Czech Technical University

Faculty of Electrical Engineering

Doctoral Thesis

April 2019   Jan Šebek

Czech Technical University

Faculty of Electrical Engineering

Department of Circuit Theory

Independent Component Analysis for EEG Pre-Processing

Doctoral Thesis

Jan Šebek

CONTENTS

Čestné prohlášení

Jméno a příjmení doktoranda: Jan Šebek

Prohlašuji, že jsem svou disertační práci vypracoval samostatně a v předložené práci důsledně citoval použitou literaturu.

Předložená práce a prezentované výsledky vznikly bez autorské spoluúčasti jiných osob.

Praha, dne: Podpis:

ABSTRACT

In this work, the major problems, like insufficient sample size, of independent component analysis (ICA) based algorithms used for electroencephalographic (EEG) processing are addressed and several robust algorithms that mitigate their negative effects on source separation performance are proposed. The performance of the proposed as well as previously published algorithms is assessed while performing the task of muscular artifact removal from short, high density EEG records.

The studied problems are denoted as ICA overlearning in the case of limited sample size and a violation of ICA model, when ICA is applied on a signal set where the number of source signals is higher than the number of measured signals, respectively. It was previously shown that due to the ICA overlearning, the ICA completely fails to separate brain signals and muscular artifacts for short EEG records with a high number of channels. Using the ICA when the theoretical ICA model is violated can further seriously degrade the quality of achieved separation.

These problems are addressed by projecting an EEG record into several subspaces with a lower dimension, performing the ICA on each subspace separately and checking the quality of separation to avoid subspaces with a low separation quality. Due to the reduced dimension of the subspaces and their automatically supervised selection, the overlearning is suppressed, ICA model violation avoided, 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. It is shown that for short signals and a high number of EEG channels presented approach outperforms the currently available ICA based algorithms for the muscular artifact removal.

Keywords: Electroencephalogram, Muscular artifacts removal, Independent Component Analysis, High density electrode arrays, ICA overlearning.

ABSTRAKT

Tato práce se zabývá hlavními problémy algoritmů analýzy nezávislých komponent (ICA), to jest například omezeným množstvím vzorků signálu, při zpracování elektroencefalografických (EEG) záznamů a návrhem robustních algoritmů, jejichž výsledná kvalita separace zdrojů není těmito problémy ovlivněna. Kvalita separace nově navržených i již dříve publikovaných algoritmů je porovnána v rámci klasické úlohy odstraňování svalových artefaktů z EEG záznamů s tím rozdílem, že jsou využity extrémně krátké záznamy z měřícího systému s vysokou hustotou elektrod.

Uvedené problémy jsou často v literatuře označovány jako ICA overlearning a porušení modelu ICA. Již dříve bylo v literatuře uvedeno, že při jevu ICA overlearning ICA algoritmy kompletně selhávají při separaci mozkových a svalových signálů pro krátké záznamy EEG signálu s větším počtem kanálů. Při porušení modelu ICA algoritmu může být výsledná kvalita separace dále závažným způsobem snížena.

Tyto problémy jsou řešeny pomocí projekce EEG záznamu do několika podprostorů s nižší dimenzí než je tomu u originálního zpracovávaného záznamu. V každém z těchto podprostorů je zvlášť provedena analýza nezávislých komponent a zaveden mechanismus kontroly kvality separace kvůli identifikaci a vyřazení podprostorů s nízkou kvalitou separace zdrojů. Kvůli snížené dimenzi jednotlivých podprostorů a jejich následnému výběru pro další zpracování je jev ICA overlearning potlačen a zároveň je menší i vliv druhého problému porušení ICA modelu, což má za následek kvalitnější separaci svalových artefaktů. Po odstranění svalových artefaktů jsou signály z jednotlivých podprostorů zpětně rekonstruovány do prostoru originálního měřeného záznamu a pomocí lineárních kombinací je zpětně sestaven měřený záznam s odstraněnými artefakty. V práci je názorně předvedeno, že pro signály s krátkým trváním a vysokým počtem elektrod navržené metody poskytují vyšší kvalitu rekonstrukce čistého EEG signálu než je tomu u momentálně známých ICA algoritmů pro odstraňování svalových artefaktů z EEG záznamů.

Klíčová slova: Elektroencefalogram, Odstraňování svalových artefaktů, Analýza nezávislých komponent, Měřící systémy s vysokou hustotou elektrod, Přeučení algoritmu analýzy nezávislých komponent.

1. INTRODUCTION

1.1 Motivation

Electroencephalography (EEG) records are often contaminated by several types of interference. Many of these unwanted signal components (artifacts) are often hard to get rid off by traditional techniques like fixed or adaptive filtering because of their spectra, which are overlapping with the EEG bands, and relatively high power compared to the brain signals. So called muscular or electromyographic (EMG) artifacts are generated by the contractions of muscles in the vicinity of the scalp. Based on their origin, two major types of muscular artifacts are recognized in the literature. First type is recognized as a muscular artifacts generated by transient movements of muscles around head, hence further denoted as transient muscular artifacts [37]. These artifacts are often very strong, highly non stationary and can completely mask the EEG activity around the scalp. The second type is recognized as muscular artifacts generated by a tonic muscular contraction. As opposed to the first type, these artifacts are typically stationary and usually weaker, so that it is difficult to recognize their presence in signal visually [37, 12].

Due to their frequent occurrence and a high potential to compromise EEG analysis [37], the aim of this work is to derive good methods for their removal. For this purpose, the algorithms which are based on a statistical technique called independent component analysis (ICA) are designed, since this technique is able to distinguish signals with overlapping frequency spectra unlike traditional or adaptive filtering.

1.2 State of the Art

The ICA algorithms are a special family of spatial filtering techniques, which estimate linear combinations of EEG signals in an analyzed signal set such that the mutual dependency between the signals resulting from these linear combinations is minimized. The main assumption about the nature of analyzed signals is that they were generated by linear combinations of underlying source signals. This mixing process is something that ICA is trying to invert, i.e. separate source signals from the linear mixtures which are being analyzed. Another critical assumptions about analyzed signals under which the ICA algorithms were derived and that are required so that optimal separation performance is guaranteed are as follow.

1. The sources are expected to be mutually independent,

2. their linear combinations which generate analyzed EEG signals are instantaneous,

3. and the number of underlying sources is equal to the number of analyzed signals (measurement channels).

The standard algorithms, which utilize independent component analysis for muscular artifact removal [16, 36, 21, 20, 22, 9, 7, 19, 6, 13, 12], were proposed in many works and are essentially based on the following three steps.

1. The analyzed EEG data with muscular artifacts are processed by the ICA algorithm and the square mixing matrix is estimated. With the use of this estimate, one attempts to invert the mixing process and consequently to estimate the underlying source signals.

2. The classifier in various forms like, for example, thresholding of the spectral slope [12], kurtosis, or entropy [10, 20] is employed to recognize, which of these estimates of source signals contain majority of muscular contamination.

3. The last part of this algorithm is responsible for projecting all of the useful sources back to the space of measured signals with the use of mixing matrix, which was estimated in the first step.

While this technique was demonstrated as a good approach for this specific task of muscular artifact removal, ICA performance can however severely deteriorate under certain circumstances. The most two critical ones, which are further studied in this work are the following. First, if the signal length is not sufficient with respect to the number of EEG channels, so called overlearning of ICA [17, 25, 26] occurs, which prevents the separation and subsequent removal of EMG artifacts. Second, the quality of source separation can be further compromised if the EEG dataset do not follow the above stated theoretical ICA model assumptions.

Let us now identify, which of the ICA model assumptions are critical for the ICA separation performance. While there is nothing one can do with the nature of the source signals, which can be actually partly related to each other as in the case of EEG and EMG signals, as one of the common tasks of EEG analysis is to estimate the coherence between EEG and EMG [5], ICA can be still employed to a task of maximizing the mutual source independence and it was already shown in all artifact rejection techniques above, that this approach still works and provides reasonable source estimates. The second assumption about an instantaneous mixing process is also not posing a severe challenge as it was proved that all signals are propagated very rapidly across the scalp and therefore no delay, which would lead to the convolution of source signals rather than to the instantaneous mixture, can be observed in the measured data [16]. Last assumption can be actually often violated when the number of underlying sources is higher than the number of measured signals, which leads to a rectangular mixing matrix, which is not easily invertible by a standard square separating matrix estimated by standard ICA or other BSS algorithms [11] and is therefore of a major concern for any of further presented developed techniques.

In the reminder of this Chapter the problems of ICA overlearning and ICA model violation are more closely specified, and a review of the current state-of-the-art methods, in which these problems are addressed, is provided.

1.3   The Overlearning Phenomenon

The above-mentioned algorithm for the removal of muscular artifacts was successfully used before in the case of measured signals which were sufficiently long compared to the number of measured channels. However, if the sample size is insufficient, this approach tends to produce artifact signals like spikes or bumps instead of actually separated sources [26]. The following two figures (Fig. 1.1, and Fig. 1.2) from [33] attempt to illustrate these situations. The purpose of this demonstration is to illustrate the effects

of the overlearning on actual source estimates rather than to provide mathematical definition of methods which will be defined in Chapter 2.

Fig. 1.1A shows a 2 second long record of 9 EEG signals from electrodes located in the frontal area of scalp with a clearly visible muscular artifact in the middle part. These signals were processed by the above mentioned algorithm for the removal of muscular artifacts using ICA (details about the utilized ICA algorithm and a classifier can be found in Sections 2.1, and 2.3 and in [33]). Fig. 1.1B shows the separated components, i.e. estimates of underlying source signals, and Fig. 1.1C shows the reconstructed signals after application of classification and reconstruction steps of the above mentioned method for suppression of EMG artifacts. In this case the separation of brain signals and muscular artifacts was successful, which consequently led to a pleasantly looking reconstruction of the original EEG without artifacts.

In Fig. 1.2, another processing attempt using the same algorithm as in the first case is demonstrated. The above mentioned method for suppression of EMG artifacts was applied on another 2s long record originating from a 111 electrode array covering the frontal, parietal, temporal and occipital regions of subject's head with a clearly visible muscular artifact. The original signals (shown in Fig. 1.2A) were processed by the ICA algorithm and source estimates are illustrated in Fig. 1.2B.

This time no clear separation between EEG and EMG was obtained and the resulting waveforms resemble neither EEG nor EMG, and are merely composed of occasionally repeated 'bumps'.

A phenomenon known as overlearning [17, 25, 26] was nicely depicted in Fig. 1.2. The overlearning occurs when the ICA algorithm has too many degrees of freedom. As a consequence, the algorithm finds estimates of source signals for which the estimated mutual dependence is even lower than in the case of estimates which would contain separated EEG and EMG signals.

Up to this date, the problem of overlearning was tackled in several studies and following approaches of how to mitigate its effects were suggested.

First and the most commonly utilized way is to use longer EEG records. This, however, should not be preferred approach. While the above-mentioned method might be able to suppress muscular artifacts, it will inevitably distort the reconstructed EEG. For example, by removing 4 of the 9 source estimates in the processing case shown in Fig. 1.1, the rank of the reconstructed data was reduced from the original 9 to 5 [1]. Thus even though the result may be looking adequate, there is surely some loss of information in the reconstructed signals. Therefore, the application of muscular artifact removal and the corresponding rank reduction should be limited only to the time intervals, where the muscular artifacts have really occurred, avoiding unnecessary distortion of the artifact free EEG.

Moreover, the interval length should follow the time duration during which the EMG artifact origin position (and the corresponding mixing matrix) is close to stationary. These durations were reported to be often less than a few seconds [34].

Following approaches were developed with a goal of overlearning suppression when only a contaminated part of EEG was processed.

Works [15, 17, 25, 26] suggested to use those contrast functions in the ICA algorithm that are more robust than a classical kurtosis as a measure of a source mutual independence. However, the improvement provided is often not sufficient to completely

_____

[1] note that by the rank of an M dimensional signal denoted by vector $\mathbf{x}[n]$ with length N, reader should understand the rank of the matrix $[x_m[n]]_{m=1...M, n=1...N}$
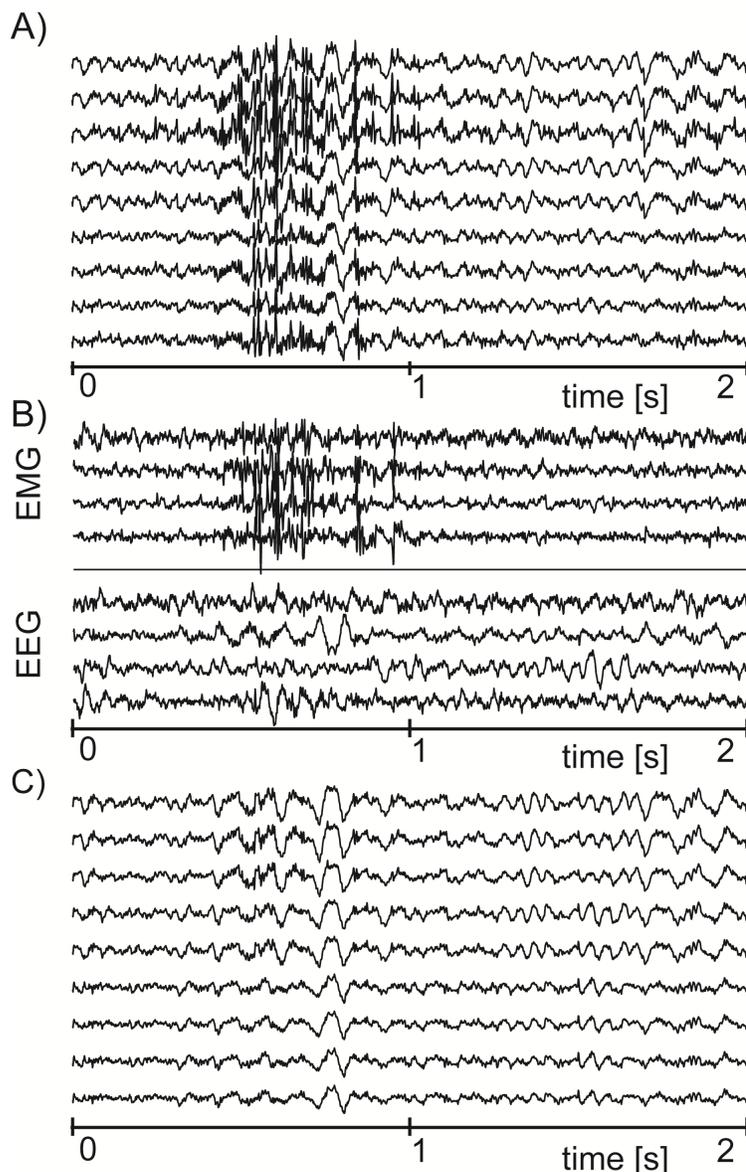
Fig. 1.1: An example of a muscular artifact suppression using the ICA. Plot A shows the original EEG signals corrupted by EMG artifacts. Plot B shows the separated source components, which were classified as brain signals (denoted as EEG) and muscular artifacts (denoted as EMG). Plot C shows the reconstructed EEG signals with muscular artifacts suppressed.

solve the problem. In processing attempt demonstrated in Fig. 1.2, the FastICA algorithm was applied with the use of robust approximation of negentropy as the contrast function (which was suggested as the best choice in [17]); however, the overlearning is still distinctly present.

Works [25, 26] suggested to use the Principal Component Analysis to decrease the dimensionality of a group of analyzed signals as a preprocessing for the ICA computation. Given that the dimension reduction is sufficient, this approach will surely suppress overlearning; however, such processing step will introduce additional EEG distortion. The brain signals may have smaller power than the EMG artifacts, thus the dimensionality reduction will inevitably remove some energy from EEG, which may cause some
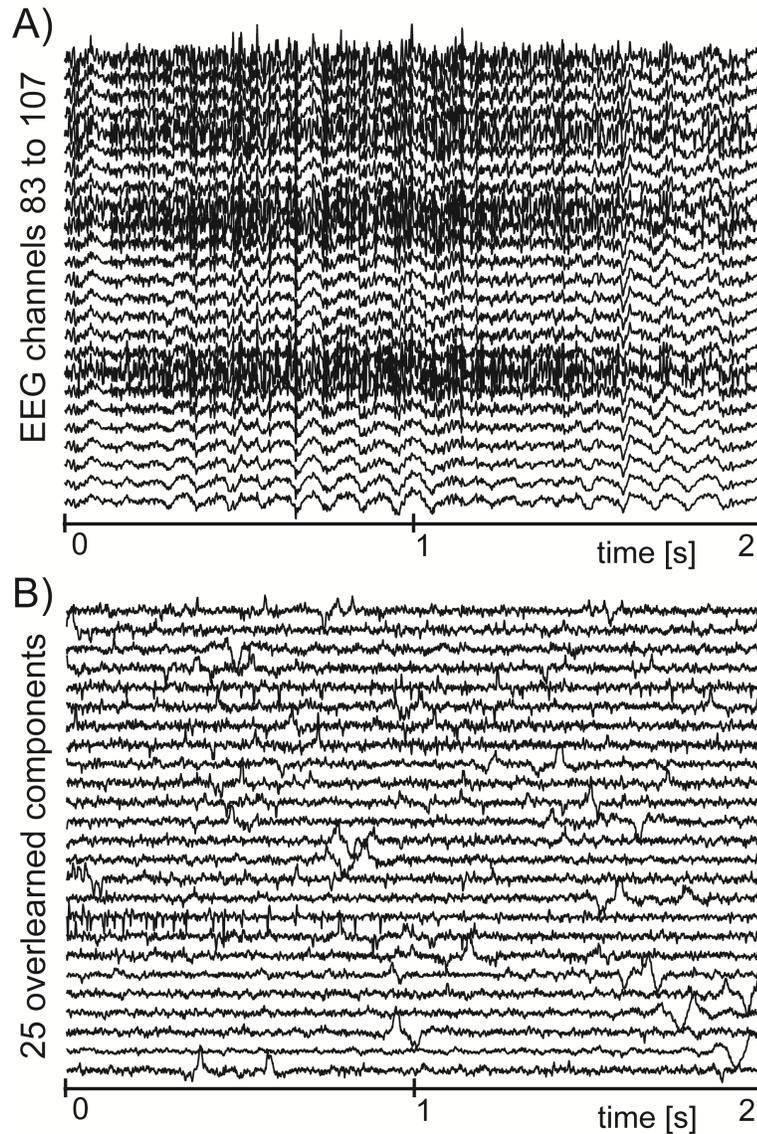
Fig. 1.2:   An example of a failed attempt to separate muscular artifacts from 111 channel EEG record (for spatial reasons only 25 EEG channels are shown). Plot A shows EEG signals corrupted by EMG artifacts. Plot B shows the source components provided by ICA (again for spatial reasons only 25 out of 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).

useful information to be missing in a subsequent EEG analysis. Therefore, the dimension reduction carried out by PCA is not always desirable, especially, if there is a way to avoid it.

Works [25, 26] also noted that the 'bumps', which are present in the separated components produced by overlearned ICA, are relatively slow compared to other signal contents and may have a dominant part of their energy associated with lower frequencies. It was therefore suggested to remove these low frequency components from the analyzed signals prior to the ICA application, which would lessen the overlearning of remaining signal contents. For this purpose, either a 1Hz fixed cut-off frequency high-pass filter or

an AR-process based high-pass filter was suggested. In my experience, this approach is not sufficient. In Fig. 1.2, the signals were already preprocessed by a high pass filtering with a 1Hz cut-off filter (to suppress the baseline wonder), but the overlearning has still occurred. The observed 'bumps' in the separated components were simply faster (with duration of about 50ms); therefore, their occurrence could not be prevented by prior high pass filtering without damaging useful brain signals.

Work [21] concentrated on the suppression of ocular artifacts. It was suggested to first extremely reduce the dimension of analyzed EEG signal with a use of PCA and to retain only 3 principal components with the highest power. Next steps were to separate these components using the ICA, and then subtract them from the original signals using a linear regression. This method uses no artifact classification. Rather, it assumes that the ocular artifacts will have the highest power, and thus will mostly populate the principal components with highest energies. Due to a large dimension reduction facilitated by the PCA, this methodology was presented as a robust to the overlearning. Its general use for other types of artifacts is however compromised, since 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. Even though, it does not seem to be a useful technique for other than ocular artifacts, it is another example of the PCA driven dimension reduction.

Last, there are some methods which build upon the basic algorithm for removal of muscular artifacts, which was briefly described in Section 1.2. The wavelet enhanced ICA [7, 19] enhances the brain signal selection by adding wavelet filtering. Surface Laplacian [13] was also proposed to filter the EEG channels after the three steps of the above mentioned basic algorithm were applied. These techniques, however, do not deal with the overlearning, which takes place in the separation step (1.2). Another technique, the ensemble empirical mode decomposition and subsequent ICA [38], first decomposes each of the measured EEG signals into the intrinsic mode functions, and then the ICA is applied. As the number of signals separated by ICA increases for this method, it is consequently making the effects of overlearning more severe.

In summary, the above-mentioned methods from the past studies do not seem to provide an efficient way to suppress the ICA overlearning without problematic side effects. New approach to the ICA overlearning suppression was therefore developed. All newly derived algorithms seem to outperform the above-mentioned methods, and provide better EEG reconstruction of short high-density EEG records with EMG artifacts.

1.4 The Violation of Theoretical ICA Model

Earlier in the Section 1.2, it was stated that the most often violated assumption of the ICA model is the number of sources, which is assumed to be equal, but is often actually higher than the number of measured EEG signals. In this Section, the mathematical definition of the ICA model is first provided and then the distortion caused by this specific model violation illustrated on a simple example with 2 measured signals and 3 sources from [11].

Let us firstly denote EEG signals measured by M EEG electrodes as $\mathbf{x}[n] = [x_1[n], \ldots, x_m[n], \ldots, x_M[n]]^T$, $n = 1, \ldots, N$. For the purpose of ICA model definition,

it is assumed that these records follow a linear mixture model

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n], \tag{1.1}$$

where $\mathbf{s}[n] = [s_1[n], \ldots, s_{M'}[n]]^T$ denote underlying source signals and $\mathbf{A} = [a_{k,l}]_{k,l}$ denotes an MxM' mixing matrix. While a common assumption in the derivation of the ICA is that M is equal to M', this cannot be realistically expected in the addressed problem. The source signals $s_m[n]$ are further assumed to be either generated by a brain or by other interfering sources (e.g. contracted muscles).

The ICA allows to estimate the separation matrix $\mathbf{B} = [b_{k,l}]_{k,l}$ that transforms signals $\mathbf{x}[n]$ into source components $\hat{\mathbf{s}}[n] = [\hat{s}_1[n], ..., \hat{s}_M[n]]^T$, which have their mutual dependency minimized

$$\hat{\mathbf{s}}[n] = \mathbf{B}\mathbf{x}[n]. \tag{1.2}$$

Now the violation of ICA model characterized by the number M' of source signals, which is greater than the number M of signals $\mathbf{x}[n]$ in the analyzed linear mixture, is of concern in this work, because the mixing process described by (1.1) and its rectangular mixing matrix A is then not simply invertible by a square demixing matrix $\mathbf{B}$ in (1.2). That means that all of the estimated source signals are distorted no matter how many of them are being actually estimated by the ICA. In many applications such as in the EMG artifact removal, one does not have to be necessarily interested in the estimation of all sources in the mixture. All that is needed in this particular task is to separate EEG from EMG; the individual EEG signals do not need to be separated, and neither the individual EMG sources. Therefore, consequently the number of estimated sources, which would be the same as the number of measured signals, might be more than sufficient for a particular task. Nevertheless, this ICA model violation has to be taken into account, because of its potentially severe degradation of the separation performance.

This can be demonstrated on a following example from work [11] with 3 source signals $\mathbf{s}[n] = [s_1[n], ..., s_3[n]]^T$, and 2 measured signals $\mathbf{x}[n] = [x_1[n], x_2[n]]^T$. When trying to estimate source signals by employing (1.2), one can easily come to the following set of conditions, which are required for the perfect reconstruction of any single source with index m that is the source estimated by using m-th row of separation matrix $\mathbf{B}$

$$\begin{aligned}
b_{m1} \cdot a_{1m} + b_{m2} \cdot a_{2m} &= c, \\
b_{m1} \cdot a_{1j} + b_{m2} \cdot a_{2j} &= 0, \\
b_{m1} \cdot a_{1k} + b_{m2} \cdot a_{2k} &= 0,
\end{aligned} \tag{1.3}$$

where $j \neq m \neq k$, and c is a non-zero real value. From this set of linear equations, it is obvious that perfect source reconstruction is not possible even in this low dimensional case with exception of a trivial case when $a_{1j} = a_{1k}$ and $a_{2j} = a_{2k}$ that is when the spatial distribution across measuring electrode array of two other sources is the same. Thus, the quality of all estimated sources and any further processing can severely deteriorate [11, 2].

Approaches labeled as overcomplete ICA like the ones based on Bayesian learning rule called geometrical ICA were proposed to minimize the inter-source interference [35, 1] for a simple cases of low dimensional signals. However, the authors of these works also stated that for high dimensional signals, the number of required samples for these algorithms is growing very rapidly, which reintroduces the problem of ICA overlearning in short data segments and useful data rank reduction in longer EEG records and potentially limits their practical usability.

Another possible effect of this violation of the ICA model can be a change of the properties of an objective space, in which the ICA is trying to find the maximum of its objective (for example negentropy). Such change can generate several additional local maxima of the ICA objective function and the iterative ICA algorithms like FastICA will then produce different results for different initializations. The interesting idea, which was introduced for example in [24], of how to deal with this problem is based on several random re-initializations of a specific iterative ICA algorithm, which then tends to converge to different source estimates with a varying level of inter-source interference at each estimation attempt. While in their work, the ICA re-initialization was used to search for few dominant sources, which were present in every estimation attempt, it is of interest in this work, because the ICA re-initialization principle will be used to choose the estimation attempt with the highest quality of source separation.

Algorithms presented in this work are therefore proposed with respect to a mitigation of deteriorating effect of ICA model violation on source signal separation. The number of estimated sources actually stays at value M for all proposed algorithms as all sources are not necessarily needed to be estimated during the task of artifact removal, especially in the case of modern high density electrode measurement systems, which provide high number of channels resulting in a sufficient number of estimated sources, so that all artifactual contents can be estimated and separated from useful signal contents.

1.5 Statement of Aims

The overall goal of this work is to improve ICA based BSS algorithm for suppression of EMG artifacts in EEG and to make this algorithm more robust against the previously mentioned ICA overlearning and the ICA model violation. To reach this goal, the following aims are defined.

The first aim is to propose a technique for ICA overlearning suppression based on a dimensionality reduction principle, since this state of the art method yields the most promising results. The newly proposed technique is however required to utilize this principle in a way which minimizes the amount of lost useful data. Consequently, this leads to a multiple subspace projection principle.

The second aim is to enhance the proposed method from the first aim by adding several estimation attempts and a separation quality control mechanism to each projection subspace. These added features are intended to discover subspaces where ICA demonstrates convergence problems or where convergence yields unsatisfactory separation performance, which can be also attributed to the ICA model violation. By identification of such subspaces and their subsequent rejection from further processing, it is expected that the separation quality degradation coming from the ICA model violation is also suppressed.

The third aim is to develop an initialization tool for ICA algorithm, which will ideally let ICA to start its estimation close to the optimum convergence point. This tool is derived based on the objective which was used as a separation quality control mechanism in the second aim. I plan to study its potential for performance enhancement when being a part of the processing chain of the algorithm from second aim or to come up with other structure of cooperation framework with ICA which will yield the best performance results.

The final aim is to merge the findings collected from the previous points and to identify the algorithm with the best overall quality of clean EEG reconstruction.

1.6   Organization of This Work

This work is organized as follows.

First, the classical use of ICA for the task of muscular artifact removal from EEG records, which is used as the reference for the performance assessment of all studied algorithms, is presented in Section 2.1. The detailed description of the utilized ICA algorithm and classifier follows in Sections 2.2 and 2.3 respectively.

Second, proposed algorithms are presented in Chapter 2 in the same order as they were derived, so the reader can easily follow the chain of introduced ideas. At the end of this chapter, the methodology for performance assessment of each proposed BSS algorithm is presented.

Third, the first section of Chapter 3 provides the results of optimization of EEG subspace size, required for nearly all derived algorithms. The second section consists of each algorithm performance evaluation and comparison via correlation and statistical analysis. Next, the impact of each algorithm on EEG contents is illustrated. This illustration is two-fold. Each algorithm is firstly applied to a real EEG record with muscular artifact contamination and its impact on temporal traces of EEG as well as its spectral profile and power distribution along the scalp surface in several spectral bands is shown. The second part of this illustration depicts the effect of processing on EEG data itself. Specifically, it is tested whether the proposed algorithm does not impair the detection of well-known alpha augmentation neurology phenomenon. Last, an impact of each method on the spectral profile and the detection of alpha suppression phenomenon is compared for each algorithm.

