Author(s): Hernández-Pavón JC, Metsomaa J, Mutanen T, Stenroos M, Mäki H, Ilmoniemi RJ, Sarvas J.

Title: Uncovering neural independent components from highly artifactual TMS-evoked EEG data

Year: 2012

Version: Post print

Please cite the original version:


This publication is included in the electronic version of the article dissertation:
Hernández-Pavón, Julio César. Transcranial magnetic stimulation and EEG in studies of brain function. Aalto University publication series DOCTORAL DISSERTATIONS, 103/2015.
Title:

Uncovering neural independent components from highly artifactual TMS-evoked EEG data

Author names and affiliations:

Julio C. Hernandez-Pavon\textsuperscript{1-3*}, Johanna Metsomaa\textsuperscript{1-2*}, Tuomas Mutanen\textsuperscript{1-2}, Matti Stenroos\textsuperscript{1}, Hanna Mäki\textsuperscript{1-2}, Risto J. Ilmoniemi\textsuperscript{1-2}, Jukka Sarvas\textsuperscript{1}

\textsuperscript{1}Department of Biomedical Engineering and Computational Science (BECS), Aalto University, School of Science, P.O. Box 12200, FI-00076 Aalto, Espoo, Finland.

\textsuperscript{2}BioMag Laboratory, HUSLAB, Helsinki University Central Hospital, P.O. Box 340, FI-00029 HUS, Helsinki, Finland.

\textsuperscript{3}Department of Physical Engineering, University of Guanajuato, Loma del Bosque 103. Lomas del Campestre 37150, Leon, Guanajuato, Mexico.

* These authors have equally contributed to this work.

Corresponding Author:

Julio Cesar Hernandez Pavon

Department of Biomedical Engineering and Computational Science (BECS), Aalto University, School of Science, P.O. Box 12200, FI-00076 Aalto, Espoo, Finland.

e-mails: julio.hpavon@aalto.fi; julio.hpavon@gmail.com

Tel: +358 503 443 470
Fax: +358 947 023 182

Number of text pages: 37
Number of figures: 7
Number of tables: 3
Abstract

Transcranial magnetic stimulation (TMS) combined with electroencephalography (EEG) is a powerful tool for studying cortical excitability and connectivity. To enhance the EEG interpretation, independent component analysis (ICA) has been used to separate the data into independent components (ICs). However, TMS can evoke large artifacts in EEG, which may greatly distort the ICA separation. The removal of such artifactual EEG from the data is a difficult task. In this paper we study how badly the large artifacts distort the ICA separation, and whether the distortions could be avoided without removing the artifacts. We first show that in the ICA separation the time courses of the ICs are not affected by the large artifacts, but their topographies could be greatly distorted. Next, we show how this distortion can be circumvented. We introduce a novel technique of suppression, by which the EEG data are modified so that the ICA separation of the suppressed data becomes reliable. The suppression, instead of removing the artifactual EEG, rescales all the data to about the same magnitude as the neural EEG. For the suppressed data, ICA returns the original time courses, but instead of the original topographies, it returns modified ones, which can be used, e.g., for the source localization. We present three suppression methods based on principal component analysis, wavelet analysis, and whitening of the data matrix, respectively. We test the methods with numerical simulations. The results show that the suppression improves the source localization.

Keywords: Transcranial magnetic stimulation; Electroencephalography; Independent component analysis; Principal component analysis; Wavelets; Dipole source localization; Broca’s area.
1. Introduction

Transcranial magnetic stimulation (TMS) is a noninvasive technique that activates the brain by means of a brief and strong magnetic pulse (Barker et al., 1985). When TMS is combined with electroencephalography (EEG) it becomes a powerful tool for studying the cortical excitability and connectivity (Ilmoniemi et al., 1997). TMS–EEG has successfully been used for studying brain areas where evoked EEG signals are not affected by large TMS-induced artifacts (Komssi et al., 2002; Kähkönen et al., 2004, 2005; Massimini et al., 2005; Silvanto and Cattaneo 2010).

Independent component analysis (ICA) has been used with EEG to find the underlying (hidden) independent components by separating the measured data matrix into two new matrices; a mixing matrix, whose columns are the topographies of the hidden sources, and a time course (waveform) matrix, whose rows are time-dependent amplitudes of the sources (Hyvärinen and Oja 2000; Vigário et al., 2000, Onton et al., 2006). A successful ICA separation can help the analysis of EEG data. The time courses tell when the sources are active. The mixing matrix is the basis for localization of the hidden sources (Hild and Nagarajan 2009; Cao and Slobounov 2010; Ventouras et al., 2010). If a source is known to lie on the cortex, then finding its spatial distribution by its topography is a rather well-posed inverse problem. If the source can be identified to be dipolar, then its localization becomes a search for a single dipolar source, which is an even better-posed, robust inverse problem.

In this paper we study the problem of applying ICA to TMS–EEG data contaminated by large muscle artifacts, in order to separate and study the neural independent components in the data. Recently, ICA has been used to identify and remove the artifact components from the TMS-evoked EEG data where the artifacts have been of moderate size (Iwahashi et al., 2008; Hamidi et al., 2010), or to remove large muscle artifacts due to TMS stimulation of Broca’s area (Korhonen et al., 2011). However, an adequate removal of very large artifacts, e.g., 2 or 3 orders of magnitude larger than the neural EEG signals, is a very difficult task. Namely, either after the removal, the remnants of artifacts may still be of the magnitude of the neural EEG, or the removal procedure may remove too much neural EEG from the data. In both cases, the corrected data is hardly good for any further EEG analysis. Because the removal is so difficult, the question arises how badly the large artifacts actually distort the ICA separation, especially that of the neural components, and whether the possible distortions could be avoided without trying to remove the artifacts. We studied this question and got the following two main results.

