of r obtained for various methods necessitates the use of statistical testing. We therefore obtained signals $\mathbf{x}[n]$ 35000 times, applied the above-mentioned methods to each of these trials, averaged the resulting values of r obtained for each trial into one mean value \bar{r} , and computed their standard deviation σ . These values were then compared with the one way ANOVA and the post-hoc Scheffe test.

The evaluation described above was used for two purposes. First, we examined the performance of the newly suggested method for various values of L and three different electrode systems (111 electrode system, 10-10 electrode system and 10-20 electrode system). The results were used to identify the values of L , for which the newly suggested method provides the best results. Second, we examined and compared the performance of all the above mentioned methods (the newly proposed method and the state-of-the-art methods)—for each method we computed the value of \bar{r} , and used these values to compare how well the individual methods remove EMG artifacts.

The evaluation is performed for EEG records that are 1s long (in practice a longer EEG record would be segmented to these shorter segments, and each segment would then be processed separately by an artifact removing algorithm). We find this segment length sufficient for the artifact removal, but short enough to accommodate the nonstationarity of EMG artifacts.

Last, to further test whether the proposed algorithm does not corrupt useful information in EEG data, we followed approach from [11], and checked whether our algorithm does not impair the detection of changes in alpha activity occurring when the measured subjects open or close their eyes. For this purpose, we used two minute long EEG records measured with subjects' eyes closed, and two minute long EEG records measured with subjects' eyes open. We used only records from the subjects that manifested distinct augmentation of alpha activity when their eyes were closed (based on visual inspection of signals in the time domain). This was the case in 13 out of 20 participants. The measured signals were processed by the proposed method using segmentation into 1 second long segments. To examine the effect of our method, we estimated PSDs before and after its application. The PSDs were computed for each subject using signals from an electrode that would correspond to the electrode OZ of the 10-10 electrode system. The PSDs estimation procedure was based on Welch's method with segmentation into 1s long segments with 75% overlap and weighting with the Hamming window. Last, the individual PSDs were averaged over subjects providing the average PSDs presented in section Results.

Results

The Fig 6 shows the values of \bar{r} computed with different subspace dimensions L for various electrode systems (ξ was chosen as 4). Note that for all electrode systems the best performance is achieved with L within interval 10-20. For further computations we chose value $L = 12$, and we suggest to use this value with future applications of our algorithm (the optimization of L does not need to be repeated).

We have also performed this optimization for the signal length of 2s and 4s, and found very similar results: the best performance was achievable for L within interval 10-20.