Fourth, Chapter 4 discuses the achieved results.

2. METHODS

This Chapter is organized as follows. Sections 2.1, 2.2, and 2.3 provide detailed description of a state-of-the-art ICA based algorithm for EMG artifact suppression. Section 2.1 describes the common framework, Section 2.2 details the ICA algorithm and Section 2.3 describes an EMG artifact classifier.

These State-of-the-art techniques are then further developed and improved in Sections 2.4, 2.5, 2.6, and 2.7, where newly developed methods are disclosed.

Last, Section 2.8 presents newly developed techniques for the evaluation of EMG artifact removal algorithms.

2.1 Classical ICA Based Algorithm for EMG Suppression

In Section 1.2, the classical ICA based algorithm for the muscular artifacts removal was briefly outlined as having three consecutive steps, which are separation of source components, their classification, and reconstruction of useful brain signal components to form EEG source components record.

The specific ICA algorithm is utilized during the first step. Based on the M observed EEG channels $\mathbf{x}[n] = [x_1[n], \ldots, x_M[n]]^T$, $n = 1, \ldots, N$, separation matrix $\mathbf{B} = [b_{k,l}]_{k,l}$ is estimated to project muscular contamination contents and useful brain activity contents to different source components $\hat{s}_m[n]$ [1].

$$\hat{\mathbf{s}}[n] = \mathbf{B}\mathbf{x}[n], \tag{2.1}$$

where $\hat{\mathbf{s}}[n] = [\hat{s}_1[n], \ldots, \hat{s}_M[n]]^T$. If the sample size in each channel of measured signals $\mathbf{x}[n]$ is sufficient, a good separation of brain signals and artifacts was reported [36, 9, 12] – this means that the most of the components $\hat{s}_m[n]$ contain either brain activity or artifacts, while the mutual intermixing of these signals is noticeably reduced. Under such circumstance, it is possible in the second step to classify (either manually [13, 12], or automatically [12]) each $\hat{s}_m[n]$ as either a brain signal or an artifact, and retain only the brain signals, which will be denoted as $\tilde{\mathbf{s}}[n] = [\tilde{s}_1[n], \ldots, \tilde{s}_M[n]]^T$

$$\tilde{s}_m[n] = \begin{cases} \hat{s}_m[n], & \text{if } \hat{s}_m[n] \text{ is a brain signal,} \\ 0, & \text{if } \hat{s}_m[n] \text{ is an artifact.} \end{cases} \tag{2.2}$$

Finally, brain signal components $\tilde{\mathbf{s}}[n]$ can be projected back to the space of the measured EEG signals to form the record without the artifacts

$$\widetilde{\mathbf{x}}[n] = \mathbf{B}^{-1}\tilde{\mathbf{s}}[n], \tag{2.3}$$

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

---

[1] For the sake of completeness, equation (1.2) from Section 1.4 is repeated here.

2.2 ICA Algorithm

To find the separation matrix $\mathbf{B}$ in the above mentioned algorithm for the muscular artifact removal, it is suggested to use the FastICA [14, Eq. (14, 21, 24)] on data that were pre-whitened by PCA [8]. This conjunction of FastICA with PCA is a popular approach, with fast convergence and previous successes in artifact separation [36, 21]. It is also noteworthy that works [17, 25, 26] reported that the overlearning does not seem to be dependent on the choice of a particular ICA algorithm. Next, the specific form of FastICA algorithm which was used in each proposed algorithm is formally described.

Specifically, the whitening step based on PCA takes the following form

$$\mathbf{z}[n] = \mathbf{\Sigma}^{-\frac{1}{2}} \cdot (\mathbf{U}^{\mathrm{T}} \cdot \mathbf{x}[n]), \tag{2.4}$$

where $\mathbf{z}[n] = [z_1[n], \ldots, z_M[n]]^T$, $n = 1, \ldots, N$ are whitened signals, $\mathbf{\Sigma}$ and $\mathbf{U}$ are matrices of eigenvalues and eigenvectors of the estimated covariance matrix $\frac{1}{N} \sum_{n=1}^{N} \mathbf{x}[n] \cdot \mathbf{x}^{\mathrm{T}}[n]$. For the purpose of defining the utilized form of FastICA algorithm, the matrix $\mathbf{Z} = [z_m[n]]$, $m = 1, \ldots, M$, and $n = 1, \ldots, N$ is denoted. If the dimension reduction to a number M'' of retained whitened signals is applied at this step, then the signals $z_m[n]$, which correspond to the $M - M''$ smallest eigenvalues in matrix $\mathbf{\Sigma}$ are excluded from further processing.

The FastICA algorithm itself starts with a random initialization of an estimate of separation matrix $\mathbf{B}$, where every coefficient $b_{k,l}$ is chosen from the standard normal distribution. When the parallel form of FastICA from [14], where all estimated vectors are being updated at the same iteration, is adopted, the following set of equations is utilized. The processing core of the FastICA consists of the following two steps. In the first step, the separation matrix is updated

$$\mathbf{B}^+ = \mathbf{B} - \mu \cdot \frac{\mathbf{Z} \cdot g(\mathbf{B}^{\mathrm{T}} \cdot \mathbf{Z})^{\mathrm{T}} - \mathbf{B} \circ [\mathbf{m} \cdot \boldsymbol{\lambda}^{\mathrm{T}}]}{\mathbf{m} \cdot [[g'(\mathbf{B}^{\mathrm{T}} \cdot \mathbf{Z})] \cdot \mathbf{n} - \boldsymbol{\lambda}]^{\mathrm{T}}}, \tag{2.5}$$

where the division of matrices is realized element wise, μ is a learning constant that is set to value 0.2, the operator $\circ$ realizes the Hadamard product of two matrices, $\mathbf{n}$ is a unit column vector with N elements, $\mathbf{m}$ is a unit column vector with M elements, the nonlinearity $g(.)$ was chosen as a hyperbolic tangent function [14], the function $g'(.)$ is the derivative of the function $g(.)$, and the matrix $\boldsymbol{\lambda}$ is defined by

$$\boldsymbol{\lambda} = [\mathbf{B}^{\mathrm{T}} \circ [\mathbf{Z} \cdot g(\mathbf{B}^{\mathrm{T}} \cdot \mathbf{Z})^{\mathrm{T}}]^{\mathrm{T}}] \cdot \mathbf{m}. \tag{2.6}$$

In the second step, the symmetric orthonormalization [14] is used to normalize and orthogonalize the updated vectors in the matrix $\mathbf{B}^+$. The utilized scheme of symmetric orthonormalization consists of two parts. First, the unit power normalization is performed

$$\mathbf{B}^+ = \mathbf{B}^+ / \sqrt{||\mathbf{B}^+ \cdot \mathbf{B}^{+\mathrm{T}}||}, \tag{2.7}$$

where $||.||$ is the row matrix norm[2].

Second, 10 iterations[3] of the following recursive formula were performed

$$\mathbf{B}^{+\mathrm{T}} = \frac{3}{2}\mathbf{B}^{+\mathrm{T}} - \frac{1}{2}\mathbf{B}^{+\mathrm{T}}\mathbf{B}^+\mathbf{B}^{+\mathrm{T}}, \tag{2.8}$$

---

[2] It is possible to use any matrix norm except for the Frobenius norm. In this work, row matrix norm was used

[3] It is usually sufficient to perform up to 10 iterations. [14]

where the matrix $\mathbf{B}^+$ consists from the separation vectors, which form its columns.

Steps (2.5), (2.7) and (2.8) are repeated until the convergence of the separation matrix $\mathbf{B}$. The convergence is usually defined as a state of computation when separation vectors, i.e. the columns of matrix $\mathbf{B}$, do not change their direction in M-dimensional space significantly after the algorithm update iteration. A specific form of the utilized stopping criterion in this work is to take an average of diagonal values of a matrix product $\mathbf{B}^{+T} \cdot \mathbf{B}$, which are scalar products being close to one if converged, and stop the computation, when this average value is higher than 0.999.

2.3 Artifact Classification

To decide whether the separated components $\hat{\mathbf{s}}[n]$ are of a brain or a muscular origin, a simple classifier that compares the power of a signal at lower and higher frequencies was used in case of each developed algorithm in Chapter 2. Compared to clean EEG, the EMG artifacts are known to have higher energy at higher frequency bands [34], which allows their easy identification.

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

$$\alpha_i = \min(P_{i,1}/P_{i,2}, P_{i,1}/P_{i,3}). \tag{2.9}$$

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

To determine the optimal value of the threshold T, value α_i was computed for 3000 separated signals $\hat{s}_i[n]$ that were manually classified as either brain signals or muscular artifacts. The computed α_i were then used to plot histograms shown in Fig. 2.1, where a clear separation can be seen between the values of α_i for the individual classes. From Fig. 2.1 one can read that the optimal value of T is about 2.5. Thus, a signal $\hat{s}_i[n]$ was classified as a muscular artifact if $\alpha_i < 2.5$.

2.4 Spatially Projected ICA (SPICA)

The first developed algorithm called spatially projected ICA (SPICA) applies ICA on data subspaces with the reduced dimension according to a first aim of this thesis. Specifically, it is based on a segregation of the M-dimensional signal $\mathbf{x}[n]$ into K subspaces with dimension L

$$\boldsymbol{\mathcal{x}}_k[n] = \mathbf{P}_k \mathbf{x}[n], \quad k = 1, \ldots, K, \tag{2.10}$$

where $\mathbf{P}_k$ are LxM subspace projection matrices (we assume that their rank is L, and their choice will be suggested in Section 2.4.2), and $\boldsymbol{\mathcal{x}}_k[n]$ are L dimensional signals.

Next, the ICA algorithm (its detailed description is provided in Section 2.2) is applied to project each signal subspace $\boldsymbol{\mathcal{x}}_k[n]$ separately into the space of independent components

$$\hat{\boldsymbol{s}}_k[n] = \mathbf{B}_k \boldsymbol{\mathcal{x}}_k[n], \tag{2.11}$$

where $\mathbf{B}_k$ is a separation matrix estimated for $\boldsymbol{\mathcal{x}}_k[n]$ [4]

---

[4] note that this means that the ICA will be computed K times for each $\boldsymbol{\mathcal{x}}_k[n]$ separately.
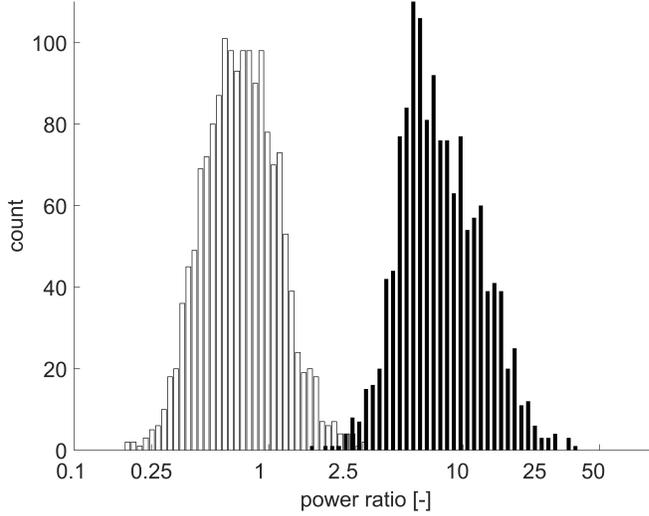
Fig. 2.1:   A histogram of power ratios α for source components classified as EMG (white) and source components classified as brain signals (black).

In next step, estimated sources $\hat{\boldsymbol{s}}_k[n] = [\hat{s}_{k,1}[n], \ldots, \hat{s}_{k,L}[n]]^T$ are classified as either useful brain signals or artifacts, and only brain signals $\tilde{\boldsymbol{s}}_k[n] = [\tilde{s}_{k,1}[n], \ldots, \tilde{s}_{k,L}[n]]^T$ are retained

$$\tilde{s}_{k,i}[n] = \begin{cases} \hat{s}_{k,i}[n] & \text{if } \hat{s}_{k,i}[n] \text{ is a brain signal,} \\ 0 & \text{if } \hat{s}_{k,i}[n] \text{ is an artifact.} \end{cases} \tag{2.12}$$

This operation can be expressed in a matrix form as

$$\tilde{\boldsymbol{s}}_k[n] = \mathbf{Q}_k \hat{\boldsymbol{s}}_k[n], \tag{2.13}$$

where $\mathbf{Q}_k$ is a diagonal matrix with its diagonal being formed by zeros and ones, positioned such that the signals $\hat{s}_{k,i}[n]$ classified as artifacts are eliminated according to (2.12).

Consequently, the reconstruction of signals $\widetilde{\boldsymbol{x}}_k[n]$, in which the muscular artifacts are suppressed, takes place

$$\widetilde{\boldsymbol{x}}_k[n] = \mathbf{B}_k^{-1} \tilde{\boldsymbol{s}}_k[n]. \tag{2.14}$$

The above-mentioned sequence of steps can be expressed as

$$\widetilde{\boldsymbol{x}}_k[n] = \mathbf{B}_k^{-1} \mathbf{Q}_k \mathbf{B}_k \mathbf{P}_k \mathbf{x}[n] = \mathbf{C}_k \mathbf{P}_k \mathbf{x}[n], \tag{2.15}$$

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

Last, all signals $\widetilde{\boldsymbol{x}}_k[n]$ are combined to reconstruct the artifact free EEG. For this purpose, the following symbols are denoted

$$\boldsymbol{\mathcal{P}} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \vdots \\ \mathbf{P}_K \end{bmatrix}, \quad \boldsymbol{\mathcal{C}} = \begin{bmatrix} \mathbf{C}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{C}_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathbf{C}_K \end{bmatrix}, \tag{2.16}$$
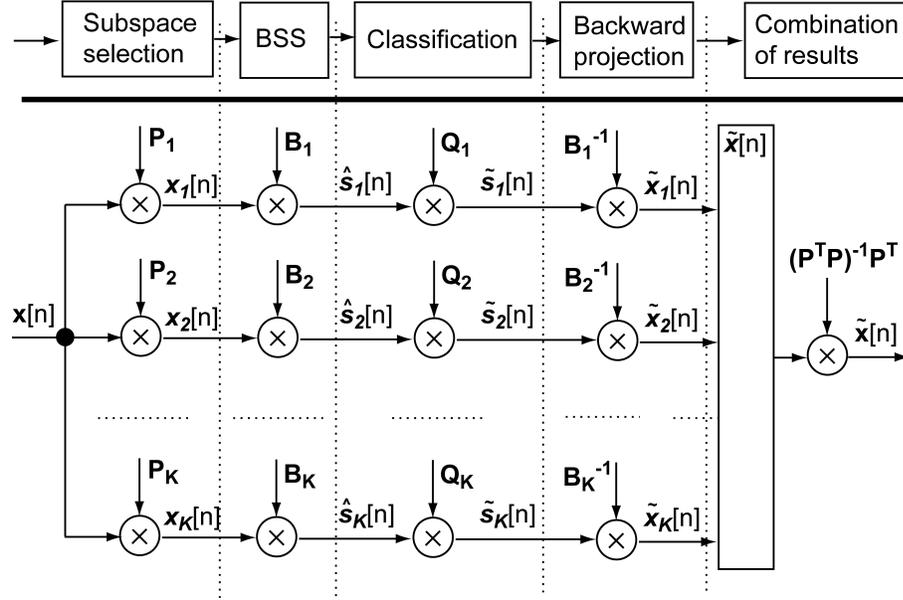
Fig. 2.2: Overall schematic of the SPICA algorithm.

$$\widetilde{\boldsymbol{x}}[n] = \begin{bmatrix} \widetilde{\boldsymbol{x}}_1[n] \\ \widetilde{\boldsymbol{x}}_2[n] \\ \vdots \\ \widetilde{\boldsymbol{x}}_K[n] \end{bmatrix}. \tag{2.17}$$

Using notations (2.16) and (2.17), all the above-mentioned steps can be expressed as

$$\widetilde{\boldsymbol{x}}[n] = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}\mathbf{x}[n]. \tag{2.18}$$

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

$$\widetilde{\mathbf{x}}[n] = \boldsymbol{\mathcal{P}}^\dagger \widetilde{\boldsymbol{x}}[n], \text{ where } \boldsymbol{\mathcal{P}}^\dagger = (\boldsymbol{\mathcal{P}}^T\boldsymbol{\mathcal{P}})^{-1}\boldsymbol{\mathcal{P}}^T. \tag{2.19}$$

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

$$\widetilde{\mathbf{x}}[n] = \boldsymbol{\mathcal{P}}^\dagger \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}\mathbf{x}[n] = \boldsymbol{\mathcal{D}}\mathbf{x}[n], \tag{2.20}$$

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

The overall flow of SPICA algorithm is finally depicted in Fig. 2.2

All further presented algorithms, which present more advanced versions of SPICA algorithm, follow the same structure, which is presented in Fig. 2.2 and equation (2.20). The improvements and modifications are made in blocks labeled as Subspace selection and BSS.

2.4.1 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 Fig. 1.2 disappear, and the ICA behaves as in low dimensional case shown in Fig. 1.1.

If the subspace projection matrices $\mathbf{P}_k$ are chosen so that $\boldsymbol{\mathcal{P}}$ has full rank, and if no artifacts are found in the separated signals $\hat{\boldsymbol{s}}_k[n]$ (i.e. $\mathbf{Q}_k$ are identity matrices), then $\tilde{\boldsymbol{s}}_k[n] = \hat{\boldsymbol{s}}_k[n]$, and $\mathbf{C}_k$ and $\boldsymbol{\mathcal{C}}$ become identity matrices, and $\boldsymbol{\mathcal{D}}$ becomes the identity matrix. Consequently, the processed signals are not distorted. 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 signals $\hat{\boldsymbol{s}}_k[n]$ the rank of $\tilde{\boldsymbol{s}}_k[n]$ and $\boldsymbol{\mathcal{C}}$ will be reduced and consequently even the rank of $\boldsymbol{\mathcal{D}}$ may also be reduced; however, in Chapter 3 it is shown that the newly proposed method causes dimensionality reduction much less severe than the one caused by the PCA based overlearning suppression.

2.4.2   Choice of Subspaces

When choosing the subspaces and therefore designing the projection matrix $\boldsymbol{\mathcal{P}}$, three points should be taken into account.

  i The dimension L should be small enough so that the overlearning is sufficiently suppressed and its effects are not observable. On the other hand, a sufficient number of signals in each subset (L) is required so that the ICA can separate brain signals and artifacts.

 ii The matrix $\boldsymbol{\mathcal{P}}$ should be full rank (otherwise $\boldsymbol{\mathcal{D}}$ would not be identity when no artifacts are being removed, and unnecessary EEG distortion would occur).

iii The orientation of subspaces should be chosen so that the individual ICAs can achieve good separation between brain signals and artifacts in separated signals $\hat{\boldsymbol{s}}_k[n]$ even with the reduced dimension L of $\boldsymbol{x}_k[n]$.

To address point (i), the optimal range of values L is identified in Section 3.1 by applying the algorithm to EEG data, and evaluating for which L we obtain the best reconstruction of brain signals. In Section 2.8 a methodology for the evaluation of the quality of brain signal reconstruction is introduced, and it will be shown at the beginning Section 3.1 of Chapter 3 that even for various EEG electrode systems a fixed range of values of L, where the algorithm always performs well, can be identified. The reader can then use the value of L suggested by the presented evaluation without any additional or repetitive optimization of L.

Point (ii) can be achieved by informed construction of $\boldsymbol{\mathcal{P}}$. For example, by choosing sufficient number of sufficiently diverse subspaces, $\boldsymbol{\mathcal{P}}$ can be full rank.

Point (iii) potentially spans several problems and is not straight forward to address. While the fact, that the separation quality in a particular EEG subspace is not satisfactory, can be attributed to several factors like insufficient amount of information about sources (i.e. too small number of channels), ICA model corruption problem might manifest itself as well when EEG subspace is chosen in a way such that the number of sources which dominate the power of all signals in this subspace is higher than the number of channels in the subspace. Here, two approaches of possible spatial sampling, that is the EEG subspace selection from a given electrode array, are presented.

Sampling strategy based on neighborhood of each electrode

First approach is based on the idea of "zooming" into a particular region of electrode array. Specifically, this idea is based on the following observation. When a high density EEG array (e.g. with 111 electrodes) is used, but the ICA processing is limited to a small group of adjacent electrodes (e.g. 10-15 adjacent electrodes), the ICA tends to provide a nice separation of brain signals and muscular artifacts. This situation was actually already illustrated in Fig. 1.1 in Chapter 1 which shows signals from a subgroup of 9 adjacent electrodes chosen from the 111 electrode array used in Fig. 1.2. The following steps are therefore suggested to choose the matrices $\mathbf{P}_k$.

First, the Euclidean distance between i-th and j-th electrode is denoted as d_{ij}. Next, the following vectors are defined

$$\boldsymbol{\ell}_k = [\ell_{k,1}, \ldots, \ell_{k,L}], \tag{2.21}$$

where $\ell_{k,i}$, $i = 1, \ldots, L$ 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, the following vectors are defined

$$\mathbf{p}_{k,i} = [p_{k,i,1}, \ldots, p_{k,i,M}], \quad p_{k,i,j} = \begin{cases} 1, & \text{if } \ell_{k,i} = j, \\ 0, & \text{otherwise,} \end{cases} \tag{2.22}$$

and the projection matrices are constructed as

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{p}_{k,1} \\ \vdots \\ \mathbf{p}_{k,L} \end{bmatrix}, \quad k = 1, \ldots, M. \tag{2.23}$$

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

Before describing the next sampling approach, few notes about the projection matrices $\mathbf{P}_k$ given by (2.23) can be made.

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 $\boldsymbol{\mathcal{P}}$ are orthogonal as well. Further, because $k = 1, \ldots, M$, each column of $\boldsymbol{\mathcal{P}}$ contains at least a single value of 1, and so has a nonzero norm. Consequently, $\boldsymbol{\mathcal{P}}$ is a full rank matrix, just as we have required in point (ii).

Moreover, this choice of subspaces can be further supported by following observations from the previous literature. 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 [23]. By choosing each subspace as the signals from a group of adjacent electrodes, the attempt of minimizing the number of sources that dominate the signals in a subspace is made, certainly at least as opposed to a subspace that would be composed of signals from nonadjacent electrodes spread all around the subject's head. Consequently, the ICA may provide better separation of brain signals and artifacts as was requested in point (iii).

This choice of $\mathbf{P}_k$ may not be optimal, but it will be illustrated that this strategy provides better results than all state-of-the-art methods mentioned in Section 1.3 available today. Further optimization of $\mathbf{P}_k$ may be possible.

The map of number of subsets, in which each electrode is located (this number is further denoted as I_m for m-th electrode), is depicted in Fig. 2.3 for the subspace
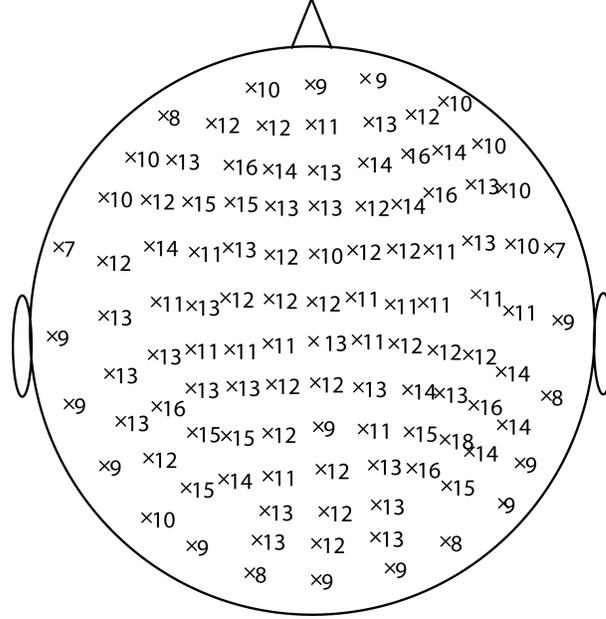
Fig. 2.3:   Overall number of times each signal occurs in any EEG subset for the sampling strategy based on neighborhood of each electrode.

size $L = 12$ electrodes and utilized 111 electrode system. As can be seen in Fig. 2.3, this sampling strategy is introducing quite redundant number of processing attempts, especially for signals from electrodes in the center of the electrode array. The next presented sampling strategy is therefore reducing this redundancy, which makes the overall processing burden smaller.

Spatially uniform sampling strategy

Due to the nature of the spatial distribution of muscular contamination, which tends to have the strongest effect in the fringe areas of the electrode array [37, 13, 12], the central area of the electrode array does not have to be so densely covered by EEG subspaces as in the case of Neighborhood of Each Electrode Sampling Strategy and a resulting algorithm can be possibly made faster without compromising the quality of reconstructed signals. Therefore, the uniform sampling approach is proposed, where the electrode array is uniformly covered by a certain number of subspaces such that each electrode is contained in a user-specified number of subspaces.

The uniform subspace selection stems from the following two steps, which are a definition of subspaces and a redistribution of individual signals from particular electrodes between neighboring subspaces.

Let us first define the term central point of an EEG subspace as the position of an electrode around which the EEG subspace is formed by taking measured EEG signals from L closest electrodes according to their euclidean distances as described in previous Subsection. When referring to a distance between two subspaces or between a subspace and an electrode, one should understand the distance between central points of subspaces or a distance between a subspace central point and a particular electrode.

The first step consists of defining a certain number of subset central points and aligning them in such a way that their uniform spread is achieved. For this, the number

of central points is estimated according to

$$N_{cp} = \left\lfloor \frac{N_{el} \cdot N_{rep}}{L} \right\rfloor, \tag{2.24}$$

where $\lfloor . \rfloor$ denotes rounding operation to the closest smaller integer, N_{cp} is an estimated number of central points, N_{el} is a number of all electrodes in a measuring array, N_{rep} stands for number of subspaces, where each electrode is expected to appear.

Location alignment of N_{cp} central points is done with the use of the basic form of k-means algorithm. The initial positions of central points can be chosen in several ways, for example to form a rectangular grid with such coordinates of individual points, which overlay the actual electrode array. In the case of the electrode array, which was utilized to measure the data for this work, the electrodes were indexed in such a way, that it was possible to choose initial positions of central points as the positions of N_{cp} electrodes from the measuring array with the indexes $1, \left\lfloor \frac{N_{el} - 1}{N_{cp} - 1} \right\rfloor, ..., N_{el}$. Once the central points were found, L electrodes in their vicinity were taken into a respective subspace according to their euclidean distances such as in previous sampling strategy.

During the second step, the following procedure is utilized. If a signal from the particular electrode is used in fewer than N_{rep} subsets, other subsets are searched, starting from the closest ones. In these subsets, algorithm finds signals/electrodes with presence in more than N_{rep} subsets. From these signals, the one coming from the closest electrode is then excluded from the subset and underused signal/electrode is included in this same subset.

Before stating the full form of this spatial sampling strategy, the following symbols are defined. The number of subspaces, which contain the signal from a particular electrode with index l, $l = 1, ..., M$, is further denoted as $N_{use,l}$. This number is denoted for signals in the particular subset as $N_{use,z_{h,j}}$, $h = 1, ..., N_{cp}$, $j = 1, ..., L$, where $z_{h,j}$ is the index of the $j - th$ electrode within the h-th subspace.