As the first result we show that, in the ICA separation the time courses of the hidden components are not affected by the large artifacts, and it is only the topographies that are distorted in the separation, especially the neural ones. As the second main result, we show how this distortion can be circumvented. Namely, we introduce a preprocessing method of the measured data, here called the suppression method, with which the data is modified so that the time courses remain unchanged, but the topographies are, in a
controlled way, modified so that the ICA separation of the modified data matrix becomes reliable. The idea in the suppression method is that, instead of removing the artifactual EEG, we suppress the data to about the same magnitude as the neural EEG and so that the neural part changes as little as possible. Thereafter, for the suppressed data, ICA still returns the original time courses, but instead of the original topographies, it returns the modified ones, which can be used in source localization. In this paper, we introduce three suppression methods based on principal component analysis, wavelet analysis, and whitening of the data matrix, respectively. In our study we use FastICA (Hyvärinen 1999; Hyvärinen and Oja 2000) as the numerical ICA algorithm, but we will show how the results can be applied to other ICA algorithms, as well.

The main application of the ICA found topographies is source localization. However, it becomes very inaccurate, if the EEG data contain large artifacts. This can be corrected by using the suppressed data and the ICA found modified topographies. We test this with data where simulated EEG is mixed with measured TMS-evoked EEG data from Broca’s area stimulation. We show that the proposed suppression methods significantly improve the localizing accuracy. The methods presented in this article open new possibilities for studying lateral areas of the brain with TMS–EEG.

2. Material and methods

2.1. Theoretical methods

2.1.1. Notation

For an $m \times n$ matrix $A$, we let $A(i,j)$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, denote the elements (entries) of $A$. The rank of $A$ is denoted by $\text{rank}(A)$ and the transpose of $A$ by $A^T$. An $m \times n$ diagonal matrix with diagonal elements $d_1, \ldots, d_p$, $p = \min(m, n)$, is denoted by $\text{diag}_{m \times n}(d_1, \ldots, d_p)$. For an $m \times n$ matrix $A$, we let $A(:,j)$ and $A(i,:)$ denote the $j$:th column and $i$:th row, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, respectively.

If $u_j$, $j = 1, \ldots, n$, are $m$-vectors, we let $[u_1, \ldots, u_n]$ denote the $m \times n$ matrix with columns $u_j$. For an index subset $J \subset \{1, 2, \ldots, n\}$, we denote the submatrix of $A$ with columns $A(:,j)$, $j \in J$, by $A(:,J)$. For integers $p < q$, the index set $\{p, p + 1, \ldots, q\}$ is briefly denoted by $p : q$; for instance, $A(:,p : q) = [A(:,p), A(:,p + 1), \ldots, A(:,q)]$.

2.1.2. Preprocessing

After the data have been averaged, and the bad channels (e.g., those with poor electrode contact) have been removed, the measured data consist of an $M \times N$ matrix $X_1$, with $M$ channels and $N$ time points. Next, the zero potential level is set at each time point by setting the average over channels equal to zero, yielding the data matrix $X_2$. 
Before applying FastICA, the matrix $X_2$ should be centered, compressed and whitened (see Appendix A for details). The centered matrix $X_3$ is obtained from $X_2$ by time centering it, i.e., by setting the time average of each row equal to zero. Thereafter, the centered data $X_3$, is compressed by letting the compressed matrix $X$ to consist of the $m$ leading principal components (PCs) of $X_3$, with $m \leq \text{rank}(X_3) \leq M$. Before applying FastICA, the whitening step must be performed on the compressed data $X$, yielding the whitened data $X_{\text{white}}$. As a result of whitening, the covariance matrix of $X_{\text{white}}$ is the identity matrix.

2.1.3. Independent Component Analysis

We assume that the compressed data matrix $X$ is due to $m$ latent (hidden) independent components so that

$$X = AS,$$

where $A$ is the $M \times m$ mixing matrix, with $\text{rank}(A) = m$, and $S$ is the $m \times N$ matrix so that $S(j, t), t = 1, \ldots, N$, is the time course (waveform) of the $j$th source. The columns of $A(:,j)$ of the mixing matrix are the topographies of the latent sources. The vectors $S(:, t), t = 1, \ldots, N$, are thought to be random samples of a random vector variable $s = [s_1, \ldots, s_m]^T$ whose components $s_i$ are statistically independent with unit variance. Accordingly, $S$ is normalized so that the covariance matrix

$$\text{Cov}(S) = \frac{1}{N}SS^T = I$$

is the $m \times m$ identity matrix $I$. The task of an ICA algorithm, like FastICA, is to find estimates for $A$ and $S$, respectively, see Appendix A.

2.1.4. Estimation errors in $A$ and $S$ due to large artifacts

Ideally, ICA would find the correct estimates $\hat{A} = A$ and $\hat{S} = S$ for the input data $X = AS$. In practice, the ICA algorithms produce some errors in the estimates both due to limitations in the algorithms and noise in the data. In this section, we show that, at the presence of large artifacts, the estimation errors in the waveforms $\hat{S}$ are not increased as compared to artifact-free data. However, the estimated topographies in $\hat{A}$ have significantly larger errors than in artifact-free data, which causes serious bias in interpreting the results.

The characteristic feature of the large artifact components is their amplitudes, which are substantially larger than those of the neural components. In the decomposition $X = AS$, this means that the artifact topographies in $\hat{A}$ are larger in magnitude than the neural ones.

However, as discussed in the previous section, the ICA algorithms are based on making the separation of components based on the independence properties of the time courses
S. In the case of artifacts, the large magnitude of some components does not affect the independence in S, but merely causes some topographies in A to grow in size. Hence, the estimation of the time courses is unaffected due to artifactual components.