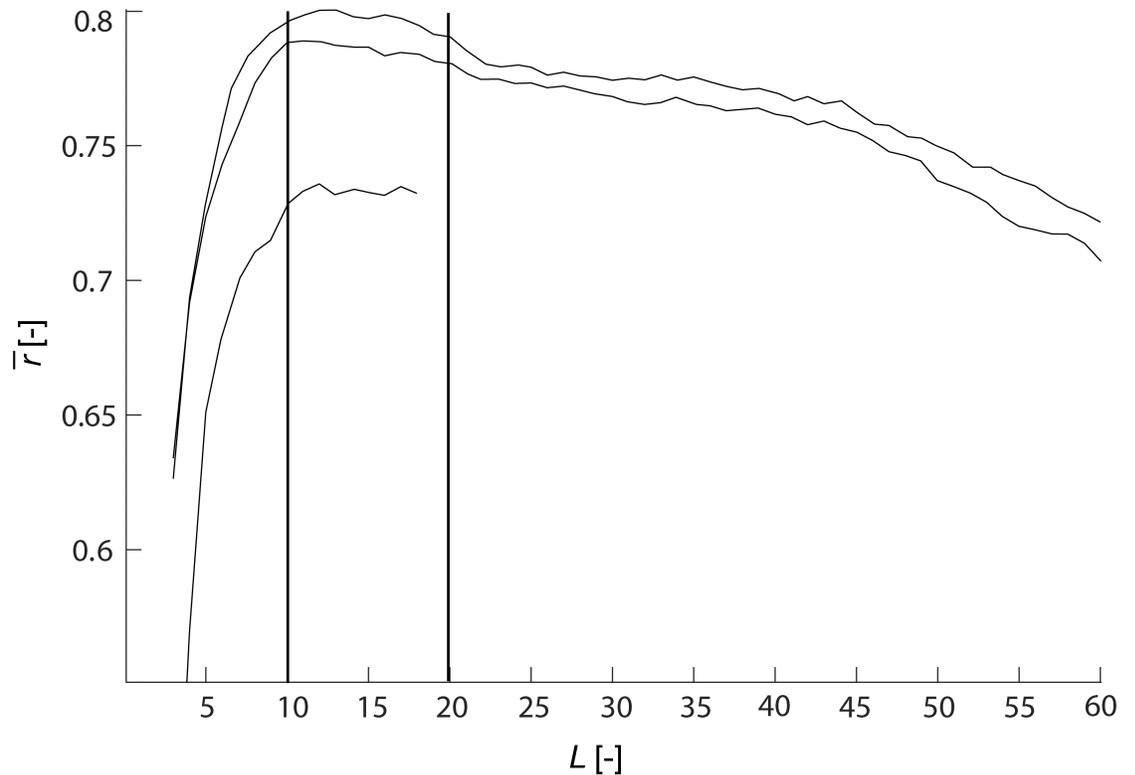


Fig 6. Mean correlation coefficients \bar{r} between the original clean EEG signals and EEG signals with suppressed muscular artifacts as a function of subspace dimension L for 111, 10/10 and 10/20 electrode systems.

<https://doi.org/10.1371/journal.pone.0201900.g006>

Tables 1, 2 and 3 show the values of \bar{r} and σ obtained for the suggested algorithm (with $L = 12$) and other state-of-the-art algorithms. The last column indicates whether a statistically significant difference (SSD) was found between the sets of values r for respective method and the newly proposed method. These results were obtained for a reference electrode placed on a forehead. We also computed these values with a common average reference and also with data

Table 1. The correlation coefficients from the processing of EEG signals without muscular artifacts ($\xi = 0$).

Method	\bar{r}	σ	SSD
None	1.0000	0.0000	no
C, 50 PCs	0.8901	0.0536	yes
C, 25 PCs	0.9209	0.0654	yes
C, 12 PCs	0.9689*	0.0465	no
D, 7 PCs	0.2295	0.0857	yes
D, 3 PCs	0.3502	0.0852	yes
D, 2 PCs	0.4053*	0.0891	yes
A	0.8495	0.0483	yes
Ba	0.8601	0.0542	yes
Bb	0.8805	0.0484	yes
Proposed method	0.9831	0.0236	-

Symbol * denotes the best achieved performance in the case of methods C and D.

<https://doi.org/10.1371/journal.pone.0201900.t001>

Table 2. The correlation coefficients from the processing of signals with equally strong muscular artifacts and EEG ($\xi = 1$).

Method	\bar{r}	σ	SSD
None	0.8738	0.0513	yes
C, 50 PCs	0.8182	0.0548	yes
C, 25 PCs	0.8253	0.0627	yes
C, 12 PCs	0.8803*	0.0536	yes
D, 7 PCs	0.2169	0.0691	yes
D, 3 PCs	0.3309	0.0832	yes
D, 2 PCs	0.3976*	0.1040	yes
A	0.7972	0.0569	yes
Ba	0.7909	0.0657	yes
Bb	0.8639	0.0328	yes
Proposed method	0.9536	0.0381	-

Symbol * denotes the best achieved performance in the case of methods C and D.

<https://doi.org/10.1371/journal.pone.0201900.t002>

re-referenced to several randomly picked electrodes and always achieved virtually identical results.

The individual state-of-the-art methods are distinguished by the following letters:

A. . . a plain FastICA [17, Eq (14, 21, 24)] used on data that were pre-whitened by PCA [22] with no overlearning suppression,

Ba. . . the high pass filtering method, using a fixed high pass filter with 1Hz cut off frequency [13, 14],

Bb. . . the high pass filtering method, using the AR model based high pass filtering [13, 14],

C. . . the PCA based method [13, 14],

D. . . the combination of the PCA and linear regression [3].

To present the best performance for the state-of-the-art methods we show the results for methods C and D with different numbers of principal components and the best performing

Table 3. The correlation coefficients from the processing of signals with muscular artifacts four times stronger than EEG ($\xi = 4$).