This sampling strategy can be summarized in following steps:

1. initialize index $l = 1$ of the current electrode

2. pick l-th electrode and compute number $N_{use,l}$

3. initialize index $k = 1$

4. if $N_{use,l} < N_{rep}$

    (a) find k-th closest subspace (according to its central point), which does not contain signal from electrode l and retrieve indexes $z_{h,j}$ for electrodes in this subspace with index h

    (b) calculate $N_{use,z_{h,1}}, ..., N_{use,z_{h,L}}$ for electrodes in this h-th subspace

    (c) if any $N_{use,z_{h,j}} > N_{rep}$

        i. choose the signal from the closest electrode, where condition $N_{use,z_{h,j}} > N_{rep}$ holds

        ii. exclude signal coming from $z_{h,j}$-th electrode from this subspace

        iii. add signal coming from l-th electrode to this h-th subspace

   iv. update values $N_{use,z_{h,j}}$ and $N_{use,l}$ for both electrodes j and l
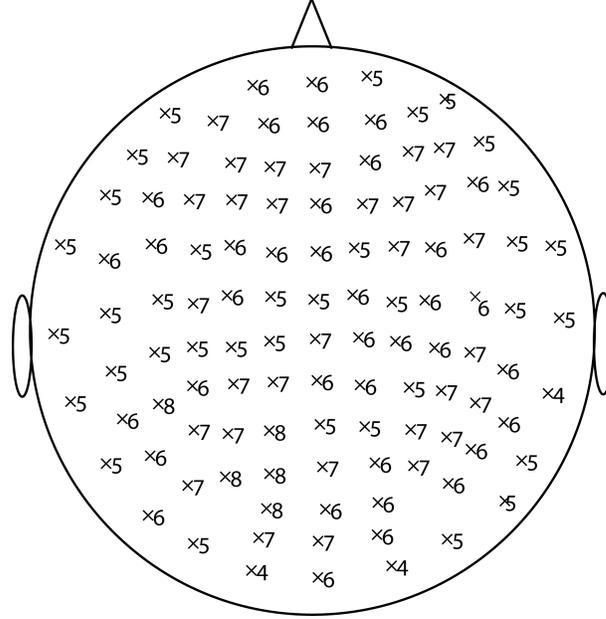
Fig. 2.4:   Overall number of times each signal occurs in EEG subset for sampling strategy based on uniform distribution.

(d) if for all j holds $N_{use,z_{h,j}} < N_{rep}$ and $k < 3$

      i. set $k = k + 1$ and return to step (a)

(e) else, continue with step 5.

5. set $l = l + 1$ and repeat steps 2 to 5

The map of number of subsets for uniform sampling, in which each electrode is located, is depicted in Fig. 2.4 for the subset size $L = 12$ electrodes and utilized 111 electrode system and number $N_{rep} = 8$.

As is clearly visible in Fig. 2.4 there is still some variation in the number I_m across the scalp, but it is smaller compared to the sampling strategy based on neighborhood of each electrode. Standard deviation of I_m computed across the scalp is in the case of sampling strategy based on the neighborhood of each electrode equal to 2.2051 and 0.9482 in the case of spatially uniform sampling strategy with parameter $N_r ep = 8$.

To verify, whether this spatially uniform sampling strategy adheres to the points (i), (ii), and (iii) made at the beginning of this Section, following notes can be made. Since the EEG subspaces chosen by this strategy retain their size and generally follow the principle of grouping the signals from nearby electrodes together, the properties of subspaces will be similar as the ones for sampling strategy based on nearest neighborhood. With respect to the point (i), the reader is again referred to Section 3.1, where the fixed range of values L, for which the algorithm SPICA always performs well for this sampling strategy, is identified. Moreover, the performance of SPICA algorithm with both EEG subspace sampling strategies is compared and it is shown that no statistically significant difference was found.

To address the point (ii), it should be noted that columns of matrix $\mathbf{P}_k$ and consequently $\mathbf{P}$ are still orthogonal due to the fact that each row of $\mathbf{P}_k$ has still only a single value of 1 and zeros otherwise just as in the case of previous sampling strategy.

As can be seen in Fig. 2.4 signal from each electrode was always utilized more than once and therefore every column has a nonzero norm resulting in a full rank matrix $\mathbf{P}$. When using different values for parameter N_{rep}, one should always check whether the sum of each column of matrix P (as shown in Fig. 2.4) is a non zero value. In the final evaluation, the resulting performance of the algorithms will be shown for parameter N_{rep} being 8, and 4, which still provided nonzero values when summing the columns of resulting matrix $\mathbf{P}$.

With respect to the point (iii), the same note can be made as in the case of the previous sampling strategy. By choosing the subspaces as signals from adjacent electrodes, the number of sources which dominate the power in the local scalp area is attempted to be minimized, which makes it easier for the ICA to estimate remaining source components. Furthermore, all developed algorithms were tested, their performance evaluated for both sampling strategies and in Section 3.1, it is shown that the performance do not decrease significantly even in the case of coarser sampling with parameter N_{rep} being as low as 4.

2.5   SPICA with Desirable Local Subset Selection (SPICA-LSS)

This section builds on the algorithm SPICA presented in Section 2.4 and introduces a subspace selection mechanism, which pinpoints those ICA estimated source subspaces, where the achieved separation quality was not sufficient, and which should therefore be rejected from clean EEG reconstruction.

This rejection of subspaces can be well tolerated, since the approaches of matrix $\mathbf{P}_k$ (2.23) assembly described in Section 2.4.2 introduce considerable level of redundancy with respect to the number of subspaces $\boldsymbol{\varkappa}_k[n]$ in which each measured signal is included (this still holds even for the uniform subspace selection, where the number of processed signal subspaces was decreased). Due to this redundancy, it is further possible to identify and reject subspaces, where the achieved separation was not successful, without loss of the full rank of matrix $\boldsymbol{\mathcal{P}}$.

To identify low separation quality in a particular signal subspace, we can use characteristics like higher order moments or power spectral densities of individual source signals.

Specifically, if the separation was successful and the source intermixing was minimized, the resulting source signals corresponding EEG and EMG will yield different values of the above mentioned properties. On the other hand, in the case of a less successful separation, with EEG and EMG intermixing in each estimated source, the resulting source estimates will have similar values of the above mentioned properties. Therefore, it can be noted that if the separation was successful, then the variance of a chosen property across all estimated sources will be higher than in the case of less successful separation.

In this work, it is suggested that quality of separation in a subset is classified as unsatisfactory if the variance of the chosen signal characteristic is less than a certain threshold. Now, the specific form of utilized signal property and a threshold is provided.

Namely, if the given task is to separate EEG and EMG sources, it is quite intuitive to take a look at the spectra of these signals as is shown in Fig. 2.5, where it can be seen that the Gabor bandwidth, which is known as the square root of ratio of the second and zero spectral moment, of the EMG signals tend to be much larger than the one of the EEG signals, which have a much narrower spectrum. These spectral properties
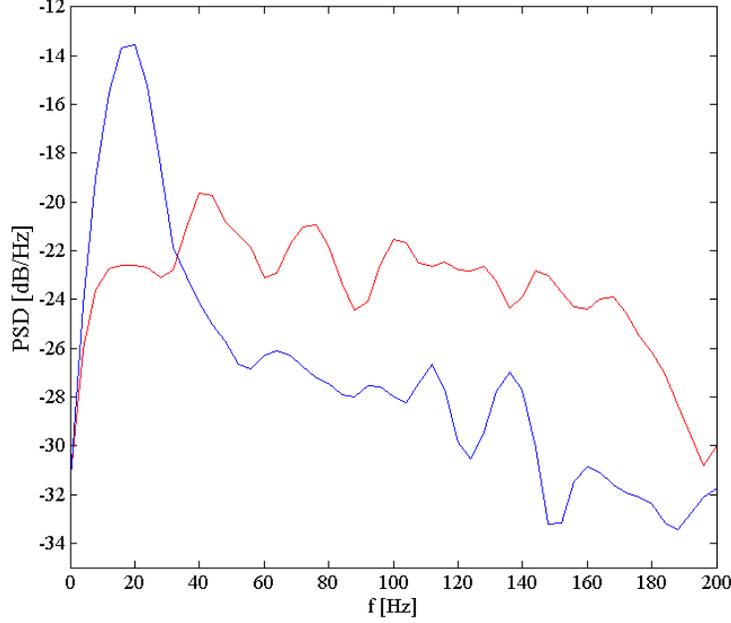
Fig. 2.5: Example of power spectral density for EEG (blue) and EMG (red) signal content.

allow EEG and EMG signals to be distinguished even when their total power is equal. The proposed signal property is therefore chosen as power in a high frequency band (the cut-off frequency of this band will be specified later in this section) for each signal $\hat{s}_{k,1}[n], \ldots, \hat{s}_{k,L}[n]$ in a subspace $\hat{\boldsymbol{s}}_k[n]$. The variance across subspace source estimates $\hat{s}_{k,i}[n]$, where $i = 1, \ldots, L$, is then used as a quantity indicating separation performance. Since the chosen signal property should be very small for clean EEG source estimates and a positive number for EMG source estimates, it is expected that the variance of this property will be larger in case of successful separation results when compared to the results from less successful separation attempts.

To compute the power values in a high frequency band for all signals from subspace $\hat{\boldsymbol{s}}_k[n]$, the high-pass filtering is firstly applied according to

$$\hat{f}_{k,i}[n] = h[n] * \hat{s}_{k,i}[n], \tag{2.25}$$

where $\hat{f}_{k,i}[n], i = 1, \ldots, L, \quad n = 1, \ldots, N$ are filtered estimates of source signals $\hat{s}_{k,i}[n], i = 1, \ldots, L, \quad n = 1, \ldots, N$ and $h[n]$ is an impulse response of the utilized finite impulse response (FIR) filter. This filter is required to retain EMG contents at higher frequencies and minimize EEG energy in the filtered signals, which leads to the use of Wiener filter of the following form

$$H(\Omega) = \frac{\bar{S}_{emg}(\Omega)}{\bar{S}_{emg}(\Omega) + \bar{S}_{eeg}(\Omega)}, \tag{2.26}$$

where $H(\Omega)$ is the frequency characteristic of the filter, $\bar{S}_{emg}(\Omega)$ denotes averaged power spectral density (PSD) of the clean measured EMG signals and $\bar{S}_{eeg}(\Omega)$ denotes averaged PSD of the measured clean EEG signals.

The average power spectral densities $\bar{S}_{eeg}(\Omega)$, $\bar{S}_{emg}(\Omega)$ were estimated for the set of EEG signals without visible muscular artifacts and the set of EMG signals measured
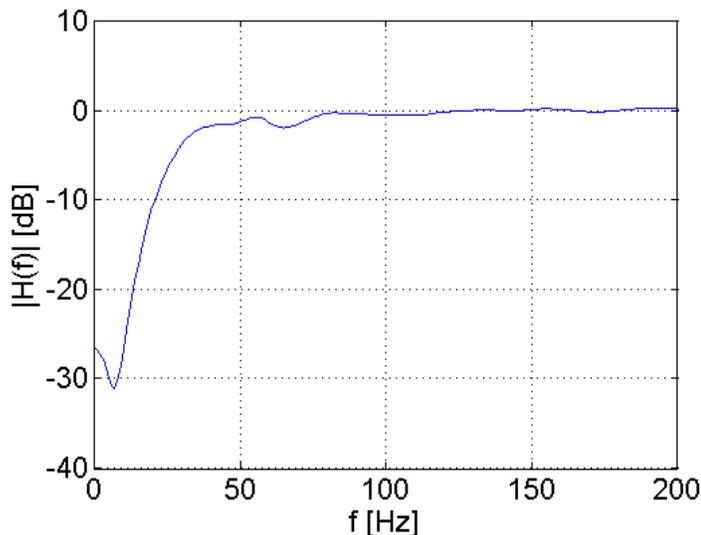
Fig. 2.6: The amplitude frequency characteristic of the utilized high-pass filter.

around cervical muscles (therefore practically without the presence of EEG). Both signal sets consisted of 20 2-second-long signals with 1024 Hz sampling frequency. Power spectral densities were estimated by means of Welch method with the utilized Hamming window of length 512 samples and 75% of segment overlap. The estimated amplitude frequency characteristic of the utilized filter is shown in Fig. 2.6.

The power in the higher frequency band is then computed for each filtered signal $\hat{f}_{k,i}[n]$ as

$$\sigma_{k,i}^2 = \frac{1}{N} \sum_{n=1}^{N} |\hat{f}_{k,i}[n]|^2, \quad i = 1, \ldots, L. \tag{2.27}$$

The variance Σ_k^2 of all power values from filtered signals in k-th subspace $\hat{f}_{k,i}[n], \quad i = 1, \ldots, L$ will be further referenced as a feature variance and is defined as

$$\Sigma_k^2 = \frac{1}{L} \sum_{i=1}^{L} \left(\sigma_{k,i}^2 - \frac{1}{L} \sum_{j=1}^{L} \sigma_{k,j}^2 \right)^2. \tag{2.28}$$

The overall flow of algorithm SPICA-LSS is similar to that of SPICA algorithm shown in Fig. 2.2. The only difference is the modified inner structure of block BSS. The computation of feature variance values and the decision processes, which are based on them, are depicted on Fig. 2.7.

In a principle, the ICA algorithm is run several times with different random initializations of estimated separation matrix $\mathbf{B}_k$ if the separation performance is recognized as unsatisfactory. To recognize unsatisfactory performance, the feature variance $\Sigma_{k,1}^2$ is computed for a set of de-correlated signals from PCA in block VE (Variance Estimation) and is compared with the feature variance $\Sigma_{k,2}^2$ representing the set of estimated independent components from ICA algorithm. The ICA is re-run with different initial separation matrix if $\Sigma_{k,2}^2$ is smaller than the scaled feature variance $c \cdot \Sigma_{k,1}^2$ from PCA. The scaling constant c was chosen as a value equal to 10, which sets the requirement that
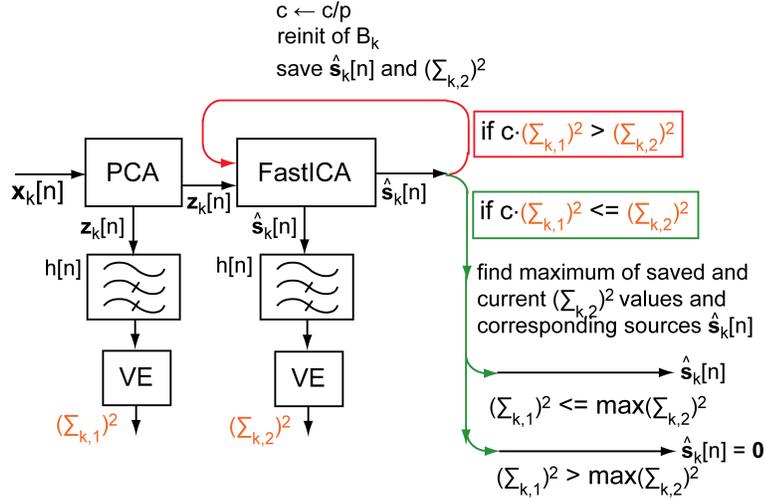
Fig. 2.7: Mechanism for checking the quality of separation for each subset.

the feature variance of estimated source signals is higher than 10 times of the feature variance of principal components to accept the separation result.

If this condition is not met, then the old subspace separation matrix is saved and a new one is randomly initialized for another ICA estimation attempt. The scaling constant c is updated by its division by a contraction factor p, which is chosen as having a value of $\sqrt{3}$. If the condition is met, then the estimated source components or the ones saved from previous ICA computation attempts, which correspond to the highest feature variance $\Sigma_{k,2}^2$, are taken for reconstruction.

The estimated source signals in each subset are discarded, if the highest achieved $\Sigma_{k,2}^2$ is smaller than the value $\Sigma_{k,1}^2$, which suggests that the separation of EEG and EMG contents was less successful than in the case of basic PCA for the case of present muscular artifact. In the case that the estimated source set is about to be discarded, it is required to check whether there is still at least one reconstructed signal corresponding to each channel in $\mathbf{x}[n]$. This can be done by defining auxiliary vector of number of processing times for each signal channel

$$\mathbf{p}_{check} = \mathbf{1} \cdot \mathbf{P}, \tag{2.29}$$

where $\mathbf{p}_{check}$ is a row vector of M non zero values of processing times for respective electrodes, $\mathbf{1}$ is a unit row vector of $L \cdot M$ values. Before any subspace is about to be discarded, values in $\mathbf{p}_{check}$ at indexes corresponding to the signals of that specific subspace are lowered by 1. If any value in $\mathbf{p}_{check}$ reaches zero, then the corresponding subspace cannot be discarded and is taken for further processing, e.g. classification and reconstruction.

The final decision rule stated in the previous paragraph also means, that for the chosen values of c and p, it makes no sense to attempt ICA estimation for more then 4 times (at the fifth time, the iteratively updated value of c, which stands for the achieved ratio between $\Sigma_{k,2}^2$ and $\Sigma_{k,1}^2$, is already smaller then 1).

2.6 Preprocessing Tool for ICA Based on Spectral Difference Based Separation (SDBS)

The goal of this technique is to pre-process the EEG signals for the ICA algorithm in such a way, which will consequently make it easier for ICA to separate the EMG artifacts from EEG signal contents. In order to do this, it is still possible to utilize the traditional PCA approach, but instead of projecting EEG signals in orthogonal directions of the highest power, an attempt is made to project the signals in the directions of EMG artifacts themselves.

This task can be accomplished by looking at this problem as an eigenvalue decomposition of a carefully chosen estimate of a covariance matrix of interest. Specifically, if the artifactual contents of EEG signals are amplified and an estimate of the covariance matrix is assembled from such modified signals, then the eigenvectors of this new estimated covariance matrix will point in directions of EMG artifacts more likely than the eigenvectors of a regular covariance matrix estimate used in standard PCA as defined in (2.4). The EMG artifacts can be amplified by a high-pass filtering of the measured signals; namely, the filter utilized in the previous section can be reused here again.

The technique stems from the following steps. Firstly, the covariance matrix of filtered signals is estimated by

$$\mathbf{R}_1 = \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}_f[n] \cdot \mathbf{x}_f[n]^{\mathrm{T}}, \tag{2.30}$$

where $\mathbf{x}_f[n]$ are EEG signals $\mathbf{x}[n], n = 1, ..., N$ that were filtered by the filter defined in Section 2.5. Since the algorithm will operate on data in a block-wise manner, a zero-phase filtering can be applied.

Second, matrices $\mathbf{U}_1$ and $\mathbf{\Sigma}_1$ of eigenvectors and eigenvalues obtained from an eigenvalue decomposition of matrix $\mathbf{R}_1$ are used as in the case of regular pre-whitening procedure of principal component analysis. This time, EEG signals are projected in a direction of their decorrelated and unit-power normalized EMG contents, which were amplified by the utilized filter prior to matrix R_1 assembly

$$\hat{\mathbf{s}}[n] = \mathbf{\Sigma}_1^{-\frac{1}{2}} \cdot \left(\mathbf{U}_1^{\mathrm{T}} \cdot \mathbf{x}[n] \right), \tag{2.31}$$

where $\hat{\mathbf{s}}[n], n = 1, \ldots, N$ are projected and normalized EEG signals. The difference of this step with respect to the standard PCA lies in the fact that the EEG signals are now projected to a mutually orthogonal normalized subspaces with a sorted power located in the frequency band of interest (according to the utilized filter).

The structure of the SDBS algorithm is depicted in Fig. 2.8.

Theoretically, SDBS algorithm could be used as a stand-alone tool for the muscular artifacts suppression as the author has already described in [32]. SDBS algorithm is in this scenario followed by a classifier of separated components and the backward reconstruction of EEG components as described in Chapter 1, equations (1.2), (2.2), and (2.3). However, it is shown in Chapter 3, that the main utility and strength of this algorithm lies in pre-processing of the data for subsequent ICA algorithm, for which it was originally derived.
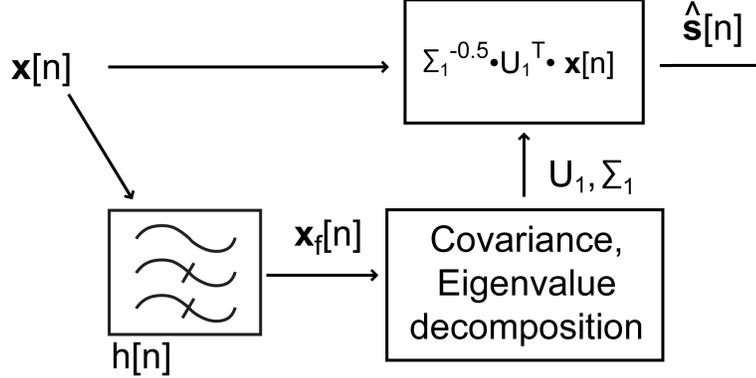
Fig. 2.8: SDBS algorithm structure.

2.7   Combined SPICA/SPICA-LSS and SDBS

In this Section, possible scenarios for a cooperation between SPICA/SPICA-LSS and SDBS algorithms as well as a rationale for such cooperation are provided. The first and the second further presented cooperation scenarios deal with the modified structure of BSS block within SPICA scheme, which is shown in Fig. 2.2. The third scenario builds on SPICA-LSS algorithm scheme which is enhanced by concurrent SDBS processing.

The SPICA-LSS algorithm, which was introduced in Section 2.5, presented two new key features when compared to basic SPICA. Namely, the technique for ICA separation quality assessment and the fact that the ICA is performed several times with random starting points if the separation quality does not meet the user-defined requirement. Since this strategy of a random sampling of the ICA objective space does not have to find the optimal separation results given the finite number of random starting points for ICA, this can lead to an idea of whether the ICA estimation cannot be directly guided in a preferable direction such that both objectives, that is the source mutual independence (ICA objective) as well as the feature variance defined by (2.28) (objective of technique for separation quality assessment or its feasible approximation), are maximized.

Algorithm SDBS introduced in Section 2.6 represents a possible tool for the correction and guidance of the ICA estimations, since it attempts to project measured signals in directions of pronounced power at higher frequencies and therefore interesting directions of possible EMG artifactual contents. The following three cooperation scenarios were developed and their behavior studied. Each of the following scenarios are representing modifications of the algorithms SPICA or SPICA-LSS, the schemes of which are shown in Fig. 2.2 and Fig. 2.7.

The first scenario, which is further denoted as an iterative SPICA-SDBS (IT-SPICA-SDBS), is depicted in Fig. 2.9. IT-SPICA-SDBS builds on the basic SPICA algorithm and modifies only its block BSS according to Fig. 2.9. It is based on the serial processing BSS chain, where in each computation step several FastICA iterations are followed by SDBS step. This means that the specific number of iterations is dedicated to FastICA algorithm and then the resulting source estimates $\hat{s}_k[n]$ are treated as an input for the SDBS algorithm according to

$$\hat{\boldsymbol{s}}_k[n]^+ = \boldsymbol{\Sigma}_1^{-\frac{1}{2}} \cdot \left(\mathbf{U}_1^{\mathrm{T}} \cdot \hat{\boldsymbol{s}}_k[n] \right), \qquad (2.32)$$
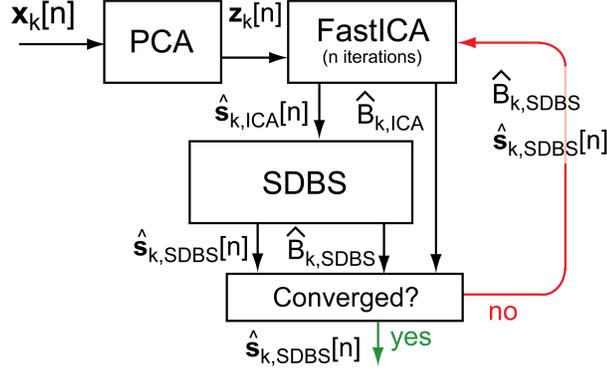
Fig. 2.9: Iterative sequential source estimation by FastICA and SDBS algorithms.

where $\hat{\boldsymbol{s}}_k[n]^+$ denotes new estimates of subset sources, and matrices $\boldsymbol{\Sigma}_1$, $\mathbf{U}_1$ contain eigenvalues and eigenvectors of the following estimate of covariance

$$\frac{1}{N} \sum_{n=1}^{N} \left(\hat{\boldsymbol{\mathcal{F}}}_k[n] \cdot \hat{\boldsymbol{\mathcal{F}}}_k[n]^{\mathrm{T}} \right), \tag{2.33}$$

where $\hat{\boldsymbol{\mathcal{F}}}_k[n] = [\hat{f}_{k,1}[n], \ldots, \hat{f}_{k,L}[n]]$ are filtered signals $\hat{\boldsymbol{s}}_k[n]$ according to

$$\hat{f}_{k,i}[n] = h[n] * \hat{s}_{k,i}[n], \quad i = 1, \ldots, L \tag{2.34}$$

where $h[n]$ is the impulse response of the filter defined in Section 2.5 by equation (2.26).

All in all, estimated sources from the FastICA are used as signals for a covariance matrix estimation in the SDBS algorithm and estimated source signals and a separation matrix from SDBS are in turn used as the initialization information for FastICA. This continuous switching between ICA iterations and SDBS processing for separation vector estimation is maintained until convergence, which is defined as a small change of the direction between separation vectors from the last and current iteration as mentioned in Section 2.2.

The second scenario, which is further denoted as an SDBS initialized SPICA, is depicted in Fig. 2.10. The algorithm SDBS is used as a starting point for ICA. Specifically, the SDBS provides decorrelated signals with maximized SDBS objective function to direct the ICA computations into the region, where a maximum of the feature variance for the given set of separated signals occurs. SDBS initialized SPICA also builds on the basic SPICA algorithm and modifies only its block BSS as is depicted in Fig. 2.10. The estimated separation matrix in each EEG subspace has therefore a following form

$$\mathbf{B}_k = \mathbf{B}_{k,ICA} \cdot \mathbf{B}_{k,SDBS}, \tag{2.35}$$
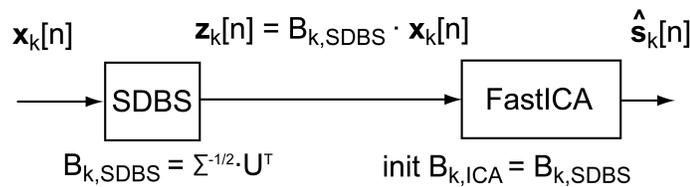

Fig. 2.10: SDBS algorithm used for initialization of the separation matrix $\mathbf{B}_k$ for each EEG subspace processing by FastICA.
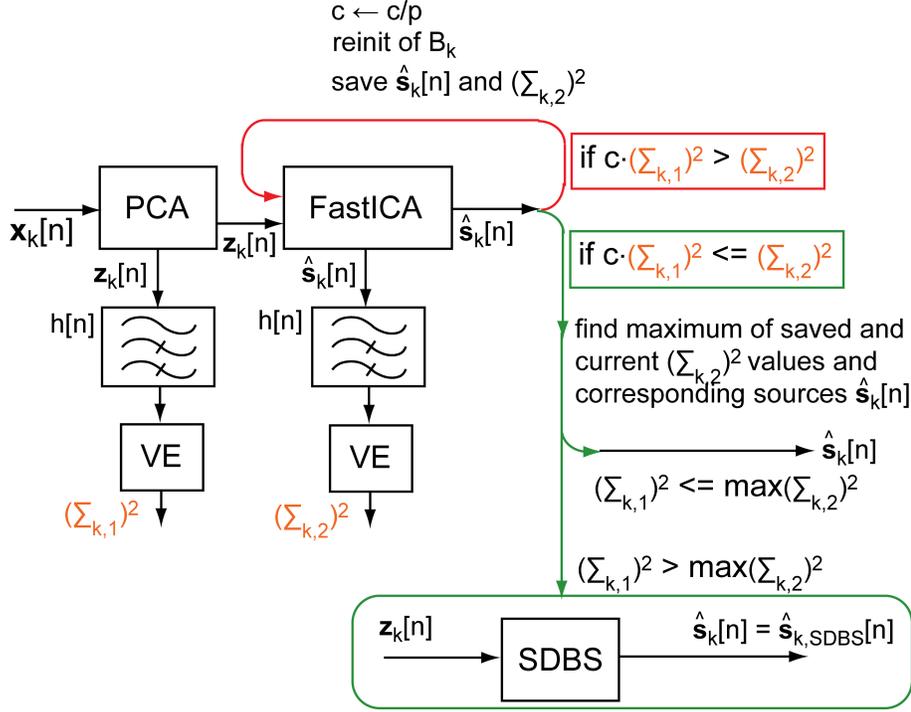
Fig. 2.11: SDBS algorithm used for separation when the ICA provides low quality estimates of sources.

where $\mathbf{B}_{k,SDBS} = \Sigma^{-0.5} \cdot (U^{\mathrm{T}})$ is the separation matrix estimated by SDBS algorithm, Σ and U are the eigenvalues and eigenvectors estimated by an eigenvalue decomposition of an estimate of covariance matrix from (2.30), and $\mathbf{B}_{k,ICA}$ is the separation matrix estimated by FastICA algorithm without PCA preprocessing as defined in Section 2.2.

The third scenario, which is further denoted as SPICA-LSS with additive SDBS (SPICA-LSS-SDBS), is shown in Fig. 2.11, the SDBS algorithm is utilized when the FastICA is not able to estimate a set of source estimates with the feature variance which is higher than the feature variance of the set of principal components provided by PCA. There could be 2 reasons for this, either a low separation of EEG and EMG contents was achieved or there was very weak muscular content in the data, for example even clean EEG itself. In both cases, the set of source estimates from SDBS algorithm is utilized for the consequent classification and reconstruction, which do not harm the EEG contents as shown in Chapter 3 and provide additional suppression of interference if muscular artifacts are present.