To further examine the effect of artifacts, if the data X is whitened, yielding the whitened data matrix $X_{\text{white}}$, then $X_{\text{white}} = WS$, where W is the new, ‘whitened’, orthogonal mixing matrix. In the whitened mixing matrix, all the topographies are of the same size, which eliminates the possible problem of large artifacts. Hence, applying ICA now to the whitened data $X_{\text{white}}$, as one does with FastICA and what can be done with any ICA algorithm, then the algorithm returns the estimate $\hat{W}$ and $\hat{S}$ for W and S with good accuracy (i.e., as good as with artifact-free data). The estimated time courses for the original, non-whitened data X are the same as for the whitened one, only the topographies are different. In that case the dewhiteing matrix, in Appendix A, needs to be applied on $\hat{W}$ to obtain the estimate for the original mixing matrix $A$. Therefore, the estimate $\hat{S}$ is not affected by large artifacts (see Appendix B for a more rigorous reasoning).

Though the large artifactual signals in X do not affect the estimated $\hat{S}$, they usually increase the relative error in the estimated topographies $\hat{A}$, particularly in the neural topographies, which are of our main interest. This can be illustrated by the following equation, using (1) and (2)

$$\hat{A} = \frac{1}{N}X\hat{S}^T.$$  

(3)

Here even small errors in $\hat{S}$, with the large artifactual EEG values in X, will cause large errors in the product $X\hat{S}$, which lead to large relative errors in the small elements $A(k,j)$.

The relative error $\rho_j$ in $\hat{A}(\cdot,j)$ can be computed as follows (see Appendix C)

$$\rho_j = \frac{\sqrt{\frac{1}{N} \text{trace}(XX^T)} \sigma}{\|\hat{A}(\cdot,j)\|} = \frac{\sqrt{\sum_{k=1}^{m} \|\hat{A}(\cdot,k)\|^2}}{\|\hat{A}(\cdot,j)\|} \sigma,$$

(4)

for all $j = 1, \ldots, m$, were $\sigma$ would be the average error in the mixing matrix element if no artifacts were present in the data. The formula shows that the error is large for those components that are small compared to other components or signal values in X. This confirms the above heuristic reasoning that relative errors are expected to be high for the neural topographies if large artifact components are present in the data.

2.1.5. Suppressing high EEG signal values in EEG data

The previous section shows that the highly artifactual EEG components, though not affecting the time courses in the FastICA outcome, greatly distort the estimated mixing matrix. An obvious remedy to that distortion is to sufficiently suppress the high EEG signal values in the data, in a controlled way, so that no information is lost and the
suppression is directed mostly to the artifactual data. In practise, such a suppression is carried out with an appropriately chosen \( L \times M \) suppression matrix \( P \) so that \( \tilde{X} = PX \) is the suppressed data matrix, where \( X \) is our preprocessed \( M \times N \) data matrix.

In order to make the suppression matrix information preserving, we require that \( P \) is rank-preserving (with respect to \( X \)) in the sense that
\[
\text{rank}(PX) = \text{rank}(X).
\]
Namely, if (5) holds, then we can show that there is a desuppression matrix \( Q \) so that
\[
X = Q(PX),
\]
i.e., \( X \) can be recovered from the suppressed data \( \tilde{X} = PX \), and thus no information is lost in the suppression. In fact, \( Q \) is given by
\[
Q = U(:,1:m) \text{pinv}(PU(:,1:m)),
\]
where \( m = \text{rank}(X) \), \( U \) is the orthogonal matrix in the singular value decomposition (SVD) \( X = UDV^T \) and \( \text{pinv}(B) \) stands for the Moore–Penrose pseudo-inverse of matrix \( B \).

Again, since the suppression only modifies the mixing matrix, \( PX = PAS \), the estimation of \( S \) is not affected. However, the suppression should be such that the relative error (4), i.e., the high EEG values and artifact components decrease, while the magnitude of the denominator, i.e., the neural component, remains about the same.

As the suppression \( P \) only affects the estimated mixing matrix, it can be applied either on data \( X \), as \( PX \), before ICA, or on the estimated \( \hat{A} \), as \( P\hat{A} \), after running ICA on the original \( X \). It still remains to find appropriate suppression matrices \( P \). We will do that in the next section, where we introduce three suppression methods and discuss their properties.

Before moving on to the derivation of the suppression methods, we present the workflow that describes how the data analysis, the suppression, and source localization are performed (see section 2.3 for source localization details).

1. The data \( X \) is assumed to be composed of independent components \( X = AS \).
2. Perform the preprocessing as required for the chosen ICA algorithm.
3. Compute the suppression matrix \( P \) based on the data and using the chosen suppression method.
4. Perform the ICA on \( X \), and obtain the estimates \( \hat{A} \) and \( \hat{S} \), and the suppressed estimated topographies \( \tilde{\hat{A}} = P\hat{A} \), respectively.
5. Choose an interesting component index \( i \) based on the estimate waveform, \( \hat{S}(i,:) \). (Do not rely on the topographies, because the topographies can be non-interpretable by inspection due to the artifacts and the suppression).
6. Pick the corresponding estimated suppressed topography $\hat{\mathbf{A}}_{\cdot i}$.
7. Use your preferred source localization strategy to identify the active areas based on the the suppressed topography.
8. Repeat steps 5–7 for all interesting components.

2.2. Suppression Methods

A good suppression should have the following properties: (a) it is rank-preserving, (b) it suppresses more the artifactual EEG than the neural EEG so that the relative error in the neural topographies is lowered, (c) the neural EEG data is changed as little as possible in the suppression. In the following, we introduce three rank-preserving suppressions, which are formed in various heuristic ways so that also properties (b) and (c) would be fulfilled satisfactorily.