Method	\bar{r}	σ	SSD
None	0.5631	0.0834	yes
C, 50 PCs	0.5408	0.0951	yes
C, 25 PCs	0.5684	0.1201	yes
C, 12 PCs	0.7061*	0.1057	yes
D, 7 PCs	0.2390	0.0912	yes
D, 3 PCs	0.3831	0.1109	yes
D, 2 PCs	0.4049*	0.0734	yes
A	0.5251	0.0641	yes
Ba	0.5237	0.0561	yes
Bb	0.5879	0.0328	yes
Proposed method	0.8053	0.0694	-

Symbol * denotes the best achieved performance in the case of methods C and D.

<https://doi.org/10.1371/journal.pone.0201900.t003>

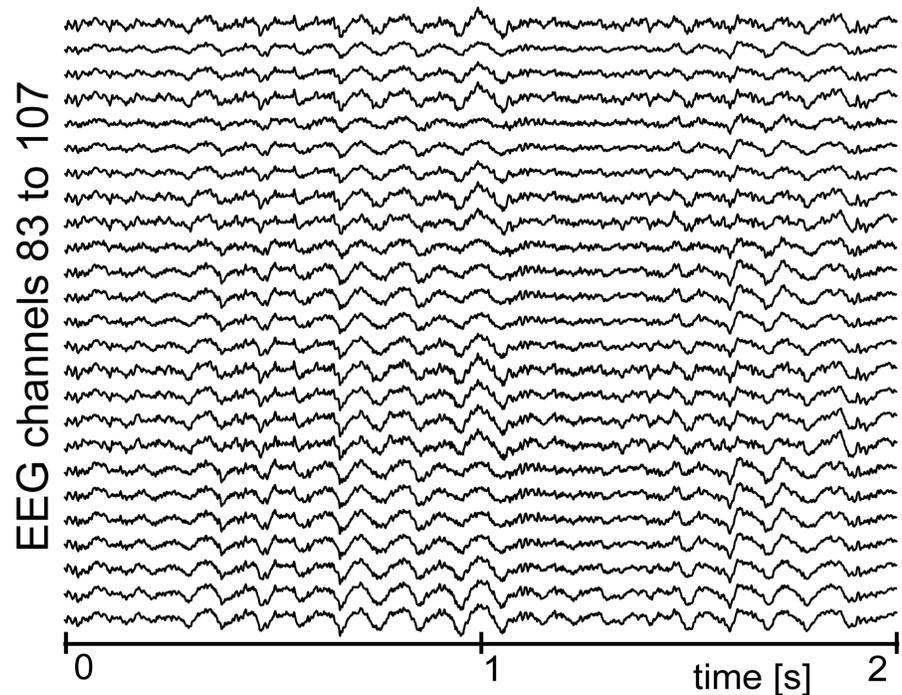


Fig 7. An example of EEG signals processed by the newly proposed method. The original signals are shown in Fig 2A. The EEG contains 111 channels, but for spatial reasons we show only 25 channels that correspond to the ones shown in Fig 2A.

<https://doi.org/10.1371/journal.pone.0201900.g007>

number of principal component (this number was found by the evaluation of all possible numbers of principal components, and cannot be realistically achieved in practice, where we have no way of knowing which result provides the highest \bar{r}).

To assess the dimensionality reduction caused by the newly proposed method, for each \mathcal{D} we computed how many strongest singular values comprise 99% of the energy of all singular values. On average, we found this to be 78 for $\xi = 0$, 67 for $\xi = 1$ and 54 for $\xi = 4$. Note that these numbers are much higher than the number of PCs retained in the application of method C (Tables 1, 2 and 3). Therefore, our approach causes much lower dimensionality reduction, and consequently smaller corruption of clean EEG data.

In addition to the quantitative analysis provided above, we also present some illustrative results achieved when our method was applied to real EEG signals.

In Fig 7 we show the results obtained when the proposed method was applied to the real EEG signals with EMG artifacts shown in Fig 2A (the signal was split to 1s long segments and our method was applied separately in each of these segments). Fig 7 shows that the EMG artifacts were nicely removed, while the effects of overlearning (shown in Fig 2B) were not observed in any of the estimated source subsets $\hat{\mathcal{S}}_k[n]$ during the processing. In addition, Fig 8 shows the topographical power distribution of these signals before and after the application of our method. In Fig 9 we also show the PSDs computed for five different positions on the scalp (these PSDs were computed using Welch's method with 0.5s long segments overlapped by 75% and weighted by the Hamming window). Note the decrease of power in fringe electrodes, which is associated with the removal of muscular artifacts. These results are similar to those previously presented in [11].