The final scenario is further denoted as SDBS initialized SPICA with SDBS at 2nd step. This technique uses SDBS initialized SPICA algorithm to process the contaminated EEG signals and suppress strong muscular artifacts in the first step. In the second step, the standalone SDBS algorithm followed by classification and reconstruction steps as described at the end of Section 2.6 is utilized to process the EEG channels $\widetilde{\mathbf{x}}[n]$ with already suppressed artifacts from the first step to search and suppress any remaining interference in resulting EEG channels $\widetilde{\mathbf{x}}^+[n]$.

The structure of this algorithm is shown in Fig. 2.12.

1st step:

$\mathbf{x}[n]$

channel 1

$\vdots$

channel M

SDBS
initialized
SPICA

$\tilde{\mathbf{x}}[n]$

channel 1

$\vdots$

channel M

2nd step:

$\tilde{\mathbf{x}}[n]$

channel 1

$\vdots$

channel M

SDBS $\rightarrow$ Classifier $\rightarrow$ Backward projection

$\tilde{\mathbf{x}}^{+}[n]$

channel 1

$\vdots$

channel M
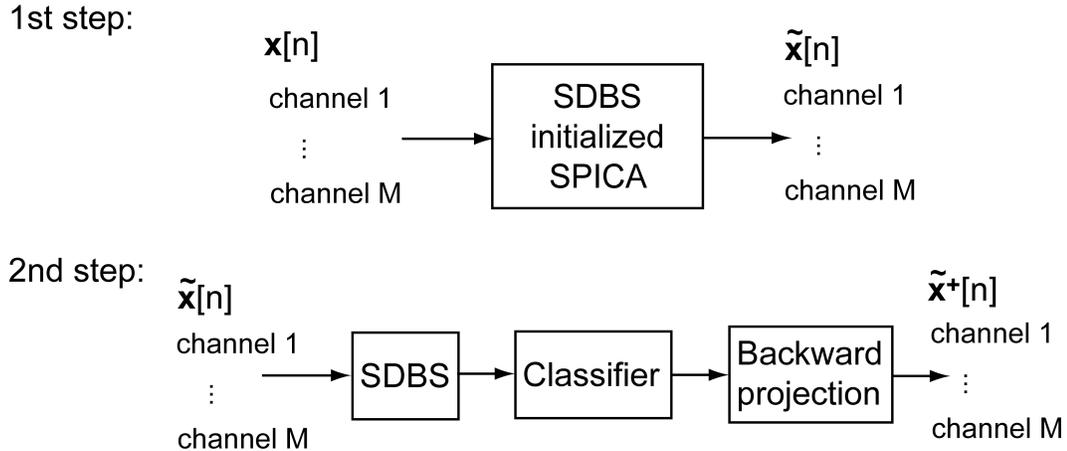
Fig. 2.12: Structure of SDBS initialized SPICA algorithm with SDBS post-processing.

2.8 Evaluation and Comparison of Each Algorithm Performance

To evaluate the properties of the suggested method, one 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 there is no way how to find out what exactly should the EEG look like, when the artifacts are removed. Therefore, proposed methods cannot be objectively evaluated on real world EEG signals corrupted by EMG artifacts.

To be able to assess the EMG suppression as well as the EEG contents preservation, while keeping the evaluation as close to the real scenario as possible, a performance testing procedure that uses real world EEG and EMG records but simulates their mixing using a realistic head model has been devised.

Specifically, the following approach was used. First, artifact free EEG signals, henceforth denoted as $\mathbf{x}_{clean}[n]$, were obtained. 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 a 10-minute long record with their eyes open and 2-minute long records with their eyes closed were obtained. 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 a 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, the distribution of EMG artifacts, which were individually measured as one channel record, across the head surface was simulated. The 111 channels of muscular artifacts obtained by a simulation on a boundary element method (BEM) based realistic head model with the same electrode
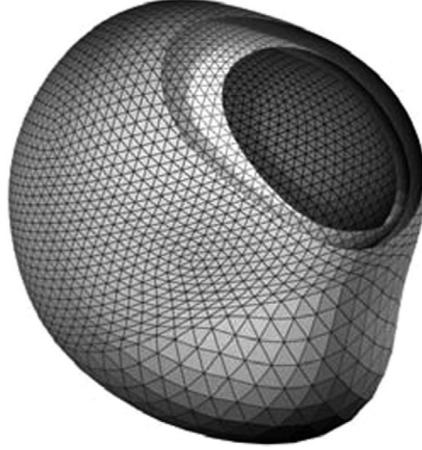
Fig. 2.13: BEM model of skin, skull and brain boundaries used for simulation of EMG surface potential distribution. This figure was published in [4]. Missing surfaces in the boundary layers serve merely for the illustration purposes and were not present in the resulting mesh for simulations.

arrangement as the one used for the EEG measurement were assembled. BEM model composed of 19440 elements arranged in 3 layers representing air-skin, skin-skull and skull-brain boundaries. This model, which is shown in Fig. 2.13 was previously used in [3, 4]. In this simulation, 5 sources representing cervical muscles and 6 sources representing mandibles were placed 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, EMG artifacts that could be realistically observed in EEG electrodes were created. 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

$$\mathbf{x}[n] = \mathbf{x}_{clean}[n] + \sqrt{\eta} \cdot \mathbf{x}_{artifact}[n]. \tag{2.36}$$

The constant η was chosen as

$$\eta = \frac{\sum_{n=1}^{N} \|\mathbf{x}_{clean}[n]\|^2}{\sum_{n=1}^{N} \|\mathbf{x}_{artifact}[n]\|^2} \xi, \tag{2.37}$$

where $\|.\|$ 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 constants, $\xi = 0$, $\xi = 1$ 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) were chosen.

Illustrations of signals $\mathbf{x}[n]$ obtained through this approach for $\xi = 1$ are in Fig. 2.14 and Fig. 2.15. Fig. 2.14 shows signals in temporal domain, and presents topographic maps of average power spectral density (PSD) in various frequency bands. Fig. 2.15 shows the PSDs of EEG with simulated EMG contamination at five different scalp locations. PSDs in Fig. 2.15 were estimated by averaging periodograms computed from 700 segments, which were 1s long and were weighted by the Hamming window.

A)

B)
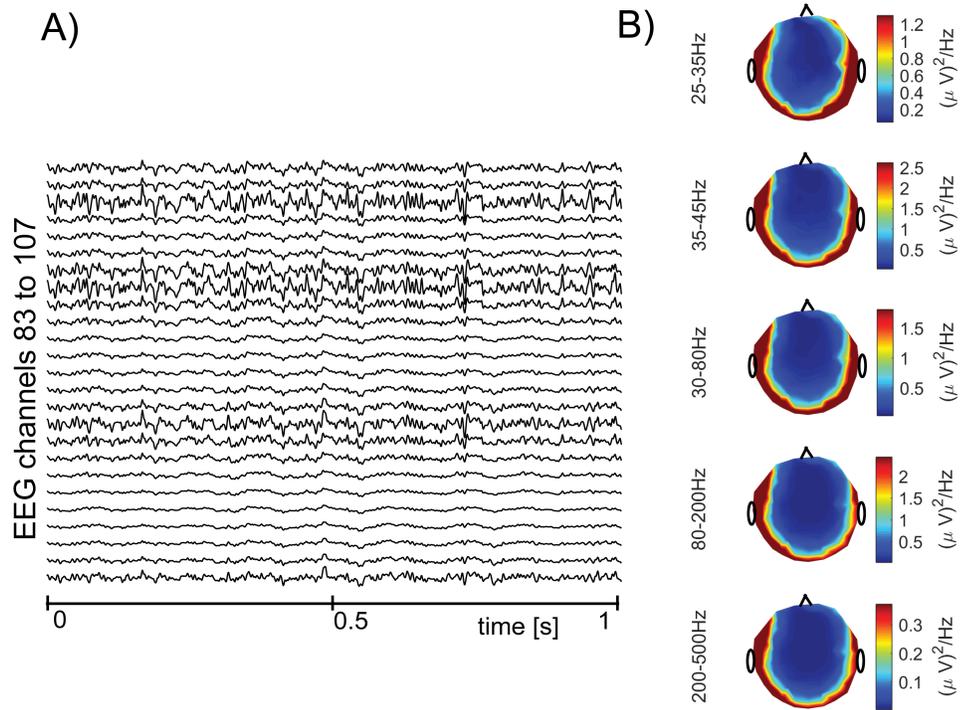
Fig. 2.14: An example of EEG with the simulated EMG contamination. (A) 25 EEG channels with simulated EMG contamination. (B) Topographic maps of average power spectral density in various frequency bands.
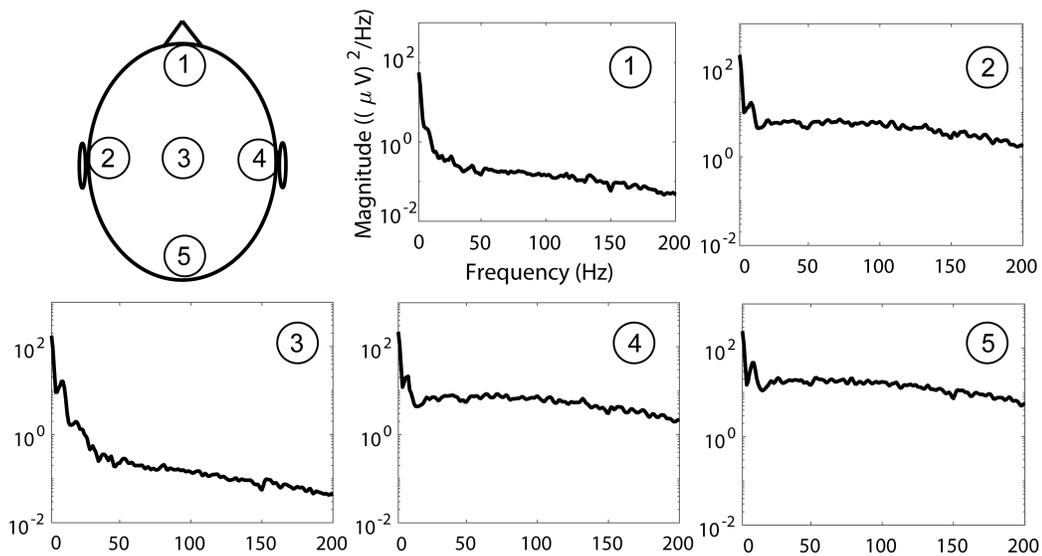

Fig. 2.15: Examples of PSDs of EEG with simulated EMG contamination at several scalp locations.

Once the reconstructed signals $\widetilde{\mathbf{x}}[n]$ were obtained, an average correlation coefficients for each channel pair between original and processed datasets were computed. These 111 values (one for each channel pair) were consequently averaged to obtain one value, which serves as a measure of EEG reconstruction quality for all channels of particular EEG record

$$r = \frac{1}{M} \sum_{m=1}^{M} \frac{\sum_{n=1}^{N} \widetilde{x}_m[n] x_{clean,m}[n]}{\sqrt{\sum_{n=1}^{N} \widetilde{x}_m^2[n] \sum_{n=1}^{N} x_{clean,m}^2[n]}}. \qquad (2.38)$$

From all subjects, 35000 samples of signals $\mathbf{x}[n]$ were gathered, and all above-mentioned methods were then applied to each of these trials. The performance measure values r were obtained for each trial and one mean value $\bar{r}$ and standard deviation σ were estimated for each tested method and strength of muscular interference as defined by (2.36). These values were then compared for all methods with the one way ANOVA and the post-hoc Scheffe test.

The evaluation described above was used for two purposes. First, the performance of SPICA algorithm used with 3 spatial sampling strategies (Nearest neighborhood and Uniform sampling with parameter $N_{rep} = 8$ from Subsection 2.4.2) for various values of L, three different electrode systems (111 electrode system, 10-10 electrode system and 10-20 electrode system) and 1,2 and 4 second long data segments was examined. The obtained results were used to identify the values of L, for which SPICA algorithm and other algorithms, which build upon it, provide the best results. Second, the performance of all the above mentioned methods (the newly proposed methods as well as the state-of-the-art methods) was examined and compared by means of finding statistically significant differences between the sets of r values for particular methods. The computed values of $\bar{r}$ for each method also served for evaluation of how well EMG artifacts were suppressed and original EEG retained.

The evaluation was carried out on 1s long EEG segments. If there is a need to process a longer record in practice, then particular record can be segmented with or without overlap between neighborhood segments with 1s long windows and resulting fractions processed separately by suggested methods. Selected segment length seems to be sufficient for the artifact removal and at the same time short enough to capture the nonstationarity of EMG artifacts.

Last, to test the ability of the algorithms to retain useful EEG information on real EEG data, the approach from [12] was followed. Specifically, the effect of processing on detection of changes in alpha activity in EEG records when measured subject closed or opened eyes was evaluated. These changes in alpha activity, when comparing EEG epochs from segments where measured subjects had their eyes opened to epochs from segments where subjects had closed eyes, will be further denoted as the alpha band power augmentation. The testing EEG dataset comprised of two-minute-long EEG records measured with subjects' eyes closed, and two-minute-long EEG records measured with subjects' eyes open. Based on visual inspection of signal waveforms, only records from 13 out of 20 participants were selected for this evaluation, since the rest of records did not manifest distinct augmentation of alpha activity when the subjects eyes were closed. The measured signals were processed by the proposed methods using segmentation into 1s long segments without overlap. To be able to study the impact of the proposed methods, PSDs before and after the processing by a specific algorithm were computed. The signals for PSD computation originated for each subject from an electrode that is corresponding to the electrode OZ of the 10-10 electrode system.

The PSDs were estimated based on Welch's method with 1s long segmentation window, segment weighting by the Hamming window and 75% overlap. Finally, the grand average of estimated PSDs over all participants was acquired and that is presented in each algorithm processing illustration section of Chapter 3.

3. RESULTS

This chapter is organized as follows. The first section reports the findings on an optimal EEG subspace sizes for efficient computations of SPICA, SPICA-LSS algorithms and their derivations. This search for optimal EEG subspace size is carried out for SPICA algorithm and two mentioned strategies of EEG subspace selection, namely nearest neighborhood and uniform sampling.

Next, comparison of each algorithm performance based on the correlation coefficients between original EEG traces and their reconstructed versions as described in Section 2.8 of Chapter 2 is provided in Section 3.2.

In further sections, each of the algorithms is applied on real EEG data with muscular artifacts. The effect of each algorithm on signal waveforms, PSDs and topographic maps of power in several spectral bands is examined. This examination consists of two steps.

First, 1s long EEG segment contaminated by muscular artifacts is processed with a selected algorithm and then the time-domain waveform of the resulting EEG signals is visually inspected for the presence of any remaining muscular activity content. Also, power spectral densities (PSD) of processed signals at several locations across the scalp and their topographic maps are compared against PSD of EEG signals with artifacts and their topographic maps.

Second, the proposed algorithms are applied to the data manifesting alpha augmentation phenomenon and their effect on the detection of power increase in alpha activity spectral band between open and closed eyes periods [12] is evaluated. It is especially verified that any of the proposed techniques as described in Section 2.8 of Chapter 2 does not impair the detection of this phenomenon.

In case of SDBS algorithm, its function is firstly tested on a synthetic EEG/EMG mixture dataset, where a mixing matrix is known. In case of SPICA/SDBS cooperation scenarios, each of the algorithms was firstly applied on a low dimensional EEG record with muscular artifact from Fig. 1.1, where approximate result after the artifact rejection is provided by standard ICA. These additional experiments allow to further study the specific algorithm properties and help with a decision whether the algorithm is applicable on a real high-dimensional EEG dataset.

These findings presented in Sections 3.3, 3.4, 3.5, 3.6, and 3.7 serve merely as a demonstration of an effect of the specific algorithm on the processed data temporal traces and a spectral profile due to the reason stated in Chapter 2, Section 2.8. In the very end of this Section, the difference in performance between basic SPICA algorithm and the best performing algorithm is further studied on a task of very strong muscular artifact suppression.

3.1 Optimal EEG Subset Size

For all tests in this section, the records with the strongest simulated EMG contamination (ξ was chosen as 4) were utilized.
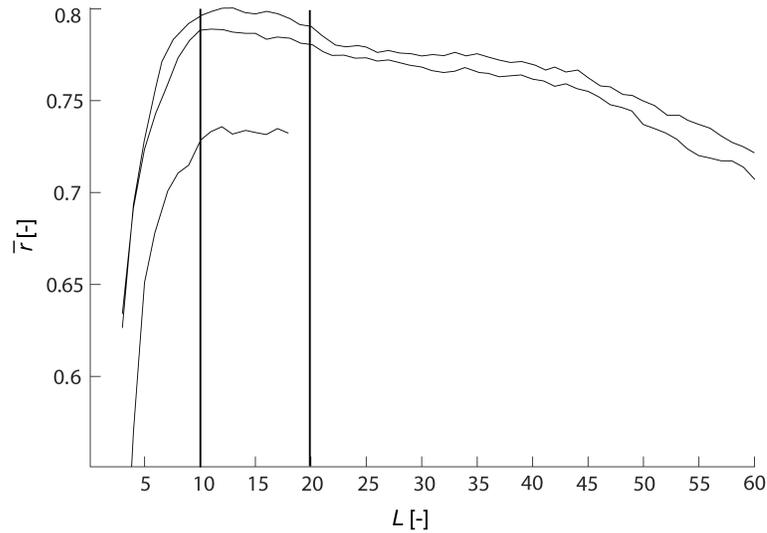
Fig. 3.1: Nearest Neighborhood Sampling: Mean correlation coefficients $\bar{r}$ between 1 second long original clean EEG signals and EEG signals with suppressed muscular artifacts as a function of subspace dimension L for 111, 10/10 and 10/20 electrode systems.
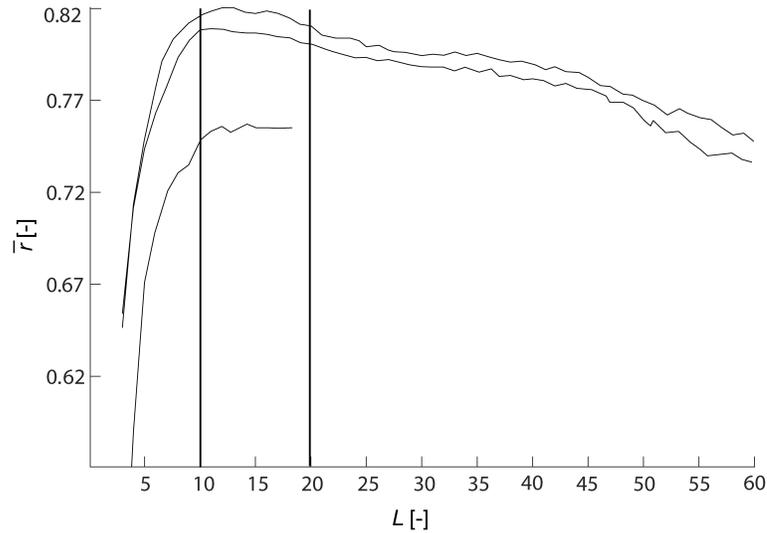

Fig. 3.2: Nearest Neighborhood Sampling: Mean correlation coefficients $\bar{r}$ between 2 second long original clean EEG signals and EEG signals with suppressed muscular artifacts as a function of subspace dimension L for 111, 10/10 and 10/20 electrode systems.

The Fig. 3.1 depicts the values of $\bar{r}$ as a function of different subspace dimensions L for various electrode systems and the Nearest Neighborhood sampling strategy for subspace selection and 1 second long EEG segments for algorithm SPICA. As can be seen for all electrode systems, the best performance is achieved with L within interval 10-20.

This best parameter L value search was also performed for signal lengths 2 and 4 seconds for SPICA and Nearest Neighborhood sampling strategy, which are illustrated in Fig. 3.2 and Fig. 3.3.
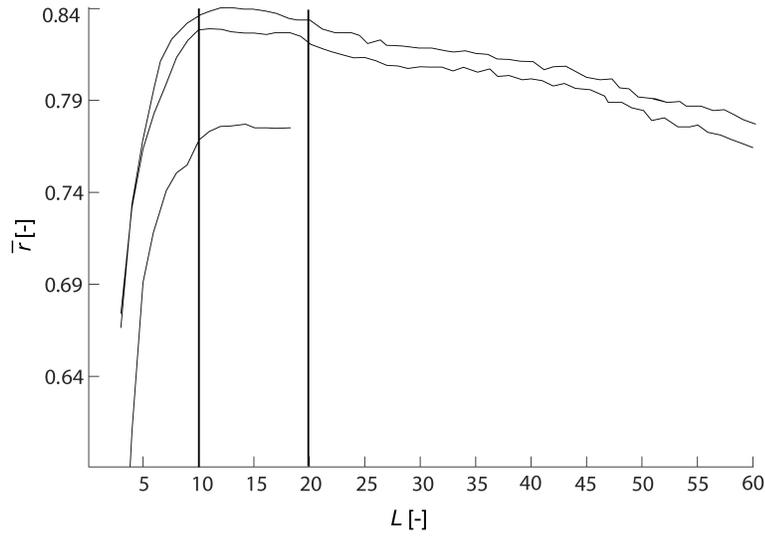
Fig. 3.3: Nearest Neighborhood Sampling: Mean correlation coefficients $\bar{r}$ between 4 second long original clean EEG signals and EEG signals with suppressed muscular artifacts as a function of subspace dimension L for 111, 10/10 and 10/20 electrode systems.
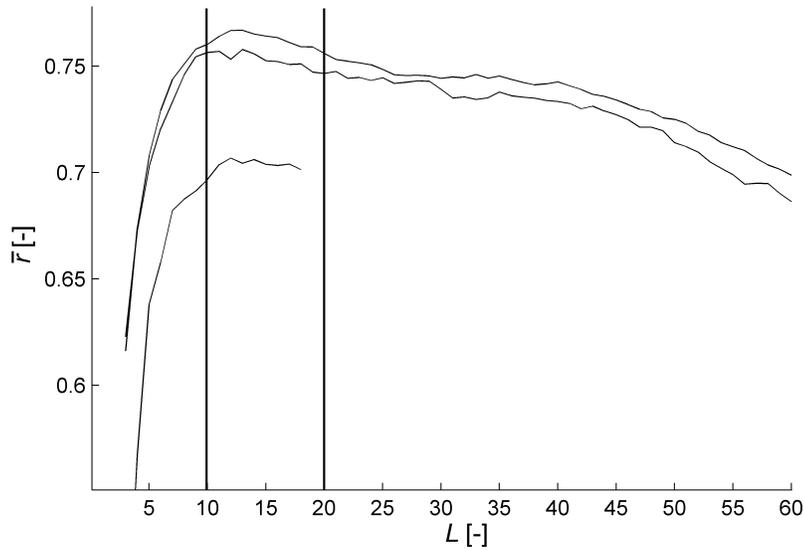

Fig. 3.4: Uniform Sampling: Mean correlation coefficients $\bar{r}$ between 1 second long original clean EEG signals and EEG signals with suppressed muscular artifacts as a function of subspace dimension L for 111, 10/10 and 10/20 electrode systems.

This search for optimal subspace size L was also repeated for the Uniform sampling strategy for subspace selection with parameter N_{rep}, which was equal to 8, and SPICA algorithm. These evaluations are shown in Fig. 3.4 for 1 second long EEG segments, in Fig. 3.5 for 2 second long EEG segments and in Fig. 3.6 for 4 second long EEG segments.

As can be seen for all electrode systems, the best performance is again achieved with L within interval 10-20.
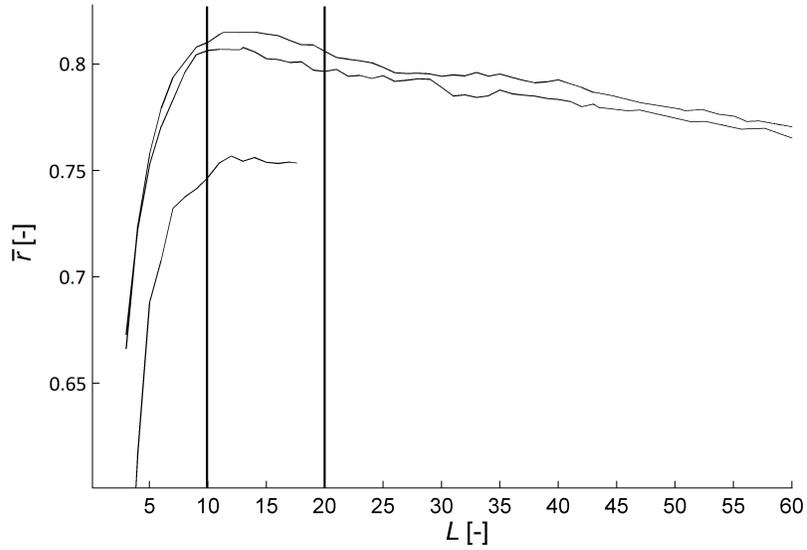
Fig. 3.5: Uniform Sampling: Mean correlation coefficients $\bar{r}$ between 2 second long original clean EEG signals and EEG signals with suppressed muscular artifacts as a function of subspace dimension L for 111, 10/10 and 10/20 electrode systems.
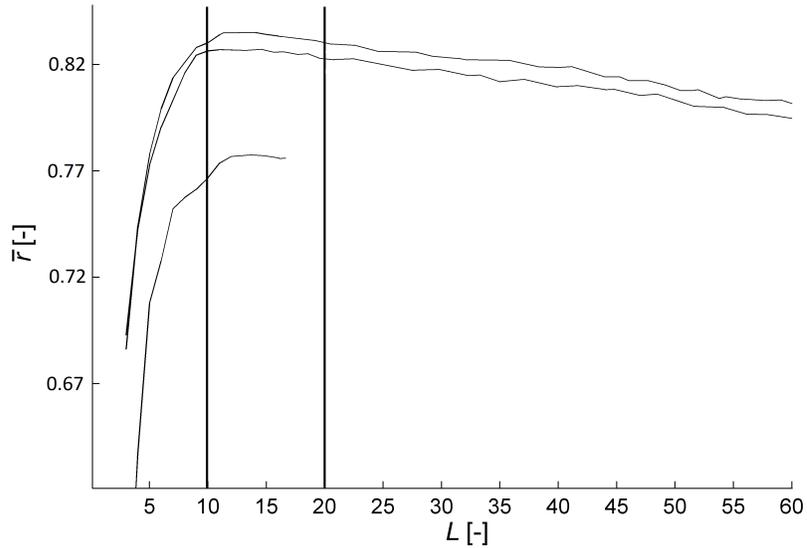
Fig. 3.6: Uniform Sampling: Mean correlation coefficients $\bar{r}$ between 4 second long original clean EEG signals and EEG signals with suppressed muscular artifacts as a function of subspace dimension L for 111, 10/10 and 10/20 electrode systems.

This best parameter L value search was also performed for signal lengths 2 and 4 seconds for SPICA with uniform sampling strategy, which are illustrated in Fig. 3.5 and Fig. 3.6.

The overall conclusion is that the best performance was achievable for L within interval 10-20 in all studied cases.

Since all other algorithms are using either of these sampling strategies and are direct modifications of the basic SPICA, obtained results are expected to be valid for them as well. For further computations, value $L = 12$ is chosen for SPICA, SPICA-LSS as well as for IT-SPICA-SDBS, SDBS initialized SPICA and SPICA-LSS-SDBS algorithms.

3.2   Performance Comparison

In this section, the algorithm SPICA is firstly compared with the state of the art techniques for processing of short EEG segments without ICA overlearning effects to provide a baseline algorithm for further performance comparison. Next, the performance of SPICA algorithm is compared against other proposed algorithms. Namely, these algorithms are SPICA-LSS (Section 2.5), SDBS (Section 2.6), SPICA with SDBS initialization (Section 2.7), SPICA-LSS with SDBS substitution at particular EEG subspaces with a low ICA separation quality (Section 2.7), and SDBS initialized SPICA followed by SDBS postprocessing. The last part of this section provides another way how to compare all proposed methods against the best performing state of the art algorithm, which is based on assessing of the EEG dimensionality reduction inherently introduced by each of the compared methods.

Tables 3.1, 3.2 and 3.3 provide the values of $\bar{r}$ and σ, which were obtained for the basic proposed algorithm SPICA (with the subspace sampling strategy based on nearest neighborhood electrodes and $L = 12$) and other state-of-the-art algorithms. The last column indicates whether a statistically significant difference (SSD) was found between the two sets of values r, which correspond to a respective method and the SPICA algorithm. No effect of different referencing techniques (forehead electrode reference, common average reference, several randomly picked electrodes) on average correlation values was observed.