We propose three suppression methods and two first ones are of the following general structure. Based on the analyzable data $\mathbf{X}$, another data matrix $\mathbf{Y}$ is computed. The idea is to construct $\mathbf{Y}$ in such a way that, in particular, the artifacts are well presented by it. Some prior knowledge of the artifact properties is used in building $\mathbf{Y}$.

After $\mathbf{Y}$ has been constructed, let $\mathbf{Y} = \mathbf{UDV}^T$ be the SVD of $\mathbf{Y}$. It follows that the unit vectors $\mathbf{u}_j = \mathbf{U}_{\cdot j}$, $j = 1, \ldots, M$, form an orthonormal basis in the $M$-space, so that

$$\mathbf{X} = \sum_{j=1}^{M} \mathbf{u}_j \mathbf{u}_j^T \mathbf{X}. \quad (8)$$

We form the suppression $\mathbf{P}$ by rescaling the first $n_s$ terms $\mathbf{u}_j \mathbf{u}_j^T \mathbf{X}$ in (8) to $\alpha_j \mathbf{u}_j \mathbf{u}_j^T \mathbf{X}$, where $\alpha$ is the median size of the remaining non-zero terms. Accordingly, $\alpha_j = \alpha/\|\mathbf{u}_j^T \mathbf{X}\|$ with $\alpha$ being the median of the non-zero $\|\mathbf{u}_k^T \mathbf{X}\|$, $n_s + 1 \leq k \leq M$. If $\|\mathbf{u}_j^T \mathbf{X}\| \gg \alpha$, this rescaling strongly suppresses the component $\mathbf{u}_j \mathbf{u}_j^T \mathbf{X}$ and the artifacts in it.

The suppression matrix is given by

$$\mathbf{P} = \sum_{j=1}^{n_s} \alpha_j \mathbf{u}_j \mathbf{u}_j^T + \sum_{j=n_s+1}^{M} \mathbf{u}_j \mathbf{u}_j^T = \mathbf{U} \text{diag}_{M \times M}(\alpha_1, \ldots, \alpha_{n_s}, 1, \ldots, 1) \mathbf{U}^T, \quad (9)$$

where $\alpha_j = \alpha/\|\mathbf{u}_j^T \mathbf{X}\|$, $j = 1, \ldots, n_s$, with $\alpha$ as above. In the following, we suggest three ways of constructing the submatrix $\mathbf{Y}$ and determining the number of suppressed components $n_s$.

2.2.1. Principal components suppression

If the artifacts are restricted to a certain time interval, the data $\mathbf{Y}$ can simply be chosen as a submatrix of $\mathbf{X}$ corresponding to that time interval. In TMS–EEG data,
the artifacts dominate in the beginning of the trial. Consequently, consider the data submatrix

$$Y = X(:, t_1 : t_2)$$

(10)

where the data are restricted to time indices from $t_1$ to $t_2$. The index $t_1$ is the starting index of the EEG analysis, and $t_2$ is chosen so that most of the artifacts after $t_1$ occur before $t_2$.

In (8), a small number of leading components $u_j u_j^T X$, $j = 1, 2, \ldots$, are expected to contain most of the large artifacts. Our choice of the number $n_s$ of the suppressed components was based on how $n_s$ affects the numerator in (4), with $X$ replaced by $PX$, $E_s = \sqrt{1/N \text{trace}(PX(PX)^T)}$. We plotted the variable $E_s$ for various $n_s$ and chose the $n_s$ at the kink of the curve, at which point the decline turned significantly slower. As we show later, with an appropriate choice of $t_2$, a very low $n_s$ can be chosen, say $n_s \leq 4$, and we obtain an efficient suppression, which also changes the data quite little because most of the components $u_j u_j^T X$ remain unchanged.

The first $u_j$’s mainly span the artifactual EEG patterns. For physical reasons the potential patterns of muscular artifact sources in the scalp are different from the potential patterns due to neuronal sources in the brain. Therefore, it is expected that by suppressing only a few leading terms $u_j u_j^T X$ by (9), the neural EEG is suppressed much less than the artifactual EEG.

2.2.2. Wavelet Suppression

In the work by Mäki and Ilmoniemi (2011), the subdata $Y$ was formed by high-pass filtering the data $X$ by the Fourier transform in time with the cut-off frequency 100 Hz, above which little neural signals appear. Thereafter, the artifacts were removed from $X$ by projecting out the components $u_j u_j^T X$, $j = 1, \ldots, m_s$, where $u_j$ were obtained from the SVD of $Y$ as above. Our wavelet suppression uses a similar idea but the filtering is based on the magnitude of the wavelet coefficients as in (Nazareth et al., 2006). In addition, the prominent components in $Y$ are rather rescaled than completely projected out.

The sampled signal measured by the $q^{\text{th}}$ channel can be written in terms of a discrete wavelet expansion as follows

$$x_q(t) = \varphi_{n,0}(t) + \sum_{j=1}^{n} \sum_{k=0}^{2^n-j} \lambda_{j,k}^{(q)} \psi_{j,k}(t),$$

(11)

where $t$ indices the time, $\varphi_{n,0}$ is the level $n$ scaling basis function, $\psi_{j,k}$ are wavelet basis functions

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k),$$

(12)

and $\lambda_{j,k}^{(q)}$ are the wavelet coefficients. We used both the Haar wavelets and the Daubechies
4-wavelets (Daubechies, 1992), which have been considered appropriate for EEG time series (Adeli et al., 2003).