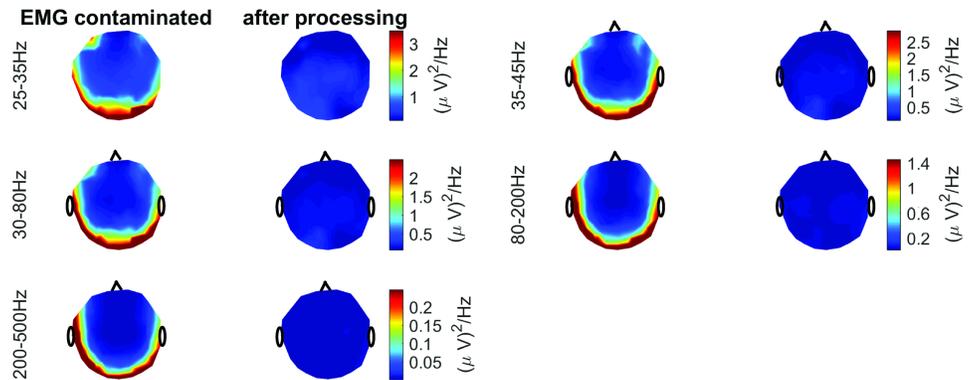


Fig 8. Topographic maps of the average power spectral density in various frequency bands before and after processing by the proposed algorithm.

<https://doi.org/10.1371/journal.pone.0201900.g008>

Overall, the results shown in Figs 7, 8 and 9 show nice suppression of EMG artifacts without any noticeable effects of overlearning. However, as stated above, this kind of application cannot be used to quantitatively evaluate the quality of EEG reconstruction; therefore, we include it for illustrative purposes only.

Last, Fig 10 illustrates the changes in the alpha band power in EEG measured on subjects with eyes open and closed. The average PSDs clearly indicate the increase of alpha band power when subjects' eyes were closed. More importantly, Fig 10 also shows that this change does not seem to be impaired by the application of our method.

Discussion

The results show that the newly proposed method outperforms all the presented state-of-the-art methods.