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

A . . . . a plain FastICA [14, Eq. (14, 21, 24)] used on data that were pre-whitened by PCA [8] with no overlearning suppression,

Ba . . . the high pass filtering method, using a fixed high pass filter with 1Hz cut off frequency [25, 26],

Bb . . . the high pass filtering method, using the AR model based high pass filtering [25, 26],

C . . . . the PCA based method [25, 26],

D . . . . the combination of the PCA and linear regression [21].

The utilized sampling strategies for the EEG subspace selection are denoted by the following abbreviations:

NNE . . . . the sampling strategy based on nearest
    neighborhood electrodes,

US(number) . . . . . the uniform sampling strategy with the
    number in brackets denoting parameter
    N_{rep}.

Tab. 3.1: **The correlation coefficients from the processing of EEG signals without muscular artifacts** ($\xi = 0$)**.** Comparison of SPICA and state of the art methods. Performance values without brackets are the ones obtained with the use of classifier defined in Section 2.3. Performance values within brackets are the ones achieved with the use of classifier from [12]. Symbol * denotes the best achieved performance in the case of methods C and D.

Method	$\bar{r}$	σ	SSD
None	1.0000	0.0000	no
C, 50 PCs	0.8901 (0.8902)	0.0536 (0.0531)	yes
C, 25 PCs	0.9209 (0.9207)	0.0654 (0.0656)	yes
C, 12 PCs	0.9689* (0.9685*)	0.0465 (0.0471)	no
D, 7 PCs	0.2295 (0.2297)	0.0857 (0.0889)	yes
D, 3 PCs	0.3502 (0.3499)	0.0852 (0.0880)	yes
D, 2 PCs	0.4053* (0.4049*)	0.0891 (0.0889)	yes
A	0.8495 (0.8492)	0.0483 (0.0481)	yes
Ba	0.8601 (0.8597)	0.0542 (0.0539)	yes
Bb	0.8805 (0.8807)	0.0484 (0.0482)	yes
SPICA, NNE	0.9831 (0.9840)	0.0236 (0.0231)	-
SPICA, US(8)	0.9761 (0.9764)	0.0243 (0.0237)	no
SPICA, US(4)	0.9701 (0.9700)	0.0254 (0.0261)	no

In order not to skew the conclusions by presenting the suboptimal performance of the state-of-the-art methods, the results for methods C and D are provided with several different numbers of principal components. Moreover, these methods were applied with all the possible values of number of principal components and the best performing case is also shown. (In practice, this search would not be possible, since there is no way of knowing which result provides the highest $\bar{r}$).

Tables 3.4, 3.5 and 3.6 show the values of $\bar{r}$ and σ obtained for all newly proposed methods as well as for the basic proposed algorithm SPICA (with $L = 12$) for comparison purposes. The last column again indicates whether a statistically significant difference (SSD) was found between the sets of values r for respective method and basic SPICA algorithm.

The information provided in Tables 3.4, 3.5 and 3.6 is summarized in Fig. 3.7, where the approximate trends of performance decrease with respect to the strength of present muscular artifacts are shown for each newly developed algorithm.

To examine the dimensionality reduction caused by all proposed algorithms, it is estimated how many strongest singular values constitute 99% of the energy of all singular values for each $\mathcal{D}$. Average counts of strongest singular values are shown in Table 3.7. Note that these numbers are much higher than the number of PCs retained in the application of method C (Table 3.1, 3.2, and 3.3). Therefore, each of the proposed algorithms cause much lower dimensionality reduction, and consequently smaller corruption of clean EEG data.

Tab. 3.2: **The correlation coefficients from the processing of signals with equally strong muscular artifacts and EEG** ($\xi = 1$)**.** Comparison of SPICA and state of the art methods. Performance values without brackets are the ones obtained with the use of classifier defined in Section 2.3. Performance values within brackets are the ones achieved with the use of classifier from [12]. Symbol * denotes the best achieved performance in the case of methods C and D.

Method	$\bar{r}$	σ	SSD
None	0.8738	0.0513	yes
C, 50 PCs	0.8182 (0.8179)	0.0548 (0.0549)	yes
C, 25 PCs	0.8253 (0.8256)	0.0627 (0.0625)	yes
C, 12 PCs	0.8803* (0.8801*)	0.0536 (0.0533)	yes
D, 7 PCs	0.2169 (0.2155)	0.0691 (0.0689)	yes
D, 3 PCs	0.3309 (0.3313)	0.0832 (0.0835)	yes
D, 2 PCs	0.3976* (0.3969*)	0.1040 (0.1051)	yes
A	0.7972 (0.7975)	0.0569 (0.0564)	yes
Ba	0.7909 (0.7913)	0.0657 (0.0651)	yes
Bb	0.8639 (0.8632)	0.0328 (0.0324)	yes
SPICA, NNE	0.9536 (0.9532)	0.0381 (0.0384)	-
SPICA, US(8)	0.9431 (0.9428)	0.0362 (0.0367)	no
SPICA, US(4)	0.9389 (0.9385)	0.0349 (0.0356)	no
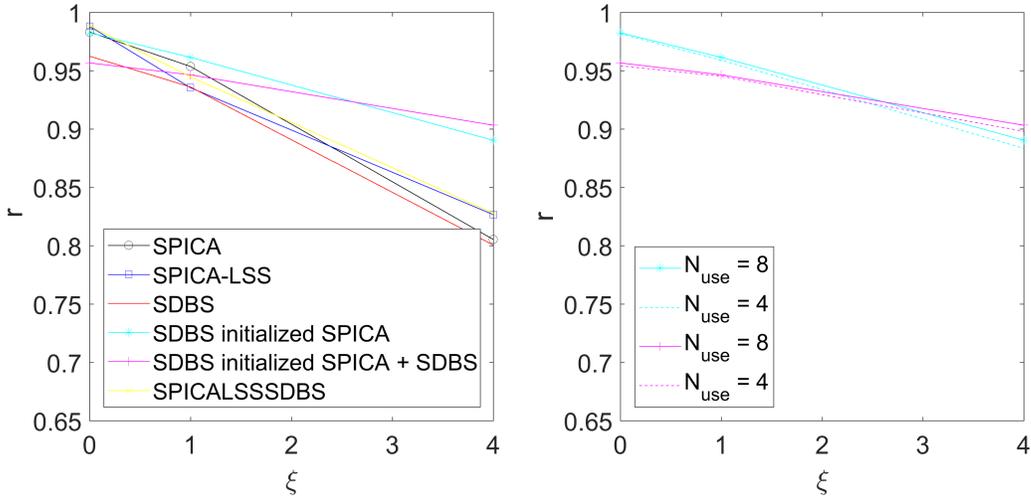
Fig. 3.7:   Mean average correlation coefficients for each of the newly presented methods. Left subplot shows performance curves for all newly presented methods, where uniform sampling strategy with parameter $N_{rep} = 8$ was applied in the case of all algorithms except for SPICA, which utilized nearest neighborhood strategy. Right subplot shows comparison of performance curves for SDBS initialized SPICA with and without SDBS post processing for parameters $N_{rep} = 8$ (solid lines) and $N_{rep} = 4$ (dashed lines).

Tab. 3.3: **The correlation coefficients from the processing of signals with muscular artifacts four times stronger than EEG** ($\xi = 4$)**.** Comparison of SPICA and state of the art methods. Performance values without brackets are the ones obtained with the use of classifier defined in Section 2.3. Performance values within brackets are the ones achieved with the use of classifier from [12]. Symbol * denotes the best achieved performance in the case of methods C and D.

Method	$\bar{r}$	σ	SSD
None	0.5631	0.0834	yes
C, 50 PCs	0.5408 (0.5413)	0.0951 (0.0965)	yes
C, 25 PCs	0.5684 (0.5681)	0.1201 (0.1207)	yes
C, 12 PCs	0.7061* (0.7058*)	0.1057 (0.1064)	yes
D, 7 PCs	0.2390 (0.2396)	0.0912 (0.0923)	yes
D, 3 PCs	0.3831 (0.3827)	0.1109 (0.1093)	yes
D, 2 PCs	0.4049* (0.4053*)	0.0734 (0.0757)	yes
A	0.5251 (0.5262)	0.0641 (0.0626)	yes
Ba	0.5237 (0.5242)	0.0561 (0.0556)	yes
Bb	0.5879 (0.5901)	0.0328 (0.0331)	yes
SPICA, NNE	0.8053 (0.8048)	0.0694 (0.0699)	-
SPICA, US(8)	0.7908 (0.7901)	0.0701 (0.0689)	no
SPICA, US(4)	0.7795 (0.7789)	0.0681 (0.0685)	no

Tab. 3.4: **The correlation coefficients from the processing of EEG signals without muscular artifacts** ($\xi = 0$)**.** Comparison of newly developed methods. Performance values without brackets are the ones obtained with the use of classifier defined in Section 2.3. Performance values within brackets are the ones achieved with the use of classifier from [12].

Method	$\bar{r}$	σ	SSD
None	1.0000	0.0000	no
SPICA, NNE	0.9831 (0.9840)	0.0236 (0.0231)	-
SPICA, US(8)	0.9761 (0.9764)	0.0243 (0.0237)	no
SPICA, US(4)	0.9701 (0.9700)	0.0254 (0.0261)	no
SPICA-LSS, US(8)	0.9880 (0.9882)	0.0103 (0.0101)	no
SPICA-LSS, US(4)	0.9798 (0.9794)	0.0115 (0.0121)	no
SDBS	0.9624 (0.9631)	0.0250 (0.0243)	no
SDBS initialized SPICA, US(8)	0.9824 (0.9819)	0.0169 (0.0165)	no
SDBS initialized SPICA, US(4)	0.9815 (0.9822)	0.0176 (0.0181)	no
SPICA-LSS-SDBS, US(8)	0.9885 (0.9882)	0.0127 (0.0131)	no
SPICA-LSS-SDBS, US(4)	0.9792 (0.9789)	0.0153 (0.0155)	no
SDBS initialized SPICA with SDBS at 2nd step, US(8)	0.9566 (0.9569)	0.0306 (0.0309)	no
SDBS initialized SPICA with SDBS at 2nd step, US(4)	0.9541 (0.9538)	0.0291 (0.0286)	no

Tab. 3.5: **The correlation coefficients from the processing of signals with equally strong muscular artifacts and EEG ($\xi = 1$).** Comparison of newly developed methods. Performance values without brackets are the ones obtained with the use of classifier defined in Section 2.3. Performance values within brackets are the ones achieved with the use of classifier from [12].

Method	$\bar{r}$	σ	SSD
None	0.8738	0.0513	yes
SPICA, NNE	0.9536 (0.9532)	0.0381 (0.0384)	-
SPICA, US(8)	0.9431 (0.9428)	0.0362 (0.0367)	no
SPICA, US(4)	0.9389 (0.9385)	0.0349 (0.0356)	no
SPICA-LSS, US(8)	0.9354 (0.9358)	0.0406 (0.0413)	no
SPICA-LSS, US(4)	0.9271 (0.9270)	0.0351 (0.0349)	no
SDBS	0.9360 (0.9362)	0.0323 (0.0334)	no
SDBS initialized SPICA, US(8)	0.9614 (0.9607)	0.0223 (0.0232)	no
SDBS initialized SPICA, US(4)	0.9590 (0.9593)	0.0234 (0.0246)	no
SPICA-LSS-SDBS, US(8)	0.9447 (0.9453)	0.0282 (0.0288)	no
SPICA-LSS-SDBS, US(4)	0.9385 (0.9381)	0.0278 (0.0283)	no
SDBS initialized SPICA with SDBS at 2nd step, US(8)	0.9464 (0.9468)	0.0325 (0.0328)	no
SDBS initialized SPICA with SDBS at 2nd step, US(4)	0.9452 (0.9451)	0.0330 (0.0328)	no

Tab. 3.6: **The correlation coefficients from the processing of signals with muscular artifacts four times stronger than EEG ($\xi = 4$).** Comparison of newly developed methods. Performance values without brackets are the ones obtained with the use of classifier defined in Section 2.3. Performance values within brackets are the ones achieved with the use of classifier from [12].

Method	$\bar{r}$	σ	SSD
None	0.5631	0.0834	yes
SPICA, NNE	0.8053 (0.8048)	0.0694 (0.0699)	-
SPICA, US(8)	0.7908 (0.7901)	0.0701 (0.0689)	no
SPICA, US(4)	0.7795 (0.7789)	0.0681 (0.0685)	no
SPICA-LSS, US(8)	0.8266 (0.8269)	0.0746 (0.0743)	no
SPICA-LSS, US(4)	0.8173 (0.8168)	0.0713 (0.0698)	no
SDBS	0.8005 (0.8003)	0.0777 (0.0769)	no
SDBS initialized SPICA, US(8)	0.8903 (0.8906)	0.0578 (0.0583)	yes
SDBS initialized SPICA, US(4)	0.8834 (0.8831)	0.0605 (0.0607)	yes
SPICA-LSS-SDBS, US(8)	0.8278 (0.8281)	0.0719 (0.0721)	no
SPICA-LSS-SDBS, US(4)	0.8182 (0.8179)	0.0682 (0.0679)	no
SDBS initialized SPICA with SDBS at 2nd step, US(8)	0.9032 (0.9035)	0.0547 (0.0542)	yes
SDBS initialized SPICA with SDBS at 2nd step, US(4)	0.8980 (0.8974)	0.0571 (0.0565)	yes

Tab. 3.7: **The number of singular values which constitute 99% of overall energy of all singular values of $\mathcal{D}$ for each proposed method.**

Algorithm	$\xi = 0$	$\xi = 1$	$\xi = 4$
SPICA, NNE	78	67	54
SPICA, US(8)	77	65	52
SPICA, US(4)	76	63	51
SPICA-LSS, US(8)	69	58	53
SPICA-LSS, US(4)	67	55	51
SDBS initialized SPICA, US(8)	79	65	51
SDBS initialized SPICA, US(4)	77	63	50
SPICA-LSS-SDBS, US(8)	80	68	56
SPICA-LSS-SDBS, US(4)	78	67	55
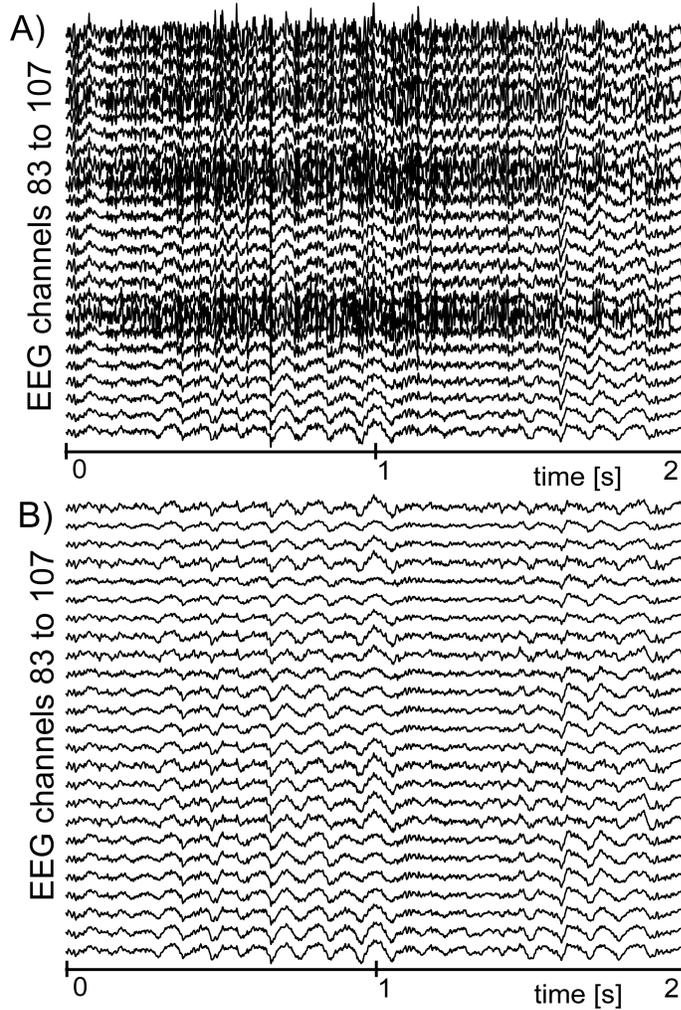
Fig. 3.8: An example of processed EEG signals after application of SPICA algorithm. A) The original EEG signals with artifact. B) EEG signals processed by SPICA. The EEG contains 111 channels, but for spatial reasons only 25 channels are shown.

3.3   SPICA Algorithm Processing Illustration

To illustrate the impact of the SPICA algorithm on processed EEG data, it was applied on a real EEG data with a muscular artifact, which are shown in Fig. 3.8A. The temporal traces and power spectral densities at different locations on scalp as well as topographic maps of power in several frequency bands of the processed signal are shown in Fig. 3.8, 3.9, and 3.10.

As can be seen in Fig. 3.8, muscular activity is strongly suppressed in individual EEG channels. Power spectral densities of processed data depicted at several locations over the scalp in Fig. 3.9 show decrease of power in higher frequency bands after the processing. Fig. 3.10 then illustrates the spatial extent of EMG artifact suppression in several spectral bands, which agrees with the findings in [12]. Overall, the results shown in Fig. 3.8, Fig. 3.9, and Fig. 3.10 indicate that the present muscular contamination was nicely removed, while the effects of overlearning (shown in Fig. 1.2B) were not present in any of the estimated source subsets $\hat{\boldsymbol{s}}_k[n]$ during the processing.
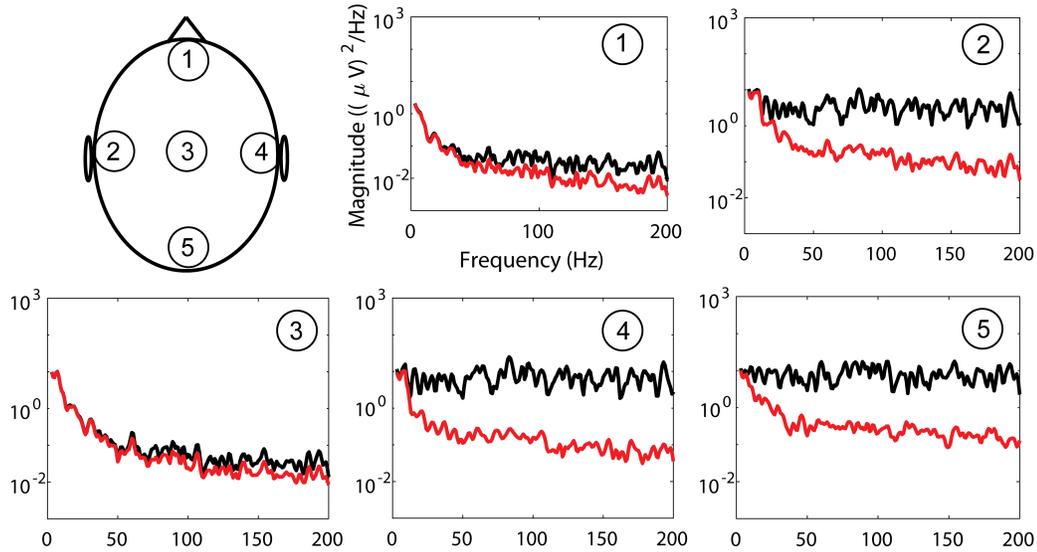
Fig. 3.9: Power spectral densities of EEG dataset with real EMG contamination at several scalp locations before and after processing by SPICA algorithm.
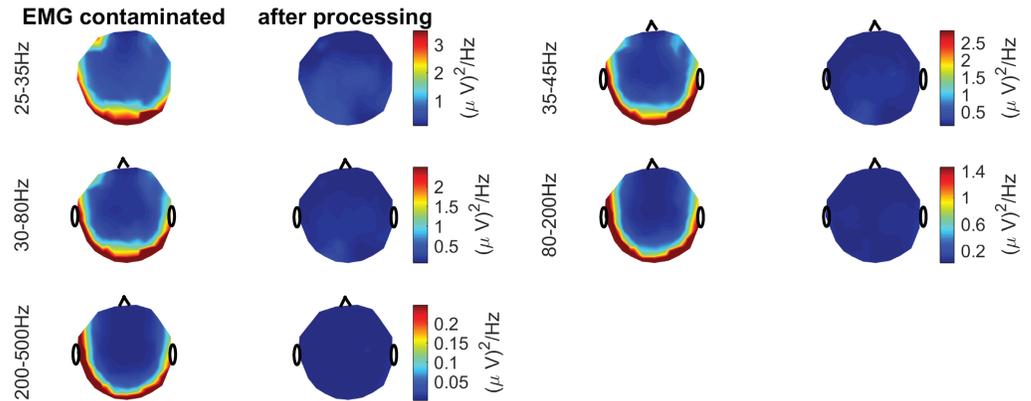
Fig. 3.10: Topographic maps of power distribution over the scalp in respective spectral bands before and after processing by SPICA algorithm.

Fig. 3.11 illustrates the power augmentation in alpha activity band when comparing the power from open vs power from closed eyes epochs of measured EEG for data with and without processing by SPICA algorithm. The shown PSDs support a claim that the detection of alpha band power augmentation phenomenon is not affected by SPICA algorithm. PSDs were averaged across the whole 2-minute interval for particular eyes condition (open, closed).

3.4 SPICA-LSS Algorithm Processing Illustration

The performance of the SPICA-LSS algorithm was tested on the same set of real EEG data with muscular artifact as in the Section 3.3, that is on 111 real EEG signals with EMG artifacts from Fig. 1.2A. The uniform subset sampling strategy is used with the parameter $N_{use} = 8$. The values of ratio c and contraction factor p from Section 2.5 were
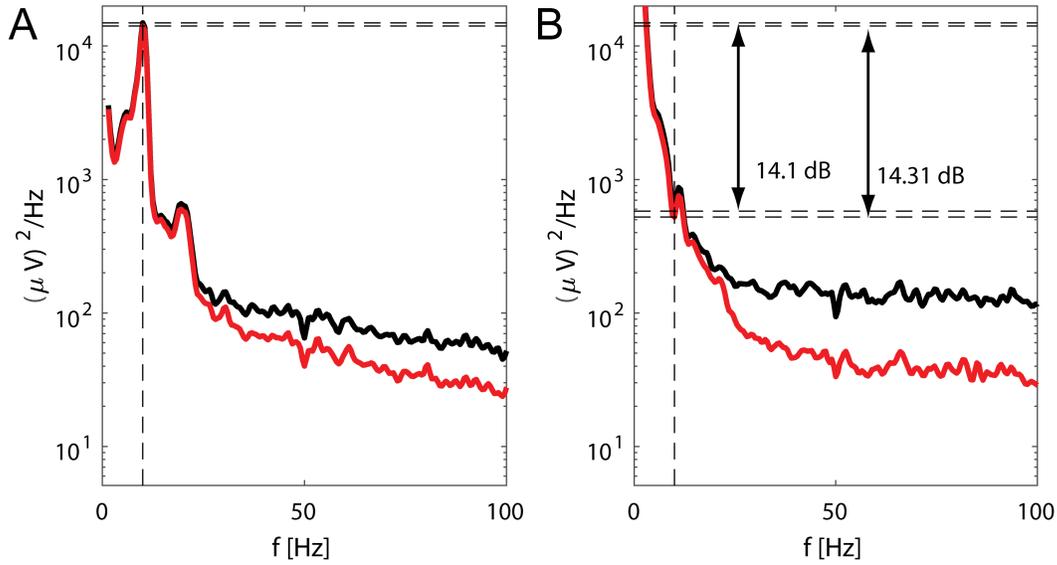
Fig. 3.11: Alpha band power augmentation when comparing EEG epochs under open eyes con-
dition to EEG epochs with closed eyes condition. (A) Average PSDs for resting EEG measured
during closed eyes. (B) Average PSDs for resting EEG measured during open eyes. Black curves
stand for the unprocessed data and the red ones for the data processed by SPICA algorithm.

chosen as 10 and $\sqrt{3}$ providing 4 re-initializations of starting point for ICA algorithm,
if necessary.

The resulting EEG traces as well as comparison of spectral profiles and topographic
maps of power in specific spectral bands are provided in Fig. 3.12, 3.13, and 3.14.
Fig. 3.12 illustrates individual EEG channels after processing. Muscular artifacts are
again quite nicely suppressed. Power spectral densities of the processed signal, which
are shown for several locations over the scalp in Fig. 3.13, indicate expected decrease of
power at higher frequencies. Fig. 3.14 then illustrates the spatial extent of EMG artifact
suppression in several spectral bands.

Overall, based on the results shown in Fig. 3.12, Fig. 3.13, and Fig. 3.14, it can be
tentatively concluded that the EMG artifacts were suppressed. Not surprisingly, the
effects of overlearning (shown in Fig. 1.2B) were not observed in any of the estimated
source subsets $\hat{\boldsymbol{s}}_k[n]$ even for this algorithm.

Fig. 3.15 illustrates the power augmentation in the alpha band when comparing
open to closed eyes epochs of measured EEG for data with and without processing by
the SPICA-LSS algorithm. The shown PSDs, which were averaged across the whole
2-minute interval for particular eyes condition (open, closed), reveal similar expected
decrease of power values, suggesting that the detection of alpha suppression phenomenon
is not affected by SPICA-LSS algorithm.