Taken into account all channels \( q \), the wavelet expansion for the whole data can be written in the vector form as

\[
X(:, t + t_1 - 1) = \Phi_{n, 0} + \sum_{j=1}^{n} \sum_{k=0}^{2^{n-j} - 1} \Lambda_{j,k} \psi_{j,k}(t), \quad t = 1, 2, \ldots, t_2 - t_1 + 1,
\]

where \( \Phi_{n, 0} = [\varphi_{n, 0}^{(1)}, \varphi_{n, 0}^{(2)}, \ldots, \varphi_{n, 0}^{(M)}]^T \) and \( \Lambda_{j,k} = [\lambda_{j,k}^{(1)}, \lambda_{j,k}^{(2)}, \ldots, \lambda_{j,k}^{(M)}]^T \). The artifactual data is presumably represented by those terms of the expansion whose coefficients \( \lambda_{j,k}^{(q)} \) are large (Nazareth et al., 2006).

In (13), we define the size \( m(j, k) \) of the \((j, k)\) summation term to be

\[
m(j, k) = \max_{1 \leq q \leq M} \left( 2^{-j/2} |\lambda_{j,k}^{(q)}| \right),
\]

since \( \max_t |\lambda_{j,k}^{(q)} \psi_{j,k}(t)| = 2^{-j/2} |\lambda_{j,k}^{(q)}| \). For \( \Phi_{n, 0} \), the size is \( m_0 = \max_{1 \leq q \leq M} |\varphi_{n, 0}^{(q)}| \).

Thereafter, we order all the summation terms in (13) according to the descending size and keep only the \( n_s \) first terms. Let the sum of those \( n_s \) terms be \( \tilde{X}(:, t) \). Next, we define \( Y \) as \( M \times (t_2 - t_1 + 1) \) matrix with columns

\[
Y(:, t) = \tilde{X}(:, t), \quad t = 1, \ldots, t_2 - t_1 + 1.
\]

The size of the parameter \( n_s \) was chosen as in PC suppression.

2.2.3. Whitening as a suppression method

The whitening is one of the preprocessing steps for FastICA. On the other hand, it can also be used as a suppression method, and the \( M \times N \) whitening matrix \( B \), see Appendix A, can be considered as a suppression matrix \( P \), which is rank-preserving. The whitened data \( X_{\text{white}} = PX \) is the suppressed data, and the weight matrix (‘whitened’ mixing matrix) \( W = PA \) is also the suppressed mixing matrix. Because of this double role of \( W \), no additional error is produced in obtaining the suppressed mixing matrix from the weight matrix. Furthermore, as the whitening scales all the topographies to the same magnitude, and so the large artifact topographies do not affect the estimated suppressed topographies, as can be seen by (4) as \( \tilde{X} \) and \( \tilde{A} \) are replaced by \( X_{\text{white}} \) and \( \tilde{W} \). In this sense, whitening is an optimal suppression. By (4), the relative error in whitened topographies is \( \sqrt{m} \sigma \) with \( m = \text{rank}(A) \), a value which should be used as a reference when relative errors of other suppression methods are considered.

On the other hand, the whitening changes the data strongly. Namely, we can also write the whitened data as \( X_{\text{white}} = U(:, 1 : M)^T \sum_{j=0}^{m} \frac{\lambda_{j}}{\sigma_j} u_j u_j^T X \) (see Appendix A). So, we see that all components \( u_j u_j^T X \), in the expansion \( X = \sum_{j=0}^{m} u_j u_j^T X \), have been
rescaled to the same magnitude $\sqrt{N}$. Accordingly, in the sense of changing the data as little as possible, whitening is not an optimal suppression, and it is not obvious whether the whitened data could be used in the EEG analysis in any other way than as a technical input in the ICA analysis. However, whitening-suppressed data turns out to be very useful in localizing dipolar IC sources, as we later show.

2.3. Measured data, simulated data and source localization

So far, the theory of the suppression has been discussed. Especially, we have shown that the suppression is needed to circumvent the errors in the mixing matrix estimated by ICA due to large artifacts. In fact, the suppression modifies the mixing matrix, and it is the modified one that ICA estimates. This raises the question whether there is any practical use for the modified mixing matrix (or is reasonably good estimate). The answer is that the use is in the localization of the latent sources, and in fact, the suppression improves significantly the localization accuracy. This is due to the fact that the relative errors in the estimated modified topographies are lower than in the estimated topographies without suppression.

We want to assess this hypothesis of improved localization accuracy with simulated data. In order to make the demonstration realistic we have constructed the data $X_{dem}$ for the demonstration by mixing measured EEG data, which contain large artifacts, with simulated EEG data of one cortical dipole, with given position and moment. Thereafter, we have applied FastICA to both $X_{dem}$ and the suppressed $X_{dem}$, and obtained two estimates for the topography of the dipole (hidden to the ICA algorithm). Next, the position of the dipole is estimated by a dipole search with the estimated topographies. Finally, comparing the found location with the correct one, the localizing accuracies without and with the suppression are found. In the next section we describe in detail how the measurement data was recorded. In the section after that, we explain how the data $X_{dem}$ is constructed and how the source localization and accuracy assessment is carried out.

We complete this section with a short note on the dipole search and how the suppression modifies it. Let $F$ be a given field solver, which computes the EEG data $X = F(r, q)$ of a given dipole with position $r$ and moment $q$. Then if only $X$ is known, the unknown dipole $(r, q)$ is found in the dipole search method by finding (in one way or another) the pair $(r, q)$ which yields the best matching $F(r, q)$ with $X$. Now, if the data is suppressed by a suppression matrix $P$, then the dipole search with the data $PX$ is carried out similarly by replacing $X$ and $F(r, q)$ by $PX$ and $PF(r, q)$.