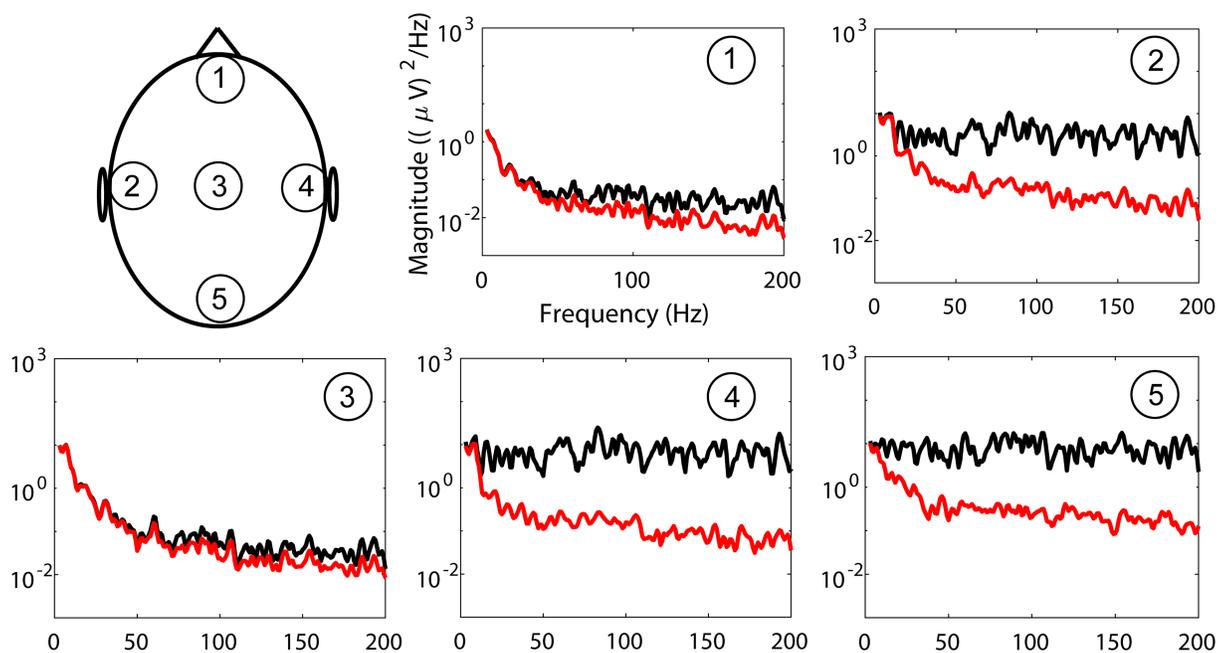


Fig 9. Power spectral densities at several scalp locations before and after processing by the proposed algorithm.

<https://doi.org/10.1371/journal.pone.0201900.g009>

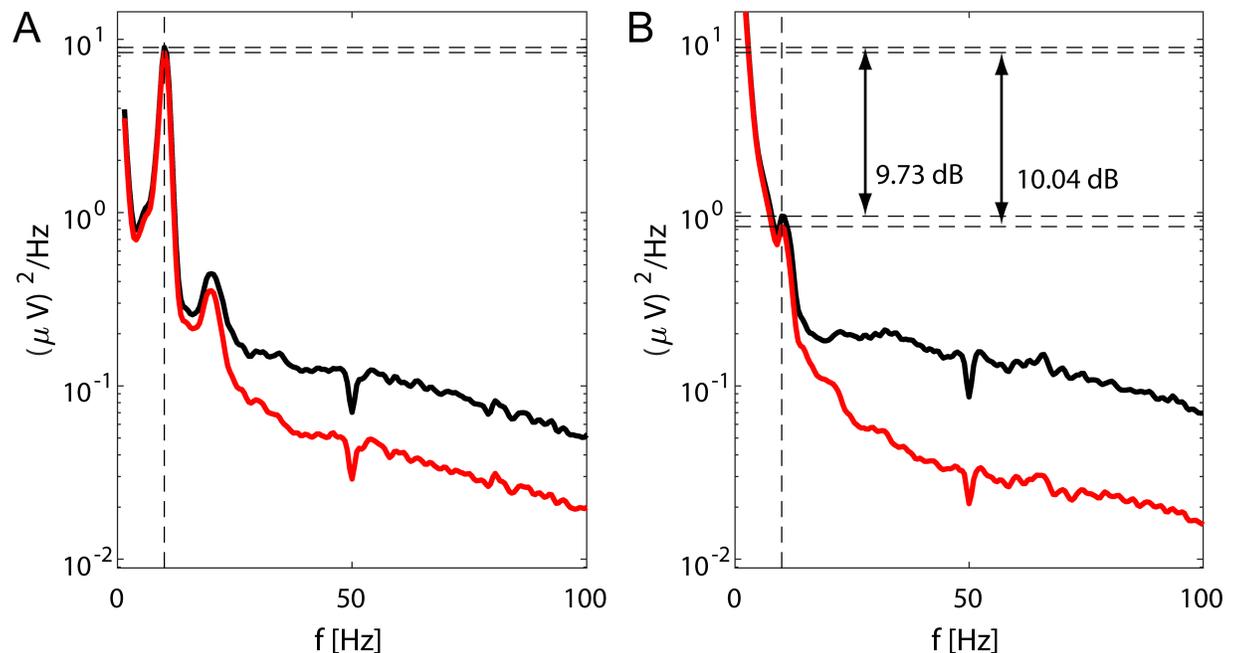


Fig 10. Average PSDs showing changes in alpha power band for eyes open and closed. (A) Average PSDs for resting EEG measured with closed eyes. (B) Average PSDs for resting EEG measured with open eyes. The black lines are PSDs computed from the unprocessed data, while the red ones are PSDs computed from the data processed by the proposed algorithm.

<https://doi.org/10.1371/journal.pone.0201900.g010>

The presented results also confirm the weaknesses of the individual state-of-the-art methods pointed out in our introduction. The PCA based dimensionality reduction (method C) distorts the EEG even if no artifacts are being removed. This can be seen as decrease of \bar{r} for clean EEG ($\xi = 0$). The high pass filtering methods (methods Ba and Bb) show consistently poor performance for all setups, because with their use the overlearning is not suppressed. Method D performs poorly for artifact free EEG ($\xi = 0$), where the strongest principal components are composed of brain signals only, and are incorrectly removed from the original EEG. For the stronger artifacts the performance improves; however, it is never as good as for the newly proposed method.