3.5   SDBS Algorithm Processing Illustration

3.5.1   Illustration of Processing on The Artificial Dataset

To test the ability of SDBS to separate the EEG and EMG subspace as a preprocessing
step for ICA, an artificial linear mixture of 16 EMG signals and 34 artificial EEG signals
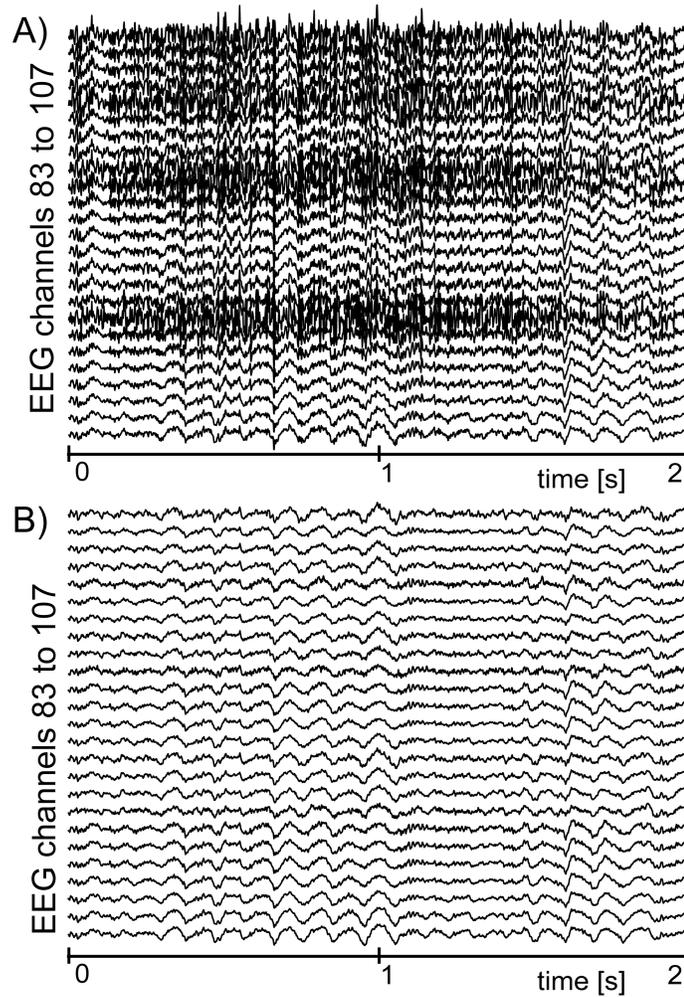
Fig. 3.12: An example of EEG signals processed by SPICA-LSS. A) Original EEG signals with muscular artifact. B) EEG signals processed by SPICA-LSS. The EEG contains 111 channels, but for spatial reasons only 25 channels are shown.

was assembled and processed by SDBS. EEG signals were generated with the use of auto regression model of 12th order, which mimicked the averaged power spectral density of clean EEG signals from open-eye measurements, which were described in Section 2.8. EMG signals were real-world signals from the measurements, which were also described in Section 2.8. All signals were normalized prior to mixing such that their standard deviations were equal. The values of the mixing matrix were set as random numbers with uniform probability density function between 0 and 1. All utilized signals are 1-second-long records of the respective type of the signal. The achieved separation between the two respective subspaces is nicely depicted in Fig. 3.16. The product of the mixing matrix **A** and the separation matrix **B** shown in Fig. 3.16 clearly manifests two mutually separated subspaces of EEG and EMG sources. The fact that the source signals of the same class, which is either EEG or EMG, are not separated, which would be demonstrated by the ideal unity matrix or its version with permutations of rows, is not concerning in the task of muscular artifacts removal, which is the main topic of this work.
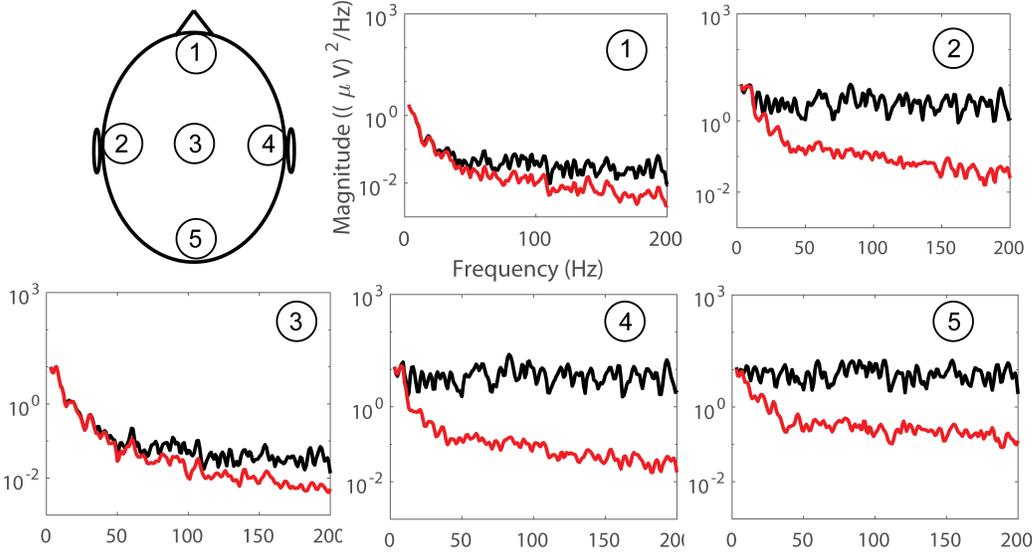
Fig. 3.13: Power spectral densities of EEG dataset with real EMG contamination at several scalp locations before and after processing by SPICA-LSS algorithm.
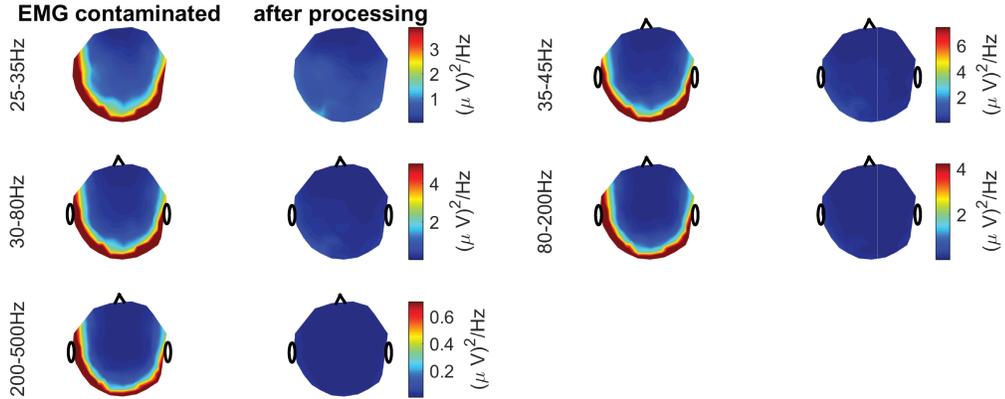
Fig. 3.14: Topographic maps of power distribution over the scalp in respective spectral bands before and after processing by SPICA-LSS algorithm.

Signals shown in Fig. 3.17 are 50 resulting sources after the processing of the above described artificial dataset. These signals also demonstrate the separation ability of the SDBS with regards to EEG and EMG signal subspaces. Two notes should be made about appearance of all estimated sources. First, the ideal situation of having the same number of sources as the measured signals was simulated, therefore no performance deterioration as a result of presence of more sources than measuring channels was observed. Second, generated EEG and selected EMG signals, which are present in the mixture, were normalized such that their power was similar and their signal characteristics made them visually easily distinguishable. This verification step was however utilized to demonstrate the ability of this technique to separate EEG/EMG subspaces under the ideal model settings, which was successfully proven in Fig. 3.16 and Fig. 3.17.
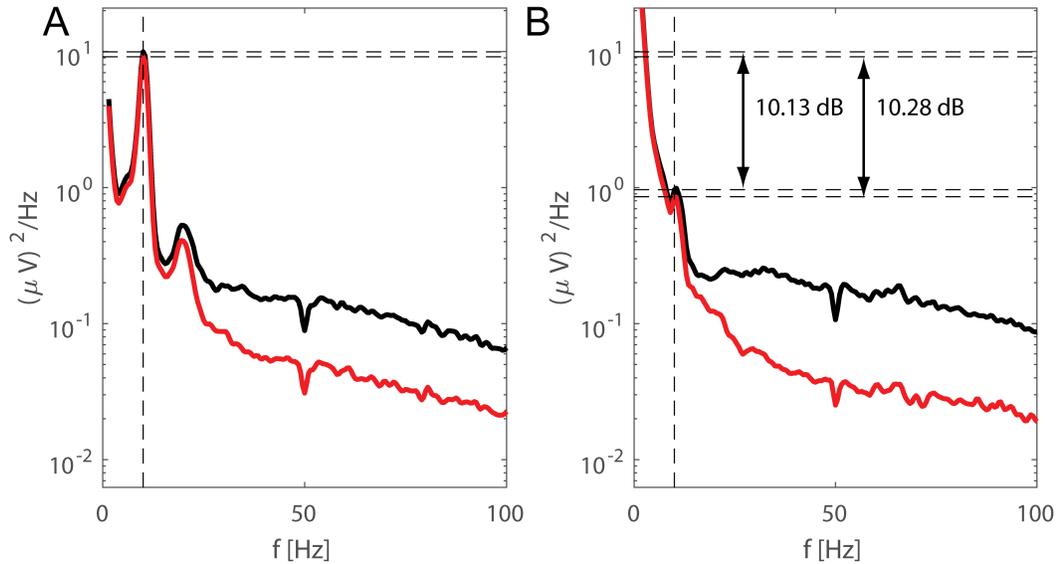
Fig. 3.15: Alpha band power augmentation in EEG record for open/closed eyes condition. (A) Average PSDs for resting EEG measured during closed eyes. (B) Average PSDs for resting EEG measured during open eyes. Black curves stand for the unprocessed data and the red ones for the data processed by SPICA-LSS algorithm.
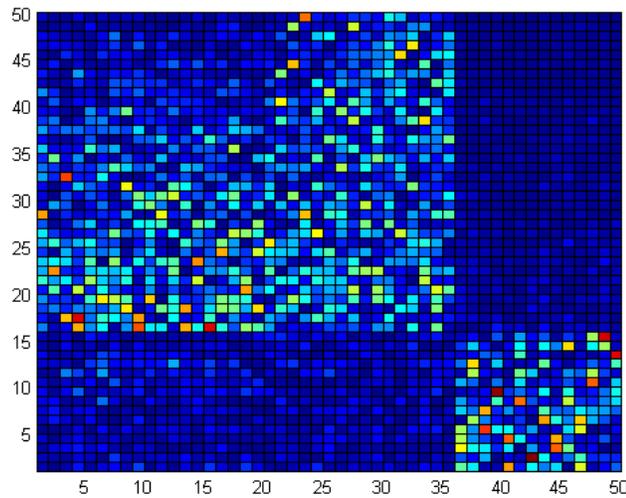
Fig. 3.16: The separation of EEG and EMG subspaces illustrated on a product of a mixing matrix **A** and estimated separation matrix **B**.

3.5.2   Illustration of SDBS Processing on The Real EEG Dataset

To further test SDBS algorithm in the case of processing of real high-dimensional EEG data with muscular artifacts, it was applied on the same real EEG data with muscular artifact as was utilized in Section 3.3.

The resulting EEG traces as well as the impact on spectral profile and topographic distribution of power in several frequency bands obtained when the SDBS algorithm
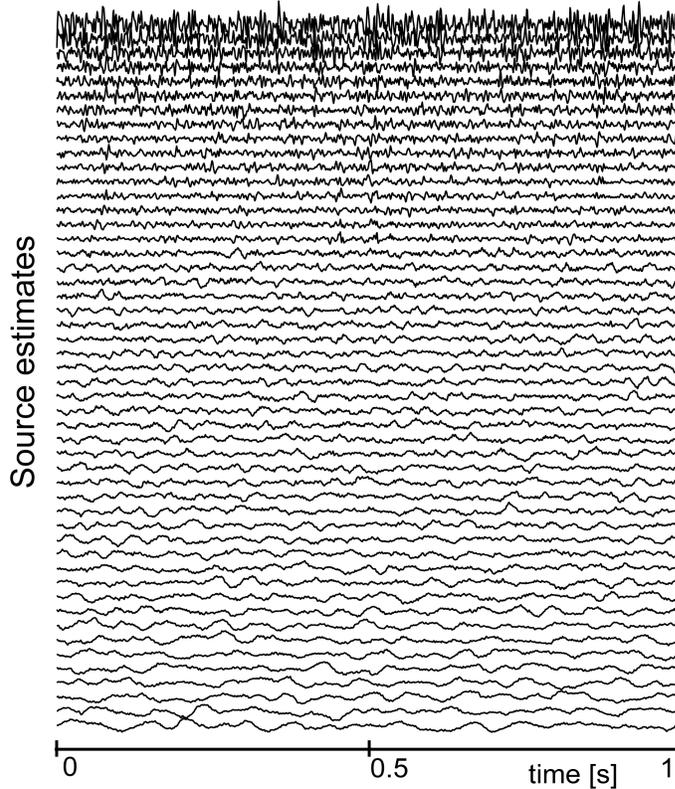
Fig. 3.17:   Example of EMG and EEG sources used for the signal mixture in this illustration.

was applied to 111 real EEG signals with EMG artifacts from Fig. 1.2A are shown in Fig. 3.18, 3.19, and 3.20.

Fig. 3.18 reveals individual EEG channels after processing. Strong muscular artifacts are no longer visible, although one can recognize higher amount of power at high frequencies located in each channel when compared with EEG traces from Fig. 3.8 and Fig. 3.12. Power spectral densities depicted at several locations over the scalp in Fig. 3.19 show similar finding as in the case of the preceding illustration Sections, that is the expected decrease of power in higher frequency bands after the processing. Fig. 3.20 then illustrates the spatial extent of EMG artifact suppression in several spectral bands. Overall, the results illustrated in Fig. 3.18, Fig. 3.19, and Fig. 3.20 show that the EMG artifacts were suppressed. No effects of overlearning were present in the set of signals $\tilde{\mathbf{s}}[n]$.

Fig. 3.21 illustrates the power augmentation in the alpha band when comparing open to closed eyes epochs of measured EEG for data with and without processing by the SDBS algorithm. Similar expected increase of power values was observed in the shown PSDs at alpha band when comparing processed and raw EEG datasets for particular eyes condition (open, closed). PSDs were averaged across the whole 2-minute interval for particular eyes condition (open, closed). Consequently, the detection of alpha power augmentation phenomenon does not seem to be affected by SDBS algorithm.
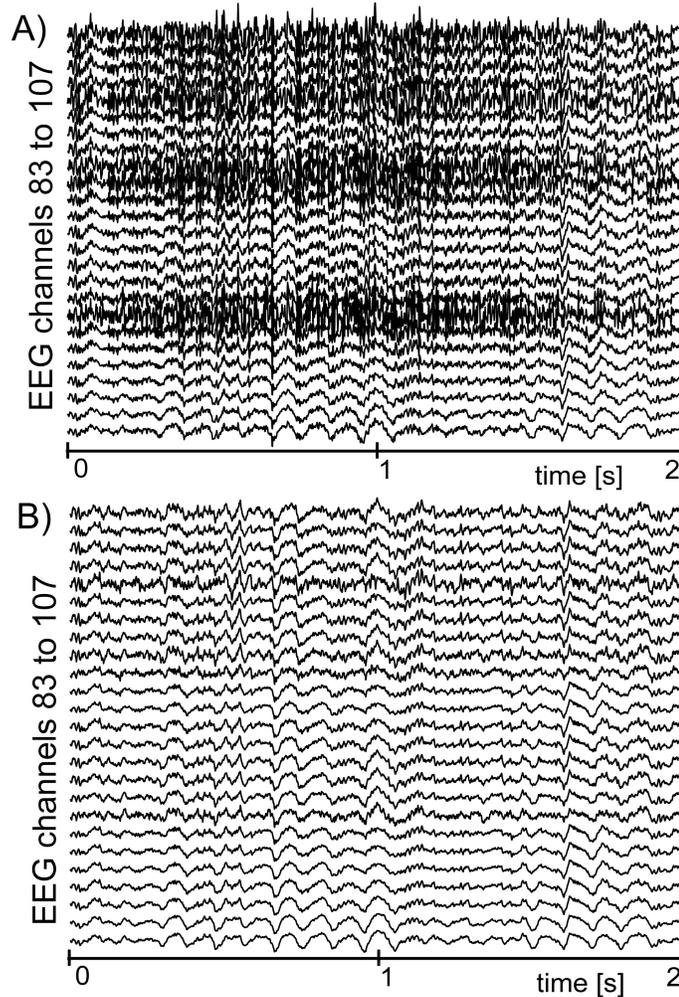
Fig. 3.18: An example of EEG signals processed by SDBS. A) Original EEG signals with muscular artifact. B) EEG signals processed by SDBS. For spatial reasons, only 25 channels out of 111 are shown.

3.6 SPICA and SDBS Cooperation Scenarios Processing Illustration

In this part of the Chapter, the impact of cooperation strategies between SPICA/SPICA-LSS and SDBS on EEG data is illustrated and the behavior of each proposed structure is discussed.

First, each of the scenarios was tested on low dimensional EEG signal with muscular artifact (9 dimensions, length 2 seconds) shown in Fig. 1.1. Consequently, if no problems with the separation were observed, then the high dimensional EEG signal (111 dimensions, length 2 seconds) shown in Fig. 1.2 was processed next.

IT-SPICA-SDBS

The algorithm IT-SPICA-SDBS appeared to be problematic even in the low dimensional signal case. Namely, algorithm was not able to reach a stable convergence, which was defined in Section 2.2 as the average value of scalar products between estimated separation vectors from last and current iteration, and the quality of resulting source
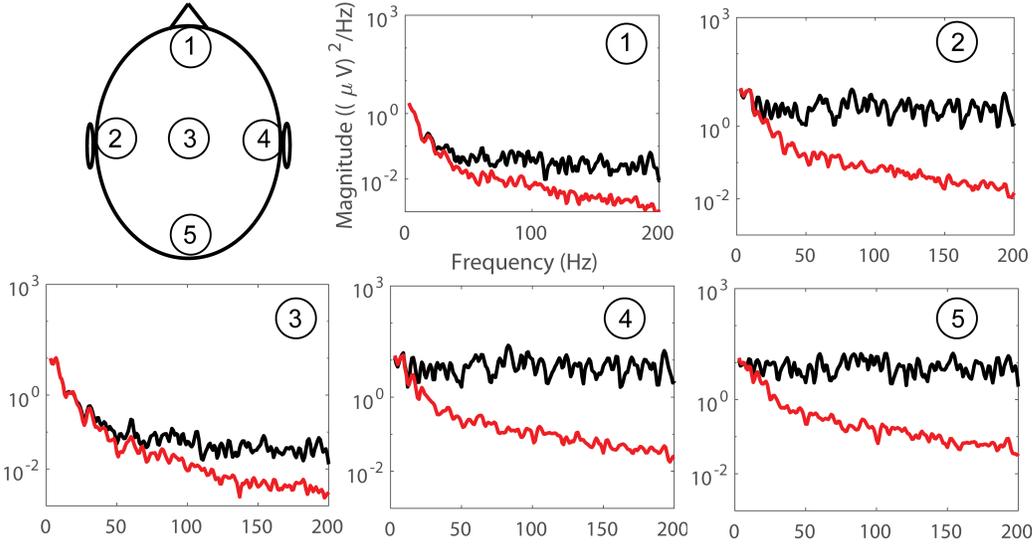
Fig. 3.19:   Power spectral densities of EEG dataset with real EMG contamination at several scalp locations before and after processing by SDBS algorithm.
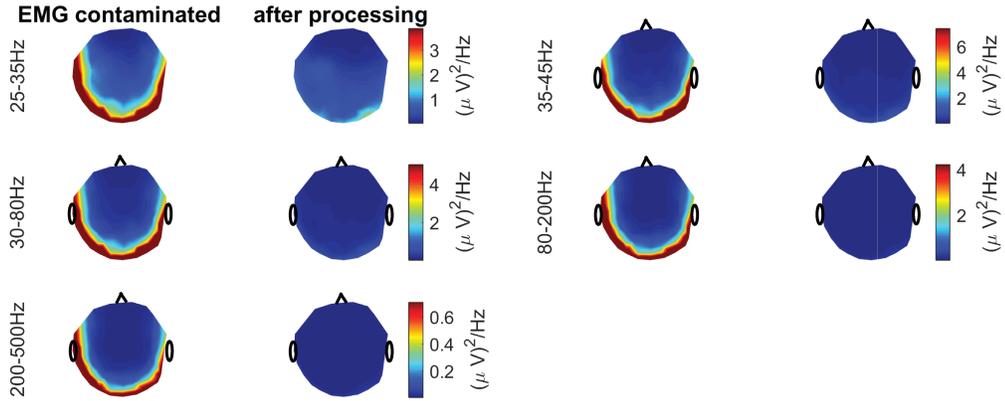
Fig. 3.20:   Topographic maps of power distribution over the scalp in respective spectral bands before and after processing by SDBS algorithm.

estimates varied significantly at each processing attempt. The Fig. 3.22 shows the convergence curve for 200 iterations after which the processing was stopped.

The possible reason for this behavior is depicted in the following simple feature space example in 2 dimensional space, which is depicted in Fig. 3.23. For the sake of simplicity, this example illustrates the behavior of IT-SPICA-SDBS in case of 2 measured EEG channels with muscular artifacts.

As it can be clearly seen, if the location of extrema of FastICA and SDBS algorithm objective functions is significantly different, then the whole scenario converges to a point between these locations and the BSS algorithms will work against each other. Moreover, even if this state could be captured and evaluated as a convergence point, there is no guarantee that it resembles any meaningful estimates of the sources, since none of the objectives was successfully maximized.
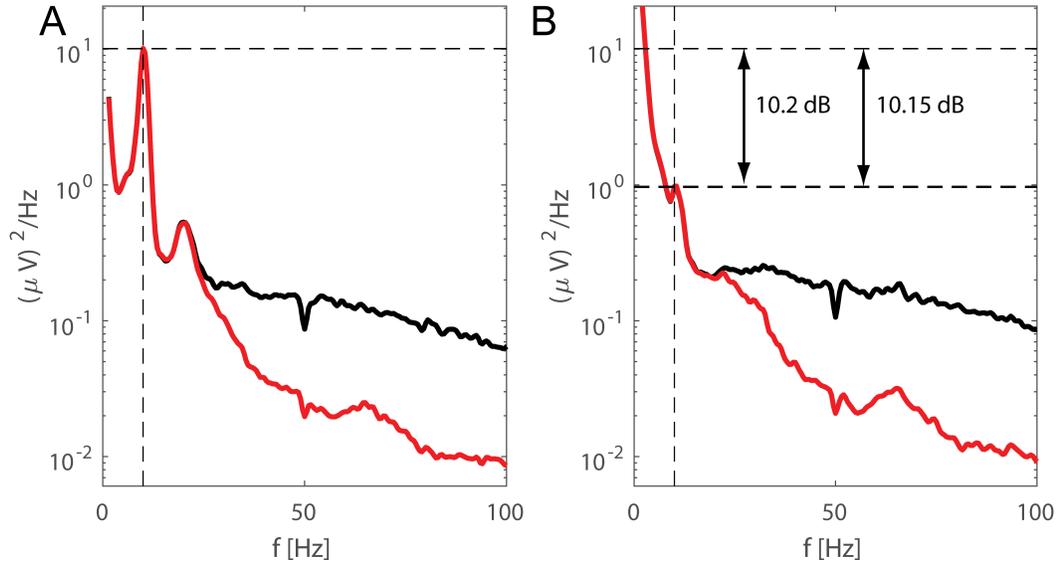
Fig. 3.21: Alpha band power augmentation in EEG record for open/closed eyes condition. (A) Average PSDs for resting EEG measured during closed eyes. (B) Average PSDs for resting EEG measured during open eyes. Black curves stand for the unprocessed data and the red ones for the data processed by SDBS algorithm.
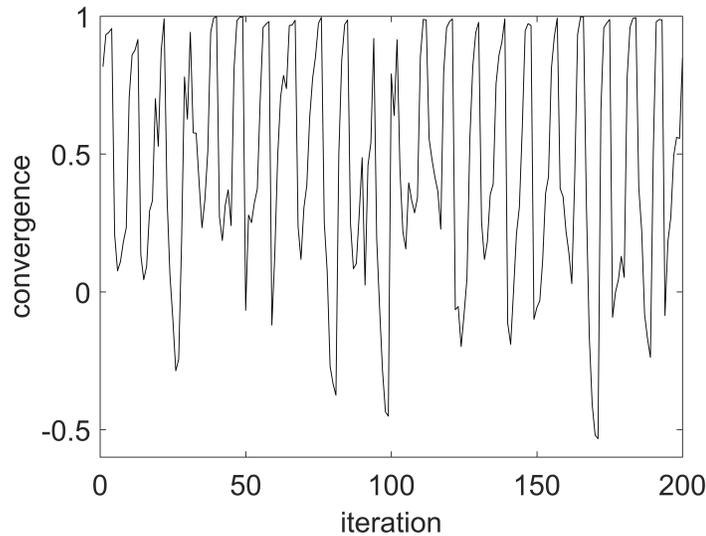

Fig. 3.22: Example of unstable convergence in case of iterative FastICA and SDBS cooperation. The convergence values on the y-axis stand for the average dot product between estimated separation vectors from last and current iteration.

SDBS Initialized SPICA

The SDBS initialized SPICA showed a stable convergence and was therefore applied to the signal dataset which is shown in Fig. 1.2A. The number of channels in each subspace L was again chosen as 12 as in the case of SPICA algorithm. Uniform subset sampling strategy is used with the parameter $N_{rep} = 8$.
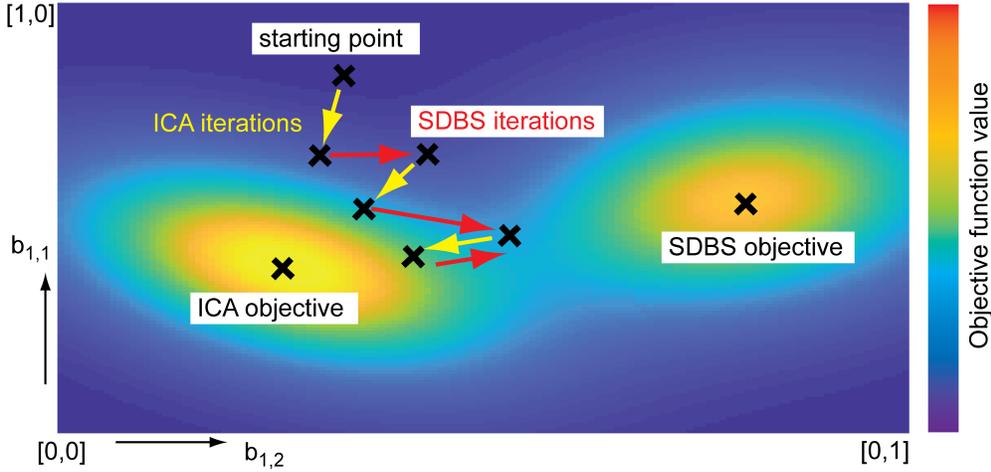
Fig. 3.23: Simple feature space example of iterative FastICA and SDBS cooperation. The values of components $[b_{1,1}, b_{1,2}]$ in one separation vector $\vec{b}$ are on horizontal and vertical axes and the color represents the value of non-gaussianity of resulting estimated source.

The results in the form of EEG traces, PSDs and topographic power distribution changes, which were obtained when the proposed method was applied to 111 real EEG signals with EMG artifacts shown in Fig. 1.2A, are shown in Fig. 3.24, 3.25, and 3.26.

Fig. 3.24 shows individual EEG channels after processing with suppressed muscular artifacts. The power spectral densities depicted at several locations over the scalp in Fig. 3.25 show decrease of power in higher frequency bands after the processing. Fig. 3.26 then illustrates the spatial extent of EMG artifact suppression in several spectral bands. Based on results provided in Fig. 3.24, Fig. 3.25, and Fig. 3.26, it appears that the muscular contamination was suppressed. This algorithm also did not manifest any effects of overlearning (shown in Fig. 1.2B) in any of the estimated source subsets $\hat{\boldsymbol{s}}_k[n]$.

In Fig. 3.27, the power augmentation in alpha band when comparing open to closed eyes epochs of measured EEG is shown. PSDs, where the alpha band power augmentation can be observed, are provided for data with and without processing by SDBS initialized SPICA. The shown PSDs, which were averaged across the whole 2-minute interval for particular eyes condition (open, closed), yield similar expected increase of power values at alpha frequency band, proving that the detection of alpha power augmentation phenomenon is not affected by SDBS initialized SPICA.

SPICA-LSS-SDBS

The SPICA-LSS-SDBS also did not exhibit any problems regarding the achievement of the convergence in the case of processing low-dimensional 9 channel EEG and was applied on the signal dataset which is shown in Fig. 1.2A. The number of channels in each subspace L was again chosen as 12 as in the case of SPICA algorithm. Uniform subset sampling strategy is used with the parameter N_{rep} being 8. The values of ratio c and lowering factor p were chosen as 10 and $\sqrt{3}$ providing 4 re-initializations of starting point for the ICA algorithm, if necessary.

The resulting EEG traces, impact on spectral profiles and topographic distribution of power over scalp obtained when the SPICA-LSS-SDBS was applied to 111 real EEG signals with EMG artifacts from Fig. 1.2A are illustrated in Fig. 3.28, 3.29, and 3.30.
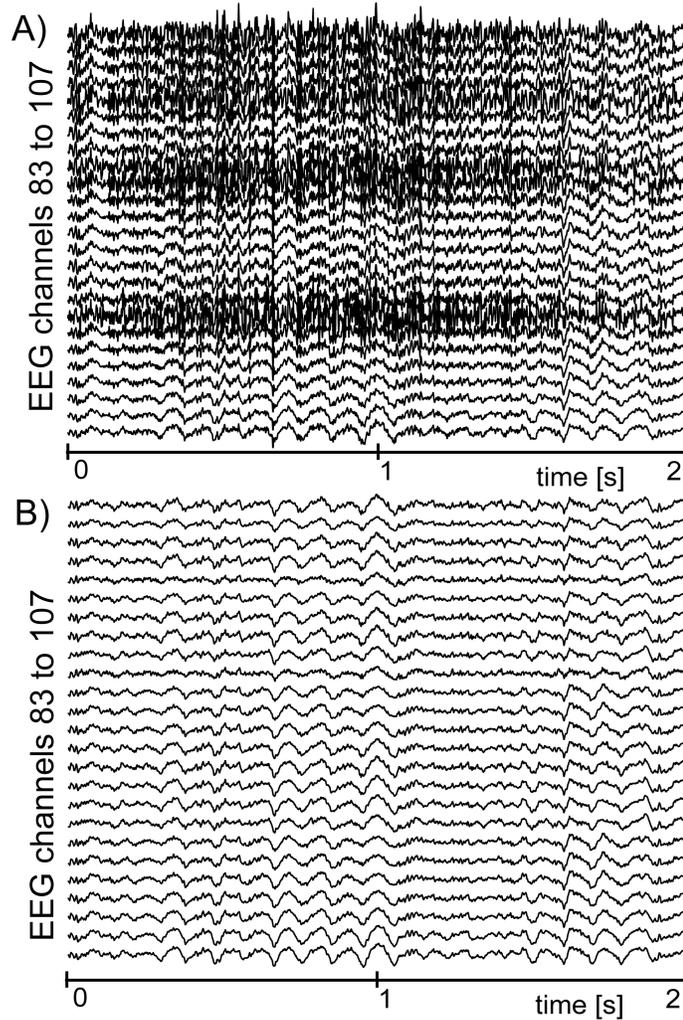
Fig. 3.24: An example of EEG signals processed by SDBS initialized SPICA. A) Original EEG signals with muscular artifacts. B) EEG signals processed by SDBS initialized SPICA. For spatial reasons, only 25 channels out of 111 are shown.