2.3.1. Measured data

The data were recorded from one healthy subject. The TMS pulses were delivered with the Nexstim eXimia TMS stimulator (Nexstim Ltd., Helsinki, Finland) and a figure-of-8 coil. The stimulation was targeted on Broca’s area with an MRI-guided navigation.
system (Nexstim eXimia NBS), see Fig. 1 (A). The subject received 110 stimuli at 100% of the motor threshold (MT) on Broca’s area. The TMS pulses were delivered at random intervals of 1.5–3.5 s. A 60-channel TMS-compatible Nexstim eXimia EEG device recorded the EEG. The signals were band-pass filtered from 0.1 to 350 Hz and digitized at 1450 Hz. Data analysis was performed on Matlab (The Mathworks, Inc., Natick, Massachusetts, USA). The evoked potentials were averaged over trials.

Broca’s area is an important language area located close to the temporal muscle. In addition to the temporal muscle, also other muscles, such as the masseter muscle can become activated. Since stimulation of Broca’s area gives rise to very large artifacts, the data analysis in the present study was restricted from 10 to 250 ms and from 15 to 250 ms. Responses earlier than 10 ms are very strongly contaminated by the artifacts, see Figs. 1 (B), (C), and (D).

2.3.2. Generating dipolar topographies with a BEM potential field solver

For dipole search, the possible locations of the EEG sources, defining the source space, were constrained to the cerebral cortex. The source space was discretized from the subject’s MRI using the FreeSurfer software (Dale et al., 1999; Fischl et al., 1999). In total, the source space comprised 5124 dipolar sources, whose locations were collected in the \((3 \times 5124)\) matrix \(M\), whose columns contained the cartesian coordinates of the sources. The orientations of the dipoles were constrained perpendicular to the cortex surface. The skull and scalp surfaces were discretized using BrainSuite (Shattuck and Leahy 2002).

Given the dipole locations and orientations together with scalp and skull surfaces, the boundary-element method (BEM) was used to compute the \((M \times 5124)\) lead-field matrix \(L\) whose each column \(L(:, i)\) is the topography caused by the unit dipole at the source space location \(M(:, i)\). The BEM field solver was built with a 3-shell volume conductor model and linear Garlerkin boundary-elements, and formulated with the isolated source approach. The conductivity-contrast between the soft tissues and skull was 40 (Hämäläinen and Sarvas 1989; Stenroos et al., 2007).

2.3.3. Simulated data

The building of the data \(X_{\text{dem}}\) was started by forming a simulated data matrix \(X_{\text{sim}}\) for a simulated dipolar source as

\[
X_{\text{sim}} = \lambda a_{\text{sim}} s_{\text{sim}},
\]

where \(a_{\text{sim}}\) is the \(M \times 1\) simulated dipolar topography, \(s_{\text{sim}}\) the \(1 \times N\) simulated waveform, and \(\lambda\) a factor describing the magnitude of the component. The simulated data were added to the measured data matrix \(X_1\), yielding the data matrix \(X_{\text{dem}}\),

\[
X_{\text{dem}} = X_1 + X_{\text{sim}}.
\]
The topography \( a_{\text{sim}} \) was chosen to be the lead-field matrix column \( L(:, i_d) \) corresponding to a cortical dipole at position \( M(:, i_d) \), i.e.,

\[
a_{\text{sim}} = L(:, i_d).
\] (18)

Three different dipole positions were chosen in the following areas: the left Broca’s area (close to the stimulation target), the left Wernicke’s area, and a position in the right hemisphere corresponding to Broca’s area.

The simulated waveform, \( s_{\text{sim}} \), had the form of the Mexican hat function, which has zero-mean in time. The duration of the waveform was defined as the time interval of the non-zero part of the signal. The simulated waveform was studied in two time intervals. In the first, difficult case, the duration was from 10 to 35 ms. In the second, easy case, it was from 15 to 40 ms. See Fig. 2 for illustrations of the simulated waveform and topographies. Factor \( \lambda \) was adjusted so that the maximum absolute amplitude among all channels was 8 \( \mu V \).

Thereafter, the analysis of \( X_{\text{dem}} \) was carried out as described section 2.1.5, and estimates \( \hat{S} \) and \( \hat{A} \) were obtained. Among the found independent components the one, with index \( j_s \), corresponding to the simulated dipole was identified by comparing the waveforms \( \hat{S}(j,:) \) to \( s_{\text{sim}} \). Note that neither the artifacts nor the suppression affected \( \hat{S} \), and therefore they do not affect the identification. More precisely, the index \( j_s \) was found as the index maximizing the correlation coefficient

\[
|\hat{S}(j,:)s_{\text{sim}}^T|/\|\hat{S}(j,:)|\|s_{\text{sim}}\|),
\]

\( 1 \leq j \leq m \). Fig. 3 shows an example of finding the best match between \( s_{\text{sim}} \) and \( \hat{S} \).

Then, without a suppression, we get the FastICA approximation \( \hat{a}_{\text{sim}} \) for \( a_{\text{sim}} \) by setting \( \hat{a}_{\text{sim}} = \hat{A}(; j_s) \). If the suppression \( P \) is used, the FastICA estimate for the suppressed topography is \( P\hat{a}_{\text{sim}} = P\hat{A}(; j_s) \) as described at the end of section (2.1.5), and therefore no extra run with FastICA is needed with the suppressed data matrix. Fig. 4 describes the flowchart of the analysis carried out.

### 2.3.4. Source localization

Next, we want to carry out the source localization based on the estimated topographies \( \hat{a}_{\text{sim}} \) and \( P\hat{a}_{\text{sim}} \). The aim is to see how accurately the position \( M(:, i_d) \) of the simulated dipole can be recovered by a dipole search using the precomputed lead-field matrix \( L \) and the estimated topographies.