Further, we would like to discuss some additional points related to our method.

In this paper we concentrate primarily on EMG artifacts caused by transient muscular contractions (henceforth, transient EMG artifacts). Besides these artifacts, EEG can also be contaminated by EMG artifacts caused by tonic muscular contraction (henceforth, tonic EMG artifacts). The tonic EMG artifacts can be quite weak, even difficult to recognize upon visual inspection, and of little concern in some applications (e.g. see a small effect of tonic EMG artifacts on evoked response potentials in [11]).

We of course expect that our method will suppress even the tonic artifacts to some extent—compared to transient EMG artifacts they are just weaker and more stationary. Therefore, as long as the ICA is able to separate them, they will be identified and removed from the EEG. In fact, the decrease of power on higher frequencies shown in Fig 10 can likely be attributed to the suppression of tonic EMG artifacts, and this suppression is similar to the results presented in [11].

However, in principle, a stationary interfering source will be better removed using a longer EEG record. Therefore, if a researcher faces a situation where any residual tonic EMG artifacts remaining after the application of our method are still concerning, the application of our

method can be followed by the application of a ‘classical’ EMG removal procedure (e.g. [11]) that operates on longer EEG records. In this arrangement our method would be better suited to deal with transient EMG artifacts, and the ‘classical’ EMG removal procedure would be better suited to remove any residual tonic EMG artifacts. In applications where the presence of residual tonic EMG artifacts is not concerning, the application of our method alone would be sufficient.

In relation to tonic EMG artifacts, one could additionally point out that the EEG records $x_{clean}[n]$ may not be completely devoid of EMG artifacts, because the tonic EMG artifacts are always present, unless the measurement is performed in paralysis [11, 23]. However, we would argue that while their presence cannot be excluded, visually unrecognizable tonic EMG artifacts have total power much smaller than the total power of EEG, and therefore the effect of these tonic EMG artifacts on the average correlation coefficients \bar{r} will be negligible, especially when we work with transient EMG artifacts with energy equal or greater than the energy of EEG.

Another point that we would like to address is related to our EMG classifier. Note that the classifier described in section Artifact classification is not the only possible approach. For example, work [11] suggested a classifier based on the slope of a power spectrum. In fact, we have also retested our method with the classifier from [11] and obtained virtually identical results. The reader should note that the classifier itself is not an unalterable part of the proposed method, and if necessary a researcher can adjust it to suit application-specific needs.

Further, in the motivation and evaluation of our method, we have limited ourselves to the removal of EMG artifacts, while making few comments about other types of artifacts. This was mostly dictated by a need to fit our presentation into a single paper, and it does not mean that the presented method cannot be generalized for other type of artifacts. In fact, as long as the ICA is able to extract the artifacts into separate components (which was already demonstrated for many different types of artifacts [3, 4, 16]), and we are able to classify these components respectively, the algorithm will allow to remove these components without the manifestation of overlearning.

In this paper we limit ourselves to EMG artifacts only, but because EMG artifacts are an omnipresent nuisance in a great number of EEG records, to have an algorithm that removes these artifacts alone, is, in our opinion, a quite useful contribution.

Regarding the future work, there are also other blind source separation algorithms that might be used for the suppression of EMG artifacts (e.g. SOBI [26], SSA [27], CCA [28]). It may be interesting to explore their overlearning properties, and compare their performance with the ICA based ones. However, before doing so we find it important to first improve the ICA based algorithm, as we attempt to do in this paper.

Conclusion

This paper presents an ICA based algorithm for the removal of muscular artifacts that is not prone to overlearning for short signals. This allows its application to short signal segments, which makes it better suited for the removal of artifacts that are distinctly non-stationary and appear, disappear or change their origin within a time frame of a few seconds or less. Thus, the algorithm can remove the artifacts present only in short segments to which it is applied. If the artifacts are not detected, the original EEG is not distorted. If the composition of artifacts is changed, the algorithm can react quickly.

We show that the proposed algorithm outperforms all the current state-of-the-art approaches, and provides the best reconstruction of the original EEG.