Fig. 3.28 provides a detailed view on individual processed EEG channels. As in the case of previous algorithms, it is challenging to recognize any residual muscular activity in Fig. 3.28B. Power spectral densities are depicted at several locations over the scalp in Fig. 3.29. As in the case of each previous algorithms, one can observe expected decrease of power in higher frequency bands after the processing. Fig. 3.30 illustrates the spatial extent of EMG artifact suppression in several spectral bands. Overall, results shown in Fig. 3.28, Fig. 3.29, and Fig. 3.30 reveal that the EMG artifacts were suppressed, while effects of overlearning (shown in Fig. 1.2B) were not observed in any of the estimated source subsets $\hat{\boldsymbol{s}}_k[n]$ during the processing.

Fig. 3.31 shows the power augmentation in alpha band when comparing open to closed eyes epochs of measured EEG. This phenomenon is illustrated for data with and without processing by SPICA-LSS-SDBS. Based on the shown PSDs, which were averaged across the whole 2-minute interval for particular eyes condition (open, closed), the detection of alpha band power augmentation phenomenon does not appear to be affected by the SPICA-LSS-SDBS as the value of power augmentation in alpha band,
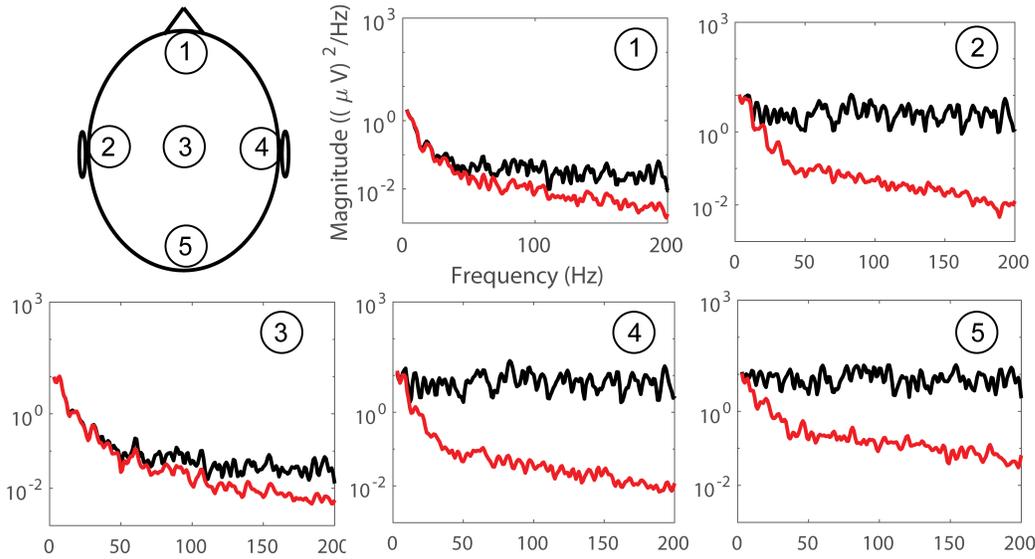
Fig. 3.25: Power spectral densities of EEG dataset with real EMG contamination at several scalp locations before and after processing by SDBS initialized SPICA.
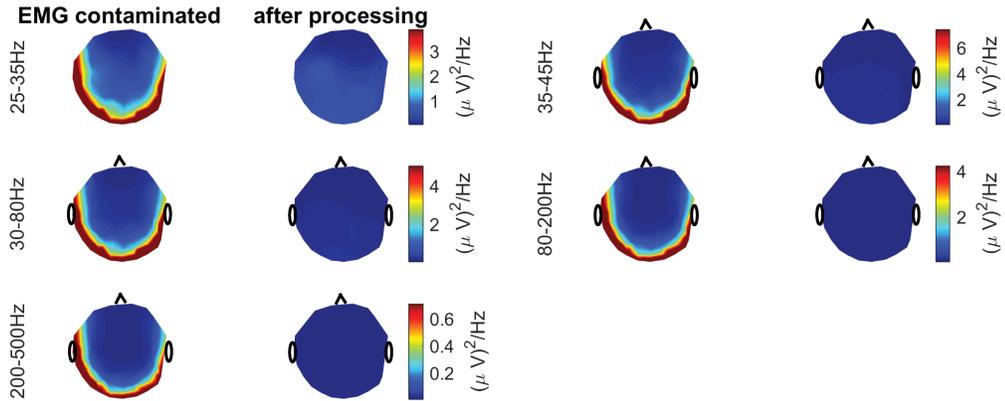

Fig. 3.26:   Topographic maps of power distribution over the scalp in respective spectral bands before and after processing by SDBS initialized SPICA.

when comparing EEG intervals for open eyes and closed eyes condition, stays virtually the same for unprocessed intervals as well as intervals processed by the SPICA-LSS-SDBS.

3.7   SDBS Initialized SPICA Followed by SDBS Algorithm - Processing Illustration

The resulting EEG traces, changes in spectral profile as well as topographic distribution of power across scalp obtained when SDBS initialized SPICA with SDBS postprocessing was applied to 111 real EEG signals with EMG artifacts from Fig. 1.2A are depicted in Fig. 3.32, 3.33, and 3.34.
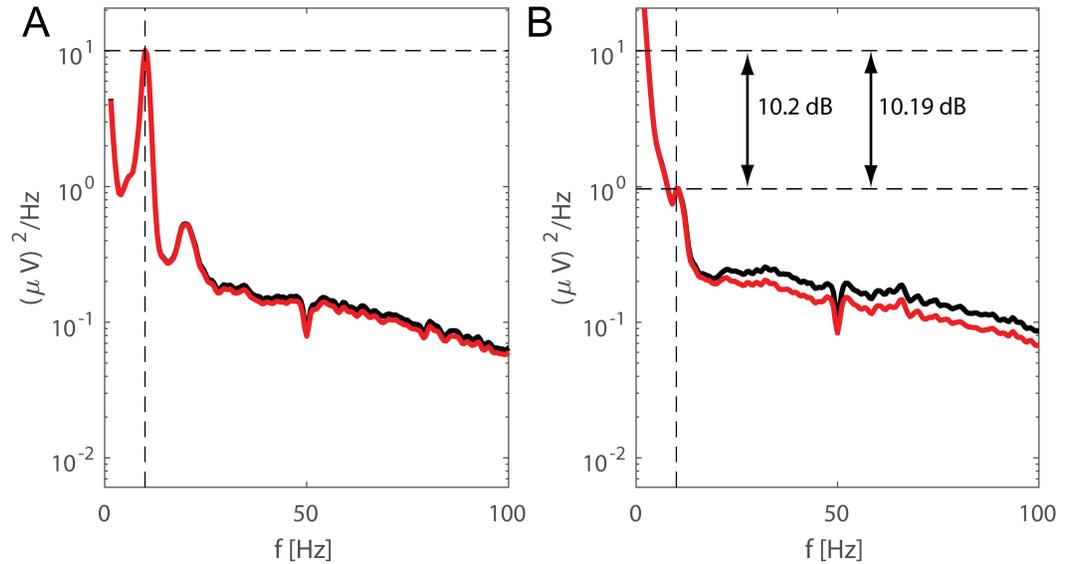
Fig. 3.27: Alpha band power augmentation when comparing EEG epochs under open eyes condition to EEG epochs with closed eyes condition. A: Average PSDs for resting EEG measured during closed eyes. B: Average PSDs for resting EEG measured during open eyes. Black curves stand for the unprocessed data and the red ones for the data processed by SDBS initialized SPICA.

Fig. 3.32 reveals individual EEG channels after processing with no visible muscular contamination. Power spectral densities are depicted at several locations over the scalp in Fig. 3.33. Fig. 3.34 then illustrates the spatial extent of EMG artifact suppression in several spectral bands. Overall, results shown in Fig. 3.32, Fig. 3.33, and Fig. 3.34 reveal the successful suppression of the EMG artifacts.

Fig. 3.35 illustrates the power augmentation in alpha band when comparing open to closed eyes epochs of measured EEG. This phenomenon was compared between data with and without processing by SDBS initialized SPICA followed by SDBS postprocessing step. The detection of alpha band power augmentation phenomenon is clearly not affected by SDBS initialized SPICA algorithm with SDBS post-processing, since the decrease of power between two conditions (closed vs open eyes) in the shown PSDs, which were averaged across the whole 2-minute interval for particular eyes condition (open, closed), was retained at very similar values for unprocessed as well as processed EEG intervals.

3.8 Comparison of Spectral Profile Impact on A Task of Alpha Augmentation

To compare the effect of newly developed algorithms on a spectral profile of processed EEG signals, the Fig. 3.36 shows the PSDs obtained when ability of each proposed method to preserve alpha band power augmentation, which was presented in Fig. 3.11, 3.15, 3.21, 3.27, 3.31, and 3.35.
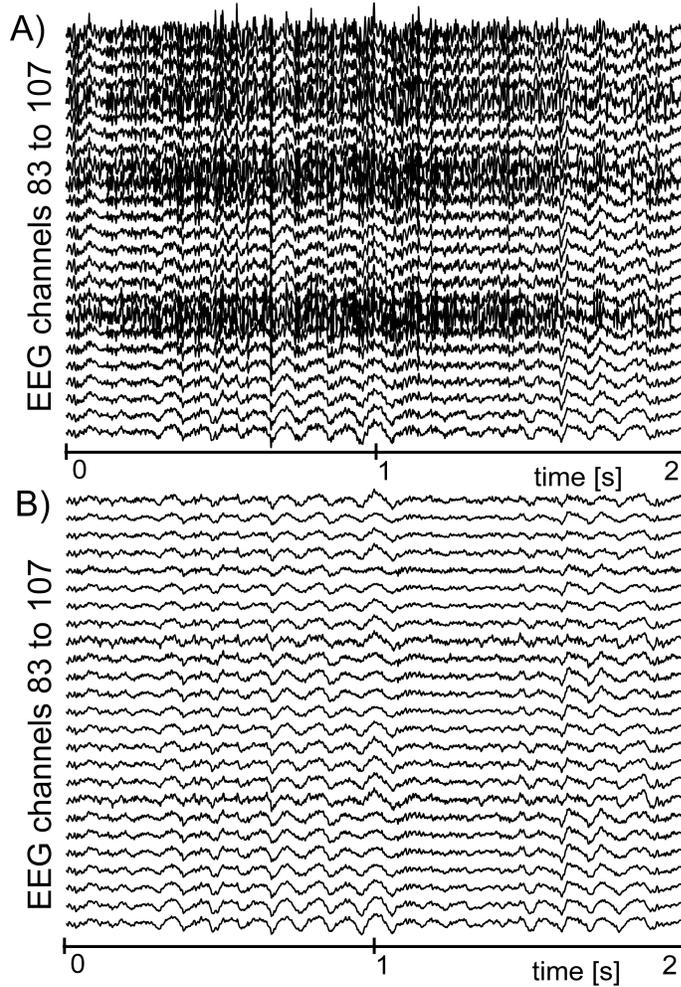
Fig. 3.28:   An example of EEG signals processed by SPICA-LSS-SDBS. The original signals with muscular artifact are shown in subplot A). EEG data processed by SPICA-LSS-SDBS are shown in subplot B). For spatial reasons, only 25 channels out of 111 are shown.

3.9   Strong Artifact Performance Comparison

To further investigate relatively large qualitative difference in performance metrics between the basic SPICA algorithm and the SDBS initialized SPICA with and without the SDBS postprocessing, we applied those algorithms to a segment of 1 s long EEG signal with strong muscular artifact, where the SNR in most contaminated channels is estimated to be -19.65 dB.

Fig. 3.37 shows individual EEG channels before and after processing with respective algorithms. As can be seen, there are still some remaining artifactual contents in the EEG channels, which were processed by the basic SPICA algorithm.

Power spectral densities depicted at several locations over the scalp are depicted in Fig. 3.38. While all utilized algorithms provide some degree of expected decrease of power in higher frequency bands after the processing, this decrease is significantly lower in case of SPICA and edge electrodes with low SNR.

Fig. 3.39 illustrates the spatial extent of EMG artifact suppression in several spectral bands for each compared algorithm.
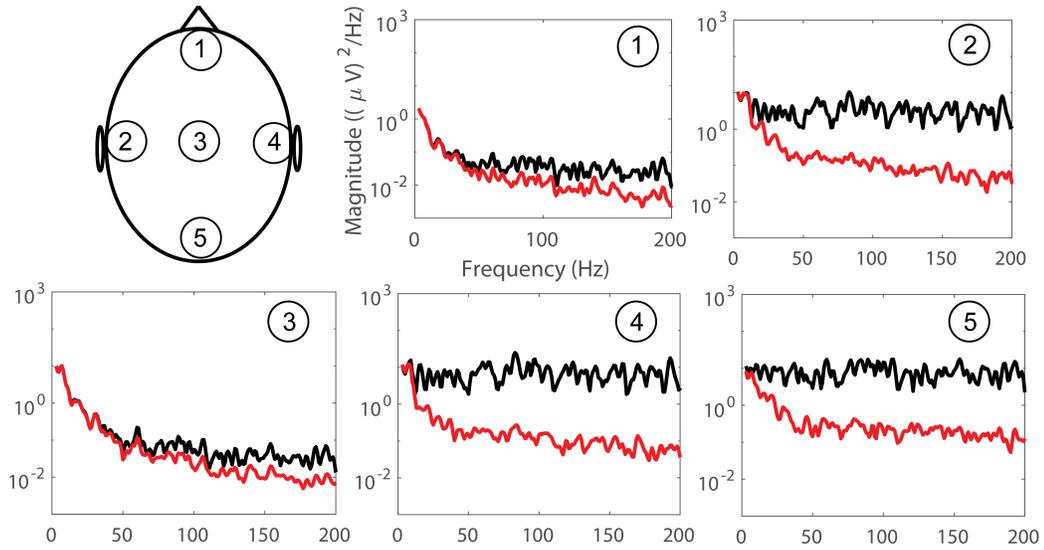
Fig. 3.29: Power spectral densities of EEG dataset with real EMG contamination at several scalp locations before and after processing by SPICA-LSS-SDBS.
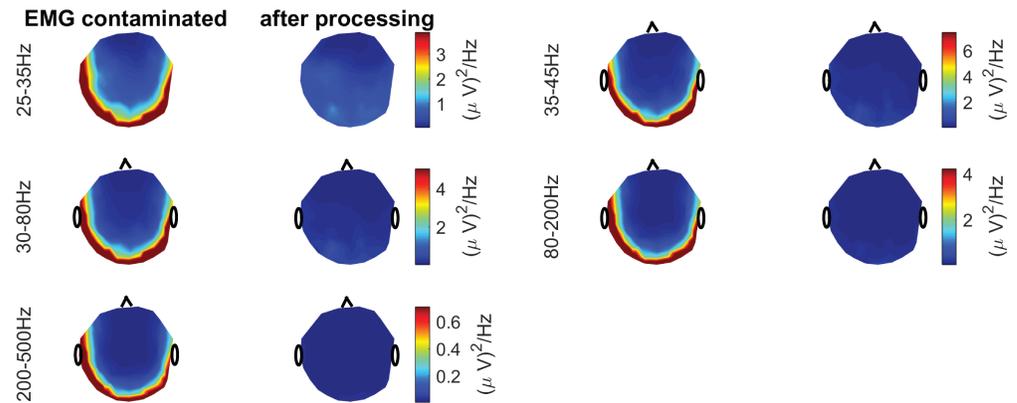

Fig. 3.30: Topographic maps of power distribution over the scalp in respective spectral bands before and after processing by SPICA-LSS-SDBS.
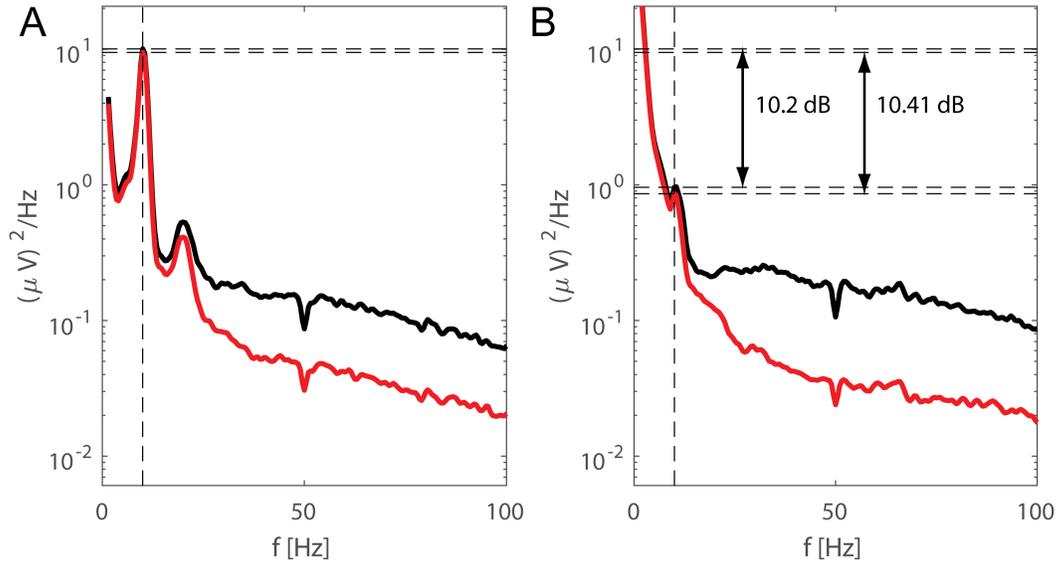
Fig. 3.31:   Alpha band power augmentation when comparing EEG epochs under open eyes condition to EEG epochs with closed eyes condition. (A) Average PSDs for resting EEG measured during closed eyes. (B) Average PSDs for resting EEG measured during open eyes. Black curves stand for the unprocessed data and the red ones for the data processed by SPICA-LSS-SDBS.
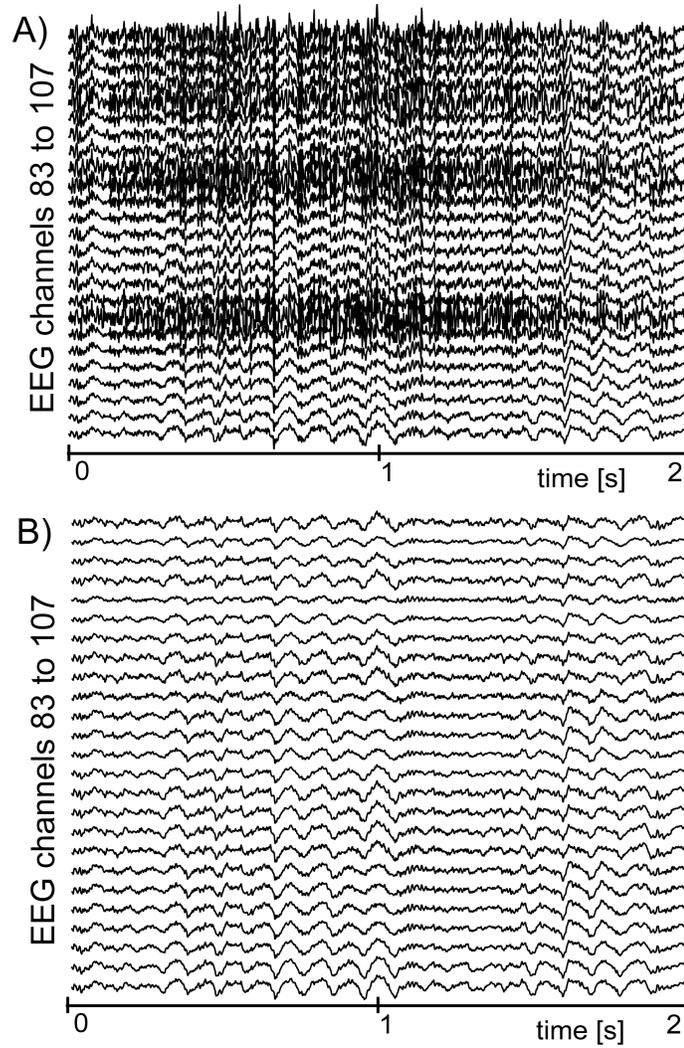
Fig. 3.32: An example of EEG signals processed by SDBS initialized SPICA algorithm with SDBS post-processing. The original EEG signals with muscular artifacts are shown in subplot A). The EEG signals processed by SDBS initialized SPICA with SDBS postprocessing are shown in subplot B). For spatial reasons, only 25 channels out of 111 are shown.
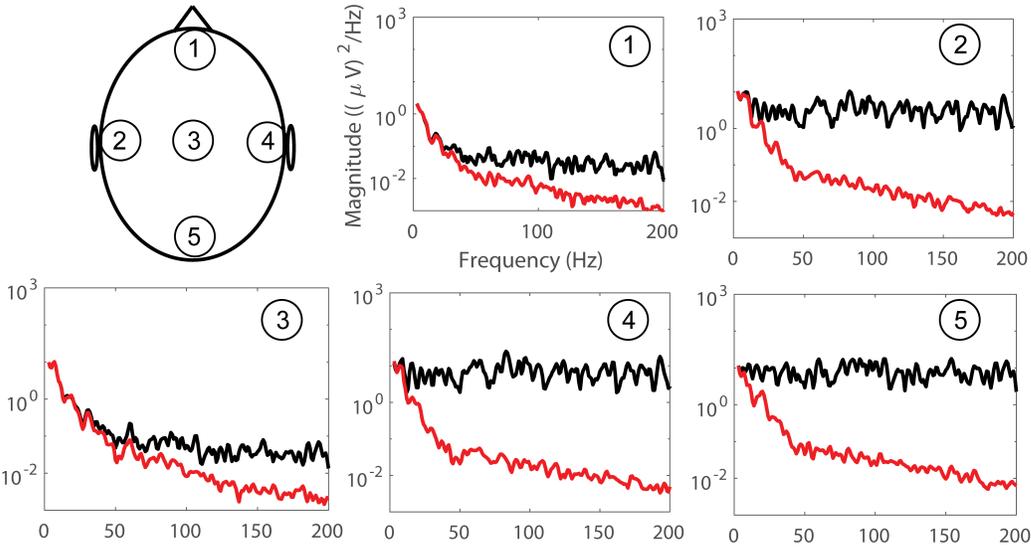
Fig. 3.33:   Power spectral densities of EEG dataset with real EMG contamination at several scalp locations before and after processing by SDBS initialized SPICA algorithm with SDBS post-processing.
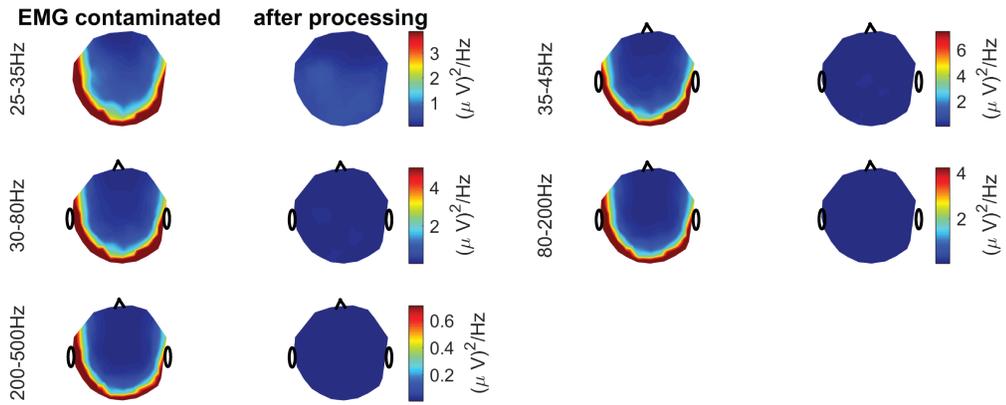