Before starting the source localization we need to modify the lead-field matrix \( L \) with the same preprocessing steps which we have performed for the data matrix \( X_{\text{dem}} \), namely, setting the zero potential level, compressing, see section 2.1.2. The resulting modified lead-field matrix is

\[
L_p = C_{\text{comp}}L_c,
\] (19)

where \( L_c \) is the zero potential level lead-field matrix, given by \( L_c(i,j) = L(i,j) - \frac{1}{M} \sum_{k=1}^{M} L(k,j) \), and \( C_{\text{comp}} \) is the compression matrix.
In the dipole search with \( \hat{a}_{\text{sim}} \), we use the information that the unknown dipole is a cortical one, and therefore we search for a dipole, with index \( \hat{i}_d \), whose topography \( L_p(\cdot, \hat{i}_d) \) best matches with \( \hat{a}_{\text{sim}} \), up to magnitude and sign, i.e., index \( i = \hat{i}_d \) maximizes the correlation coefficient

\[
\frac{\|L_p(\cdot, i)^T \hat{a}_{\text{sim}}\|}{\|L_p(\cdot, i)\| \|\hat{a}_{\text{sim}}\|}.
\]

(20)

The position of this best matching dipole is \( M(\cdot, \hat{i}_d) \), and therefore the localizing error is

\[
e_d = \|M(\cdot, \hat{i}_d) - M(\cdot, \hat{i}_d)\|.
\]

(21)

Similarly, for the estimated suppressed topography \( P\hat{a}_{\text{sim}} \) the best matching dipole with index \( \hat{k}_d \) is found by replacing \( L_p \) and \( \hat{a}_{\text{sim}} \) by \( PL_p \) and \( P\hat{a}_{\text{sim}} \), respectively in (20). The corresponding localization error \( e_d \) is given by (21) with \( \hat{i}_d \) replaced by \( \hat{k}_d \).

3. Results

3.1. Suppression of topographies and relative errors in estimated topographies

In this section we show with measured data that the suppression methods diminish the relative error of the estimated topographies. In PC and wavelet suppression, where only a few leading components are suppressed, the neural EEG will be suppressed much less than the artifactual ones. Fig. 5 shows how the topographies are suppressed with respect to different number of PCs, and that the small topographies are suppressed less than the largest EEG. In fact, here we can assume that the neural topographies are among the small ones. In this figure, the analysis time interval was chosen from 15 to 250 ms.

The main purpose of the suppression is to correct the high relative errors of the ICA-estimated topographies due to large artifacts. The success of the correction was assessed by computing the relative errors, on the measured data, both without and with the suppression, using the equation (4). We did this analysis for both the PC and wavelet suppressions. In Fig. 6, we present the (normalize) relative errors for a PC suppression. The analysis time interval was chosen to be 15–250 ms. For computing the suppression matrix, the restricted time subinterval was 15–50 ms and the number of suppressed PCs was \( n_s = 3 \); see section 2.2.1. The rank of the compressed data was 30. The relative errors were also computed for the wavelet suppression and the results were similar to those of the PC suppression.

When no suppression is performed, the relative errors become rather high with decreasing \( \|A(\cdot, j)\| \). For these components, the suppression works effectively by decreasing the relative error. For the large components, the error might even increase. However, they are mostly artifactual components. For the whitened data, the relative error is equal, \( \sigma \sqrt{m} \) according to (4), for all topographies. To compare the error of the other
suppression methods to that of whitening, in the Figure 6, we have divided the relative errors by $\sigma \sqrt{m}$.

3.2. Finding a dipolar source and comparing the suppression methods using the simulation

To evaluate how the suppression methods improve the source localization in a realistic context, we formed the $X_{dem}$ data by mixing the simulated data of a given source dipole with the measured data with large artifacts, as described in section 2.3.3. The combination of three described dipole positions and two waveforms at different time intervals resulted in six simulations. The time interval of the analysis was 10–250 ms, in the difficult case, for the simulated waveform $s_{sim1}$ (10–35 ms), and 15–250 ms, in the easy case, for the simulated waveform $s_{sim2}$ (15–40 ms).

FastICA was then used to separate 30 independent components yielding the estimated mixing matrix $\hat{A}$ given by (3). The three suppression methods, described in section 2.2, with suppressing matrices $P$, were then applied to $\hat{A}$ giving the suppressed mixing matrices as $P\hat{A}$.

In the estimated mixing matrix $\hat{A}$, the estimate for the simulated topography $\hat{a}_{sim} = \hat{A}(; i_s)$ was identified by letting $i_s$ be the index $j$ which maximized correlation coefficients $|\hat{S}(j,:)s_{sim}^T|/(||\hat{S}(j,:)||||s_{sim}||)$, which coefficients were not effected either by large artifacts or suppression. On the contrary, the estimated topography $\hat{a}_{sim}$ was expected to be distorted by the artifacts but the estimated suppressed topography $P\hat{a}_{sim}$ much less. In order to assess the effect of all three suppression methods, the correlation coefficients $||\hat{a}_{sim}||/||a_{sim}||$ and $||P\hat{a}_{sim}||/||Pa_{sim}||$ were computed and tabulated in Tables 1-3 for comparison. A correlation coefficient equal to 1 indicates a perfect match and a coefficient equal to 0 no match at all.

Fig. 3 shows an example of the set of time courses found by FastICA for the combined data $X_{dem}$. In this data, there were large artifacts due to the TMS stimulus, and the one known neural component, whose amplitude was 3 orders of magnitude lower than that of the artifacts. Among the estimated time courses, the one corresponding to the simulated time course was found in the upper right corner in Fig. 3 (A) (in red), with the good correlation coefficient of about 0.85, despite the presence of artifacts.