The algorithm was presented for the purpose of muscular artifact removal; however, we note that it can be easily extended for the removal of other types of artifacts.

Because the muscular artifacts are a recurring problem in a great number of EEG records, we believe that the presented algorithm will provide a useful and efficient tool for their removal.

Author Contributions

Conceptualization: Radoslav Bortel.

Data curation: Jan Sebek.

Formal analysis: Radoslav Bortel.

Funding acquisition: Radoslav Bortel, Pavel Sovka.

Investigation: Jan Sebek, Pavel Sovka.

Methodology: Jan Sebek.

Project administration: Radoslav Bortel, Pavel Sovka.

Resources: Radoslav Bortel.

Software: Jan Sebek.

Supervision: Radoslav Bortel, Pavel Sovka.

Validation: Jan Sebek.

Visualization: Jan Sebek.

Writing – original draft: Jan Sebek, Radoslav Bortel.

Writing – review & editing: Radoslav Bortel.

References

1. Hyvärinen A, Oja E. Independent component analysis: algorithms and applications. *Neural networks*. 2000; 13(4):411–430. PMID: [10946390](https://pubmed.ncbi.nlm.nih.gov/10946390/)
2. Vigário R, Sarela J, Jousmiki V, Hamalainen M, Oja E. Independent component approach to the analysis of EEG and MEG recordings. *IEEE transactions on biomedical engineering*. 2000; 47(5):589–593. <https://doi.org/10.1109/10.841330> PMID: [10851802](https://pubmed.ncbi.nlm.nih.gov/10851802/)
3. Cheok NS, Raveendran P. Removal of EOG artifacts using ICA regression method. In: 4th Kuala Lumpur International Conference on Biomedical Engineering 2008. Springer; 2008. p. 226–229.
4. Mammone N, Morabito FC. Enhanced automatic artifact detection based on independent component analysis and Renyi's entropy. *Neural networks*. 2008; 21(7):1029–1040. <https://doi.org/10.1016/j.neunet.2007.09.020> PMID: [18430547](https://pubmed.ncbi.nlm.nih.gov/18430547/)
5. Ng SC, Raveendran P. Enhanced μ Rhythm Extraction Using Blind Source Separation and Wavelet Transform. *IEEE Transactions on Biomedical Engineering*. 2009; 56(8):2024–2034. <https://doi.org/10.1109/TBME.2009.2021987> PMID: [19457744](https://pubmed.ncbi.nlm.nih.gov/19457744/)
6. Crespo-Garcia M, Atienza M, Cantero JL. Muscle artifact removal from human sleep EEG by using independent component analysis. *Annals of biomedical engineering*. 2008; 36(3):467–475. <https://doi.org/10.1007/s10439-008-9442-y> PMID: [18228142](https://pubmed.ncbi.nlm.nih.gov/18228142/)
7. Castellanos NP, Makarov VA. Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis. *Journal of neuroscience methods*. 2006; 158(2):300–312. <https://doi.org/10.1016/j.jneumeth.2006.05.033> PMID: [16828877](https://pubmed.ncbi.nlm.nih.gov/16828877/)
8. Mammone N, La Foresta F, Morabito FC. Automatic artifact rejection from multichannel scalp EEG by wavelet ICA. *IEEE Sensors Journal*. 2012; 12(3):533–542. <https://doi.org/10.1109/JSEN.2011.2115236>
9. Cassani R, Falk TH, Fraga FJ, Kanda PA, Anghinah R. The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis. *Frontiers in aging neuroscience*. 2014; 6. <https://doi.org/10.3389/fnagi.2014.00055> PMID: [24723886](https://pubmed.ncbi.nlm.nih.gov/24723886/)