Fig. 3.34:   Topographic maps of power distribution over the scalp in respective spectral bands before and after processing by SDBS initialized SPICA algorithm with SDBS post-processing.
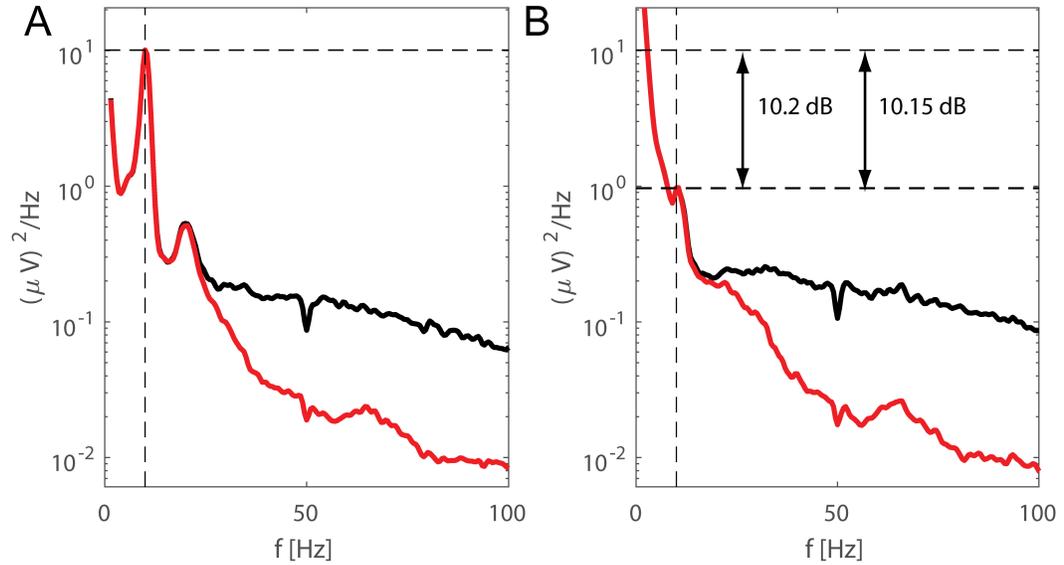
Fig. 3.35: Alpha band power augmentation for closed/open eyes condition. (A) Average PSDs for resting EEG measured during closed eyes. (B) Average PSDs for resting EEG measured during open eyes. Black curves stand for the unprocessed data and the red ones for the data processed by SDBS initialized SPICA algorithm with SDBS post-processing.
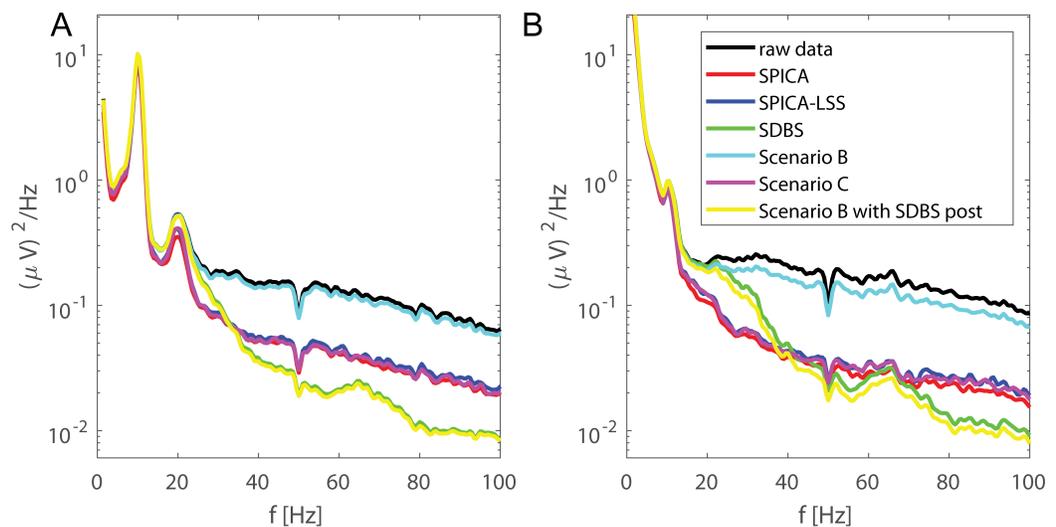
Fig. 3.36: Alpha band power augmentation when comparing EEG epochs under open eyes condition to EEG epochs with closed eyes condition. (A) Average PSDs for resting EEG measured during closed eyes. (B) Average PSDs for resting EEG measured during open eyes.
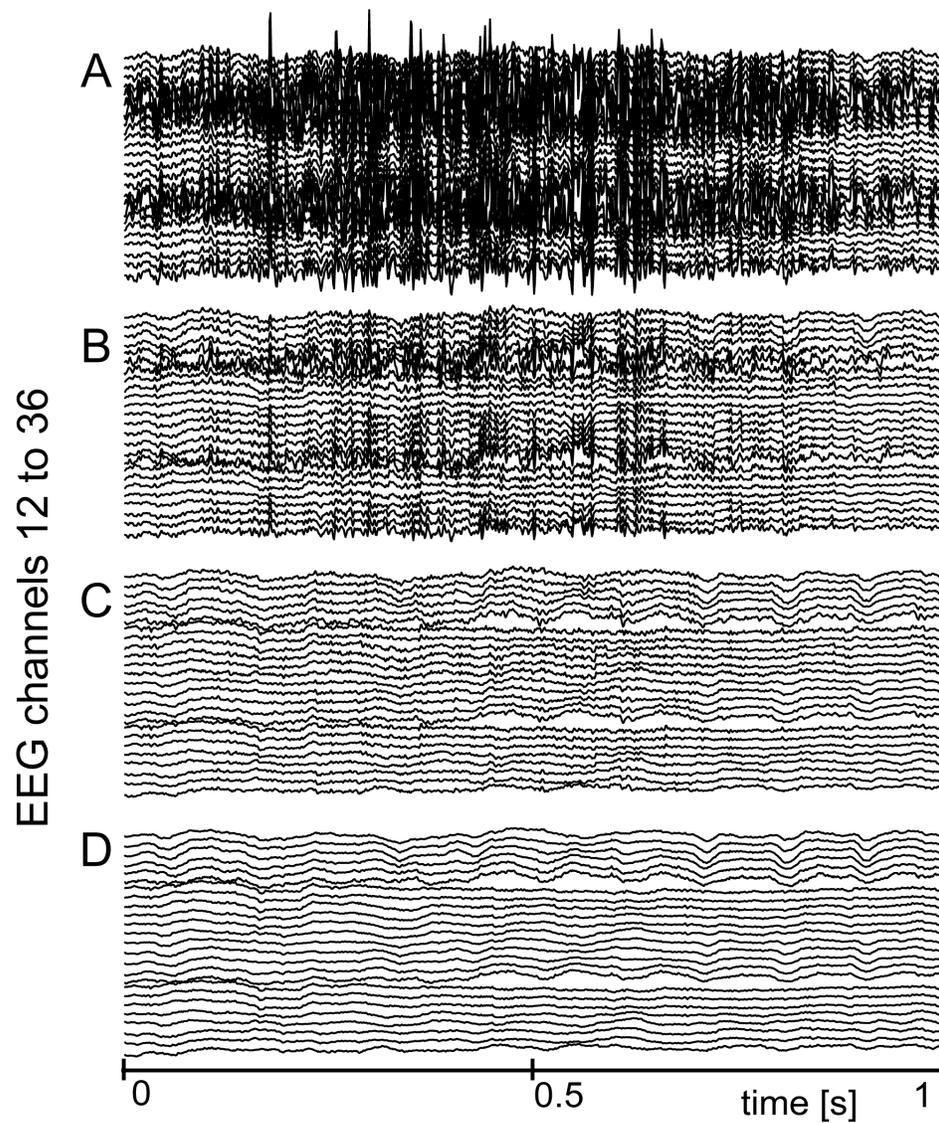
Fig. 3.37: Temporal traces of contaminated EEG segment before and after processing. (A) EEG segment with present muscular artifact. (B) EEG segment processed by SPICA. (C) EEG segment processed by SDBS initialized SPICA. (D) EEG segment processed by SDBS initialized SPICA with SDBS postprocessing.
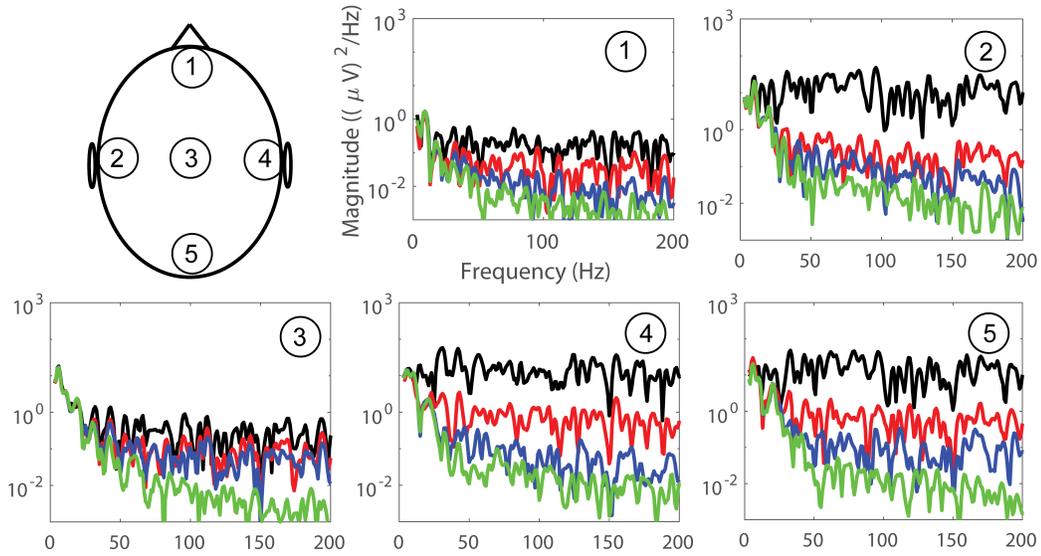
Fig. 3.38: Power spectral densities at several different scalp locations. Black curves stand for unprocessed EEG segment with muscular artifact. Red spectra belong to the EEG segment which was processed by SPICA. Blue curves represent EEG segment processed by SDBS initialized SPICA and green spectral profile is depicted for EEG segment processed by SDBS initialized SPICA with SDBS postprocessing.
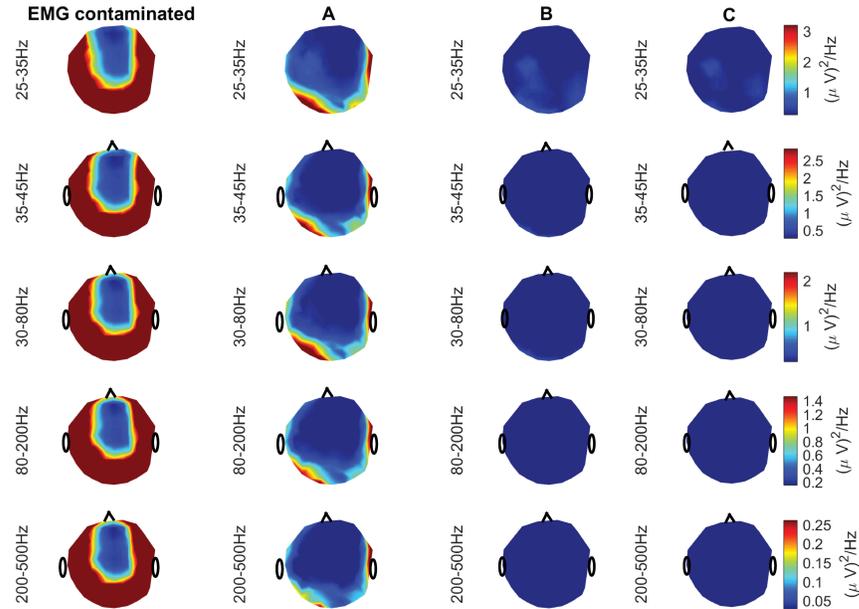
Fig. 3.39: Topographic maps of power distribution over the scalp in respective spectral bands before and after processing by (A) SPICA, (B) SDBS initialized SPICA and (C) SDBS initialized SPICA with SDBS postprocessing.

4. DISCUSSION

4.1 Performance Comparison

The results show that each of the newly proposed methods outperforms all the presented state-of-the-art methods. Furthermore, two algorithms, namely SBDS initialized SPICA with uniform spatial sampling (Section 2.7) and SDBS initialized SPICA with SDBS postprocessing (Section 3.7) and uniform spatial sampling (Subsection 2.4.2), achieved statistically significant superior performance when compared with other proposed algorithms based on data from tables 3.4, 3.5, and 3.6. Specifically, it appears that their ability of source separation and consequently also the reconstruction of EEG data without muscular artifacts is much less affected by the strength of present muscular artifacts than it is in the case of all other presented techniques. This finding can be attributed to the utilized SDBS guided-ICA principle, where ICA is provided with a feasible starting point for iterative estimation of source estimates, where the SDBS objective, which is the projection of the signal mixture in a direction of the strongest EMG artifacts, is met. According to the achieved high performance metric values in Tables 3.4, 3.5, 3.6, the ICA with the conveniently chosen starting point by SDBS also provided better source estimates than in the case of random starting point such as in SPICA or SPICA-LSS algorithm cases. Consequently, smaller number of subsets compared to standard SPICA with the nearest neighborhood sampling strategy was sufficient to achieve much better quality of separation and suppression of present muscular artifacts.

With regards to SDBS initialized SPICA with SDBS postprocessing, two points can be highlighted. Based on the results provided in Sections 3.5 and 3.6, it appears that while the SDBS algorithm does not provide completely perfect separation of muscular artifacts and EEG sources, compared to all other algorithms it manifested the strongest suppression of high frequency components in alpha augmentation experiment data in Fig. 3.21, which might be attributed to tonic muscular artifacts suppression. On the other hand, while SDBS initialized SPICA was able to remove the strong transient muscular artifact leaving only nice looking EEG traces in Fig. 3.24, it appears that this algorithm affords the smallest suppression of high frequency contents when compared with all other algorithms, which is depicted in Fig. 3.27. SDBS initialized SPICA with SDBS postprocessing consequently inherited the advantageous properties of both utilized algorithms.

Selection of various starting points for ICA estimation was also applied in the case of SPICA-LSS algorithm 2.5 and SPICA-LSS cooperation with SDBS 2.7. While the random reinitialization of ICA starting point and subsequent choice of the optimal source set or its rejection in case of imperfect separation had some beneficial impact on the quality of reconstructed signals, this was achieved on expense of higher computation burden when compared to the basic SPICA algorithm. This number was on the other hand bounded to be low (4 in our case) to maintain acceptable overall computation burden. The cooperation with SDBS algorithm as supposed in Section 2.7 for algorithm

SPICA-LSS-SDBS proved to be very useful in situations when no satisfactory separation result could not be found within 4 ICA attempts for respective EEG subspace as SDBS provided at least slightly better EEG and EMG separation in such cases.

The SDBS algorithm, when utilized as a standalone separation algorithm, afforded comparable performance as the basic SPICA. Its inability to achieve better separation results can be attributed to the fact that the sources with the overlapping spectra and a similar power cannot be separated by a PCA approach with the given estimate of the covariance matrix. It is therefore suggested to use it as a preprocessing tool for ICA algorithms, where it showed the significant merit.

The presented results also confirm the weaknesses of the individual state-of-the-art methods pointed out in the introduction Section 1. The PCA based dimensionality reduction (method C) is sometimes able to suppress the muscular contamination, but because all artifactual content cannot be estimated within the drastically lowered dimension space, the reconstruction of a clean EEG record is compromised, which is demonstrated by a low value of average correlation coefficients in case of EEG with muscular artifacts ($\xi = 1, \xi = 4$). 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 any of the newly proposed methods.

4.2   Relation of This Work to Tonic Muscular Contraction Artifacts

Further, some additional points related to the presented methodology need to be discussed. This work concentrates 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 [12]).

It is of course expected that the presented methods will suppress even the tonic artifacts to some extent as they are just weaker and more stationary when compared to transient EMG artifacts. Therefore, as long as the ICA is able to separate them in each EEG subset, they will be identified and removed from the EEG. In fact, the decrease of power on higher frequencies shown in Fig. 3.11, 3.15, 3.21, 3.27, 3.31, and 3.35 can likely be attributed to the suppression of tonic EMG artifacts, and this suppression is similar to the results presented in [12] for vast majority of these methods.

The only exception is SBDS initialized SPICA algorithm which afforded the smallest decrease of power on higher frequencies, which is shown in Fig. 3.27. It appears that while the SDBS algorithm, when used for simultaneous processing of all EEG channels, provided formidable decrease of power on higher frequencies in Fig. 3.21, its ability to achieve the same in EEG subspaces with lower dimension is impaired when used as a preprocessing for FastICA algorithm. However, the combination of SBDS initialized SPICA with SDBS used as a post-processing step seems to be an effective solution, if tonic muscular contamination is of issue. While the SDBS initialized SPICA deals with the strong transient muscular contamination in the first step, the SDBS can then extract remaining tonic muscular contamination from transient EMG-free EEG data in

the second step. The lower separation ability of the SDBS algorithm around the cut-off frequency of the utilized FIR filter, which was mentioned earlier, is of little concern within this scenario, because the separation of components with non negligible spectral contents located around this frequency is handled by the SDBS initialized SPICA in the first step.

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 presented methods are still concerning, the application of any of the presented methods can be followed by the application of a 'classical' EMG removal procedure (as defined in (1.2), (2.2) and (2.3)) that operates on longer EEG records. In this arrangement any of the presented methods 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 any of the presented methods 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 [37, 12]. However, 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.

4.3   Relation of This Work to Other Than Muscular Types of Artifacts

In the motivation and evaluation of newly developed methods, the removal of EMG artifacts was pointed out as the goal of the work, while few comments about other types of artifacts were made. This does not mean that any of the presented methods 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 [20, 18, 21]), and these components can be classified respectively, the algorithm will allow to remove these components without the manifestation of overlearning. This can be directly attributed to algorithms SPICA, and SPICA-LSS.

In the case of other algorithms, where SDBS algorithm is either applied on its own or in cooperation with SPICA, the SDBS algorithm might not provide sufficient separation performance due to the fact that it was derived specifically for the separation of EEG and EMG subspace. However, approach which is similar to SDBS can be derived for other types of artifacts as well. Quick option would be simply to change the utilized filter of SDBS to accommodate the frequency band where the majority of power of the specific type of artifact is expected.

Although this work is primarily targeted to EMG artifacts only, because EMG artifacts are an omnipresent nuisance in a great number of EEG records, to design an algorithm that removes these artifacts alone, is a quite useful contribution.

4.4 Utilized Classifier

The last point to discuss is related to the utilized EMG classifier. Note that the classifier described in Section 2.3 is not the only possible approach. For example, work [12] suggested a classifier based on the slope of a power spectrum. In fact, the presented methods were also retested with the classifier from [12] and virtually identical results were obtained (See Tables 3.1, 3.2, 3.3, 3.4, 3.5, and 3.6). 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.

5. CONCLUSION

This work presents an ICA and PCA based algorithms for the removal of muscular artifacts that are not prone to ICA overlearning for short signals with a high number of measurement channels. This allows them to be applied to short signal segments and to accommodate highly non-stationary nature of artifacts, which can appear, and vanish in a few seconds or less. The algorithms for artifact removal, which make use of the proposed ICA techniques, are able to remove the artifacts present only in short segments to which they are applied. Furthermore, if there are no artifacts detected, the original EEG is not affected. If the composition of artifacts suddenly changes, the algorithms can adapt quickly.

It is demonstrated that the proposed algorithms outperform all the current state-of-the-art approaches, and provide significantly better reconstruction of the original EEG. Of all proposed algorithms, the best reconstruction performance, especially in case of very strong muscular artifacts, was achieved by the SDBS initialized SPICA with spatially uniform EEG subspace sampling with or without a consequent SDBS post processing step.

The algorithms were presented for the purpose of muscular artifact removal; however, it was noted that they can be easily extended for the removal of other types of artifacts and suggestions for such modifications were made.

Because the muscular artifacts are quite common interference of the EEG, it is expected that the SDBS initialized SPICA with or without SDBS post processing will serve as an efficient tool for the reconstruction of artifact-free EEG.

6. IMPROVEMENTS PROVIDED BY THIS WORK

In this work I developed the ICA-based algorithms for suppression of muscular artifacts in short EEG records, which are robust to the ICA overlearning. All the presented algorithms provide better reconstruction of EEG data when compared to the state of the art techniques. Furthermore, the best overall results were achieved when the SDBS initialized SPICA with or without SDBS postprocessing was applied. These two methods showed not only a good suppression of ICA overlearning but also increased separation performance, which was less affected by the power of present muscular artifacts than it was in the case of all other methods.

6.1 Publications of The Author

The SPICA algorithm proposed herein was published in an international peer-reviewed journal [33]. The predecessor of the SDBS algorithm was published in a peer-reviewed domestic journal [32].

Besides the publications in the journals, the provided results were also presented on domestic conferences [27, 29, 30, 31, 28].

Jan Šebek

student FEE CTU

e-mail: sebekja4@fel.cvut.cz

List of published work

Dissertation related work

Impact factor peer-reviewed journals, Quartile Q1

1) Sebek, J.; Bortel R.; Sovka P.: Suppression of overlearning in independent component analysis used for removal of muscular artifacts from electroencephalographic records. In *PloS one.* 2018, vol. 13, no. 8, e0201900.

<div align="right">(70%)</div>

Citations: 0

Peer-reviewed journals

2) Šebek, J.; Bortel, R.: NEW ALGORITHM FOR EEG AND EMG SEPARATION. *Lékař a technika*. 2015, vol. 45, no. 2, p. 43-47. ISSN 0301-5491.   (70%)

Citations: 0

Other conference abstracts

1) Šebek, J.: Možná řešení problému přeučení algoritmu FastICA při zpracování EEG. In III. Letní doktorandské dny 2013. Praha: ČVUT FEL, Katedra teorie obvodů, 2013. ISBN 978-80-01-05251-8.   (100%)

2) Šebek, J.: Classification of the EEG and EMG sources estimated by ICA. In POSTER 2013 - 17th International Student Conference on Electrical Engineering. Prague: Czech Technical University, 2013, ISBN 978-80-01-05242-6.   (100%)

3) Šebek, J.: Metody odstraňování svalových artefaktů ze záznamů EEG. In IV. Letní doktorandské dny 2014. Praha: ČVUT FEL, Katedra teorie obvodů, 2014, díl 4, s. 17-20. ISBN 978-80-01-05506-9.   (100%)

4) Šebek, J.: Removal of Muscular Artifacts from EEG Records. In POSTER 2014 - 18th International Student Conference on Electrical Engineering. Prague: Czech Technical University, 2014, ISBN 978-80-01-05499-4.   (100%)

5) Šebek, J.: New Algorithm for EEG and EMG Separation. In POSTER 2015 - 19th International Student Conference on Electrical Engineering. Prague: Czech Technical University, 2015, ISBN 978-80-01-05728-5   (100%)

Dissertation non-related work

Impact factor peer-reviewed journals, Quartile Q1/Q2

1) Sebek, J.; Curto, S.; Bortel, R.; Prakash, P.: Analysis of minimally invasive directional antennas for microwave tissue ablation. *International Journal of Hyperthermia*. 2017, vol. 33, no. 1, pp. 51-60.   (70%)

   Citations: 2

Impact factor peer-reviewed journals, Quartile Q2

1) Šebek, J.; Albin, N.; Bortel, R.; Natarajan, B.; Prakash, P.: Sensitivity of microwave ablation models to tissue biophysical properties: A first step toward probabilistic modeling and treatment planning. *Medical Physics*. 2016, vol. 43, no. 5, pp. 2649-2661. http://dx.doi.org/10.1118/1.4947482.   (20%)

   Citations: 11

2) Deshazer, G.; Hagmann, M.; Merck, D.; Sebek, J.; Moore, K. B.; Prakash, P.: Computational modeling of 915 MHz microwave ablation: Comparative assessment of temperature-dependent tissue dielectric models. In *Medical physics*. 2017, vol. 44, no. 9, 4859-4868.   (16.67%)

   Citations: 0

Impact factor peer-reviewed journals, Quartile Q4

1) Sebek, J.; Curto, S.; Eaton-Evans J.; et al.: Feasibility Assessment of Microwave Ablation for Treating Esophageal Varices. In ASME. *J. Med. Devices.* 2017, vol. 11, no. 3, ISSN 031013-031013-8. doi:10.1115/1.4037187.   (76%)
   Citations: 0

Other conference abstracts

1) Šebek, J.; Albin, N.; Bortel, R.; Natarajan, B. and Prakash, P.: Sensitivity of microwave ablation models to tissue biophysical properties: application to model-based treatment planning. In: *International Congress of Hyperthermic Oncology*. New Orleans, LA, USA; 2016.   (50%)

2) Šebek, J.; and Prakash, P.: Automatic evaluation of ablation zone boundary: a tool for quantitative evaluation of ex vivo ablation zone maps and comparison against numerical modelling results. In: *International Congress of Hyperthermic Oncology*. New Orleans, LA, USA; 2016.   (70%)

3) Šebek, J.; and Prakash, P.: Analysis of improved microwave applicators for creating directional tissue ablation profiles. In: *World Conference on Interventional Oncology* (*WCIO*). Boston, MA, USA; 2016.   (70%)

4) Fallahi, H.; Šebek, J.; Frattura, E.; Schenck, J.; Prakash, P.: Global microwave endometrial ablation for menorrhagia treatment. In *Proc. of SPIE Vol.* 2017, vol. 10066, pp. 100660K-1.   (20%)

   http://dx.doi.org/10.1117/12.2254097.

5) Sebek, J.; Prakash, P.: Broadband Dielectric Properties of Porcine Lung as a Function of Temperature. Accepted for publication in EUCAP2019 conference and IEEE Xplore.   (85%)

7. ACKNOWLEDGMENT

BIBLIOGRAPHY

[1] P Bofill and M Zibulevsky. Blind separation of more sources than mixtures using sparsity of their short-time fourier transform. In *Proc. ica*, volume 2000, pages 87–92, 2000.

[2] M Borschbach and I Hahn. Separation capability of overcomplete ica approaches. In *Proceedings of the 6th Conference on 6th WSEAS International Conference on Signal Processing*, volume 6, pages 146–151. Citeseer, 2007.

[3] R Bortel and P Sovka. Electrode position scaling in realistic laplacian computation. *IEEE Transactions on Biomedical Engineering*, 55(9):2314–2316, 2008.

[4] R Bortel and P Sovka. Potential approximation in realistic laplacian computation. *Clinical Neurophysiology*, 124(3):462–473, 2013.

[5] R Bortel and P Sovka. Statistical evaluation of coherence estimated from optimally beamformed signals. *Computers in biology and medicine*, 43(9):1286–1292, 2013.

[6] R Cassani, TH Falk, FJ Fraga, PAM Kanda, and R Anghinah. The effects of automated artifact removal algorithms on electroencephalography-based alzheimer's disease diagnosis. *Frontiers in aging neuroscience*, 6, 2014.

[7] NP Castellanos and VA Makarov. Recovering eeg brain signals: artifact suppression with wavelet enhanced independent component analysis. *Journal of neuroscience methods*, 158(2):300–312, 2006.

[8] P Comon. Independent component analysis, a new concept? *Signal processing*, 36(3):287–314, 1994.

[9] M Crespo-Garcia, M Atienza, and JL Cantero. Muscle artifact removal from human sleep eeg by using independent component analysis. *Annals of biomedical engineering*, 36(3):467–475, 2008.

[10] A Delorme, T Sejnowski, and S Makeig. Enhanced detection of artifacts in eeg data using higher-order statistics and independent component analysis. *Neuroimage*, 34(4):1443–1449, 2007.

[11] D Djuwari, DK Kumar, and M Palaniswami. Limitations of ica for artefact removal. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 4685–4688. IEEE, 2006.

[12] SP Fitzgibbon, D DeLosAngeles, TW Lewis, DMW Powers, TS Grummett, EM Whitham, LM Ward, JO Willoughby, and KJ Pope. Automatic determination of emg-contaminated components and validation of independent component analysis using eeg during pharmacologic paralysis. *Clinical Neurophysiology*, 127(3):1781–1793, 2016.

[13] SP Fitzgibbon, D DeLosAngeles, TW Lewis, DMW Powers, EM Whitham, JO Willoughby, and KJ Pope. Surface laplacian of scalp electrical signals and independent component analysis resolve emg contamination of electroencephalogram. *International Journal of Psychophysiology*, 97(3):277–284, 2015.

[14] A Hyvärinen et al. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, 1999.

[15] A Hyvärinen, J Karhunen, and E Oja. *Independent component analysis.* John Wiley & Sons, 2004.

[16] A Hyvärinen and E Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.

[17] A Hyvärinen, J Särelä, and R Vigário. 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)*, pages 425–429, 1999.

[18] TP Jung, S Makeig, C Humphries, TW Lee, MJ Mckeown, V Iragui, and TJ Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.

[19] N Mammone, F La Foresta, and FC Morabito. Automatic artifact rejection from multichannel scalp eeg by wavelet ica. *IEEE Sensors Journal*, 12(3):533–542, 2012.

[20] N Mammone and FC Morabito. Enhanced automatic artifact detection based on independent component analysis and renyi's entropy. *Neural networks*, 21(7):1029–1040, 2008.

[21] SC Ng and P Raveendran. Removal of eog artifacts using ica regression method. In *4th Kuala Lumpur International Conference on Biomedical Engineering 2008*, pages 226–229. Springer, 2008.

[22] SC Ng and P Raveendran. Enhanced μ rhythm extraction using blind source separation and wavelet transform. *IEEE Transactions on Biomedical Engineering*, 56(8):2024–2034, 2009.

[23] F Perrin, O Bertrand, and J Pernier. Scalp current density mapping: value and estimation from potential data. *IEEE Transactions on Biomedical Engineering*, (4):283–288, 1987.

[24] L Ruckay, J Stastny, and P Sovka. Movement-related eeg decomposition using independent component analysis. In *Applied Electronics, 2006. AE 2006. International Conference on*, pages 149–152. IEEE, 2006.

[25] J Särelä and R Vigário. The problem of overlearning in high-order ica approaches: analysis and solutions. In *International Work-Conference on Artificial Neural Networks*, pages 818–825. Springer, 2001.

[26] J Särelä and R Vigário. Overlearning in marginal distribution-based ica: analysis and solutions. *Journal of machine learning research*, 4(Dec):1447–1469, 2003.

[27] J Sebek. Classification of the eeg and emg sources estimated by ica. In *POSTER 2013 - 17th International Student Conference on Electrical Engineering. Prague: Czech Technical University, ISBN 978-80-01-05242-6*, 2013.

[28] J Sebek. Možná řešení problému přeučení algoritmu fastica při zpracování eeg. In *III. Letní doktorandské dny 2013. Praha: ČVUT FEL, Katedra teorie obvodů. ISBN 978-80-01-05251-8*, 2013.

[29] J Sebek. Metody odstraňování svalových artefaktů ze záznamů eeg. In *IV. Letní doktorandské dny 2014. Praha: ČVUT FEL, Katedra teorie obvodů, díl 4, p. 17-20. ISBN 978-80-01-05506-9*, 2014.

[30] J Sebek. Removal of muscular artifacts from eeg records. In *POSTER 2014 - 18th International Student Conference on Electrical Engineering. Prague: Czech Technical University, ISBN 978-80-01-05499-4*, 2014.

[31] J Sebek. New algorithm for eeg and emg separation. In *POSTER 2015 - 19th International Student Conference on Electrical Engineering. Prague: Czech Technical University, ISBN 978-80-01-05728-5*, 2015.

[32] J Šebek and R Bortel. New algorithm for eeg and emg separation. *Lékař a technika-Clinician and Technology*, 45(2):43–47, 2015.

[33] J Sebek, R Bortel, and P Sovka. Suppression of overlearning in independent component analysis used for removal of muscular artifacts from electroencephalographic records. *PloS One*, 13(8):e0201900, 2018.

[34] E Shwedyk, R Balasubramanian, and RN Scott. A nonstationary model for the electromyogram. *IEEE Transactions on Biomedical Engineering*, (5):417–424, 1977.

[35] FJ Theis, EW Lang, and CG Puntonet. A geometric algorithm for overcomplete linear ica. *Neurocomputing*, 56:381–398, 2004.

[36] R Vigário, J Sarela, V Jousmiki, M Hamalainen, and E Oja. Independent component approach to the analysis of eeg and meg recordings. *IEEE transactions on biomedical engineering*, 47(5):589–593, 2000.

[37] EM Whitham, KJ Pope, SP Fitzgibbon, T Lewis, CR Clark, S Loveless, M Broberg, A Wallace, D DeLosAngeles, P Lillie, et al. Scalp electrical recording during paralysis: quantitative evidence that eeg frequencies above 20 hz are contaminated by emg. *Clinical Neurophysiology*, 118(8):1877–1888, 2007.

[38] K Zeng, D Chen, G Ouyang, L Wang, X Liu, and X Li. An eemd-ica approach to enhancing artifact rejection for noisy multivariate neural data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(6):630–638, 2016.