10. Fitzgibbon S, DeLosAngeles D, Lewis T, Powers D, Whitham E, Willoughby J, et al. Surface Laplacian of scalp electrical signals and independent component analysis resolve EMG contamination of electroencephalogram. *International Journal of Psychophysiology*. 2015; 97(3):277–284. <https://doi.org/10.1016/j.ijpsycho.2014.10.006> PMID: 25455426
11. Fitzgibbon S, DeLosAngeles D, Lewis T, Powers D, Grummett T, Whitham E, et al. Automatic determination of EMG-contaminated components and validation of independent component analysis using EEG during pharmacologic paralysis. *Clinical Neurophysiology*. 2016; 127(3):1781–1793. <https://doi.org/10.1016/j.clinph.2015.12.009> PMID: 26780994
12. Hyvärinen A, Särelä J, Vigário R. Spikes and bumps: Artefacts generated by independent component analysis with insufficient sample size. In: *Proc. Int. Workshop on Independent Component Analysis and Signal Separation (ICA'99)*; 1999. p. 425–429.
13. Särelä J, Vigário R. The problem of overlearning in high-order ICA approaches: analysis and solutions. In: *International Work-Conference on Artificial Neural Networks*. Springer; 2001. p. 818–825.
14. Särelä J, Vigário R. Overlearning in marginal distribution-based ICA: analysis and solutions. *Journal of machine learning research*. 2003; 4(Dec):1447–1469.
15. Delorme A, Sejnowski T, Makeig S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage*. 2007; 34(4):1443–1449. <https://doi.org/10.1016/j.neuroimage.2006.11.004> PMID: 17188898
16. Jung TP, Makeig S, Humphries C, Lee TW, Mckeown MJ, Iragui V, et al. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*. 2000; 37(2):163–178. <https://doi.org/10.1111/1469-8986.3720163> PMID: 10731767
17. Hyv A, et al. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*. 1999; 10(3):626–634. <https://doi.org/10.1109/72.761722>
18. Shwedyk E, Balasubramanian R, Scott R. A nonstationary model for the electromyogram. *IEEE Transactions on Biomedical Engineering*. 1977;(5):417–424. <https://doi.org/10.1109/TBME.1977.326175> PMID: 892834
19. Hyvärinen A, Karhunen J, Oja E. *Independent component analysis*. John Wiley & Sons; 2004.
20. Zeng K, Chen D, Ouyang G, Wang L, Liu X, Li X. An EEMD-ICA approach to enhancing artifact rejection for noisy multivariate neural data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2016; 24(6):630–638. <https://doi.org/10.1109/TNSRE.2015.2496334> PMID: 26552089
21. Perrin F, Bertrand O, Pernier J. Scalp Current Density Mapping: Value and Estimation from Potential Data. *IEEE Transactions on Biomedical Engineering*. 1987; 34(4):283–288. <https://doi.org/10.1109/TBME.1987.326089> PMID: 3504202
22. Comon P. Independent component analysis, a new concept? *Signal processing*. 1994; 36(3):287–314. [https://doi.org/10.1016/0165-1684\(94\)90029-9](https://doi.org/10.1016/0165-1684(94)90029-9)
23. Whitham EM, Pope KJ, Fitzgibbon SP, Lewis T, Clark CR, Loveless S, et al. Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical Neurophysiology*. 2007; 118(8):1877–1888. <https://doi.org/10.1016/j.clinph.2007.04.027> PMID: 17574912
24. Bortel R, Sovka P. Electrode position scaling in realistic Laplacian computation. *IEEE Transactions on Biomedical Engineering*. 2008; 55(9):2314–2316. <https://doi.org/10.1109/TBME.2008.921168> PMID: 18713701
25. Bortel R, Sovka P. Potential approximation in realistic Laplacian computation. *Clinical Neurophysiology*. 2013; 124(3):462–473. <https://doi.org/10.1016/j.clinph.2012.08.020> PMID: 23040294
26. Belouchrani A, Abed-Meraim K, Cardoso JF, Moulines E. A blind source separation technique using second-order statistics. *IEEE Transactions on signal processing*. 1997; 45(2):434–444. <https://doi.org/10.1109/78.554307>
27. Hara S, Kawahara Y, Washio T, Von BüNau P, Tokunaga T, Yumoto K. Separation of stationary and non-stationary sources with a generalized eigenvalue problem. *Neural networks*. 2012; 33:7–20. <https://doi.org/10.1016/j.neunet.2012.04.001> PMID: 22551683
28. Janani AS, Grummett TS, Lewis TW, Fitzgibbon SP, Whitham EM, DelosAngeles D, et al. Improved artefact removal from EEG using Canonical Correlation Analysis and spectral slope. *Journal of neuroscience methods*. 2018; 298:1–15. <https://doi.org/10.1016/j.jneumeth.2018.01.004> PMID: 29408174