In Fig. 7, we are illustrating the corresponding simulated and the estimated topographies, $a_{sim}$ and $\hat{a}_{sim}$, respectively. The identification of the estimated topography was based on matching the time courses, as was shown in Fig. 3. Despite the well-found time course, the estimated topography can be seen badly distorted due to artifactual dominance. The wavelet suppression matrix $P$ was then computed based on the combined data $X_{dem}$, and both the simulated and the estimated topographies were suppressed by it. The effect of suppression is clearly shown in the visualizations of the topographies $Pa_{sim}$ and $P\hat{a}_{sim}$. The suppressed topographies have evident resemblance with each
other, which is essential for correct source localization. In the figure, the analysis time interval was chosen to be 15–250 ms, and the simulated dipole had a waveform from 15 to 40 ms.

To evaluate the usefulness of the estimated topography in the source localization, we performed the dipole search, described in section 2.3.4. The localization error was described as the distance between the simulated and estimated dipole locations according to (21). The results for the correlation coefficient and the dipole search errors are represented in tables 1–3. In general, all of the suppression methods decreased the dipole localization error. The subinterval for computing the suppression $P$ was from 10 to 60 ms in the difficult case, and from 15 to 60 ms in the easy case. The number of PCs, $n_s$, selected for rescaling in the PC and wavelet suppressions was chosen as explained in section 2.2.1. For both types of suppression, the kink of $E_s$ was identified at $n_s = 4$, in the difficult case, i.e., when $s_{sim1}$ was placed from 10 to 35 ms, and $n_s = 3$, in the easy case, when $s_{sim2}$ was placed from 15 to 40 ms.

For the difficult case, the PC method had errors between 0.53 cm (left Wernicke’s area) and 4.3 cm (left Broca’s area). For the easy case, the errors ranged between 0.67 cm (left Broca’s area) and 0.85 cm (left Wernicke’s area). The wavelet suppression provided an error of 0 cm for the easy case in all of the three positions, but for the difficult case, the errors were up to 6.2 cm. With the whitening suppression, the errors were 0 cm for the easy case, whereas for the difficult case the errors ranged from 0 to 1.2 cm. The correlation coefficients ranged between 0.50–0.83 for the PC suppression, 0.43–0.92 for the wavelet suppression, and 0.54–0.62 for the whitening. The results shown in table 1–3 were obtained with the Haar wavelets. We also experimented with the Daubechies wavelets, but the results were not as good as with the Haar wavelets.

4. Discussion and conclusion

The independent component analysis (ICA) has been used in the EEG analysis by separating the data into independent components. However, when applying ICA to TMS–EEG data, the large artifacts in EEG may greatly distort the ICA separation. A successful removal of such artifactual EEG from the data is a difficult task. In this paper we studied how badly the large artifacts distort the ICA separation, and whether the distortion could be avoided without removing the artifacts.

We showed that in the ICA separation the time courses of the independent components are not affected by the large artifacts, while their topographies could be greatly distorted. Next, we showed how this distortion can be circumvented. We introduced three suppression methods based on principal component analysis, wavelet analysis and whitening of the data matrix, respectively. The suppression modifies the EEG data so that the ICA separation of the suppressed data becomes reliable. The suppression, instead of removing the artifactual EEG, rescales the data to about the same magnitude as the neural EEG. For the suppressed data, ICA returns the estimates of the original
time courses, but instead of the estimates of the original topographies, it returns the 
estimates of the modified ones. For the neural components, however, the relative errors 
in the estimated modified topographies have been greatly diminished, and thus those 
estimates, when used in source localization, yield better localizing accuracy than the 
estimates without suppression.

We tested the suppression methods with measured and simulated data. First, with 
measured TMS–EEG data containing large artifacts, we studied the estimation errors of 
the neural components (or rather, components with small topographies) in a FastICA 
run. The diminishing of the errors was confirmed by computing the relative errors in the 
estimated topographies with and without suppression.

Next we, by using simulated data, examined the effect of the suppression methods to 
estimated topographies and evaluated the source localizing accuracy. In order to form a 
realistic testing environment, the test data was constructed by mixing the simulated EEG 
data of a given cortical dipole with measured TMS–EEG data containing large muscle 
artifacts in the left Broca’s area. The test data consisted of six different combinations; 
two time intervals of the given dipole, an easy case (15–40 ms) and a difficult case (10–35 ms), and three dipole positions, the left Broca’s area, the left Wernicke’s area 
and a position in the right hemisphere corresponding to Broca’s area. All of the three 
suppression methods were tested and the results compared with those obtained without 
suppression. For the comparison, the correlation coefficients (CC) between the estimated 
dipole topographies and the true ones and the localizing errors were computed. Note that 
in the tests it was possible to reach zero localizing error, because the topography of the 
given dipole in all cases was exactly one of the lead-field matrix columns.

In the localization task the wavelet suppression worked ideally for the easy case $s_{\text{sim2}}$, 
while it showed a poor performance with the difficult case $s_{\text{sim1}}$ putting its usefulness 
into doubt in its present form. The PC suppression performed well in the easy case 
and in the difficult case much better than the wavelet suppression. The number $n_s$ of 
the suppressed PCs was low, $n_s = 3, 4$, and therefore the data was not changed very 
much in the suppression. The whitening suppression performed best in localization in 
al cases, and it even worked well in localizing the dipole in Broca’s area in the difficult 
case, where the localization error provided by the other two suppression methods was 
not satisfactory. The benefit of whitening suppression is also that it is readily available 
without extra parameter adjusting. Its drawback is that it greatly changes the data, and 
thus visually, the whitened topography is not informative.

The correlation coefficients between the true and estimated dipole topographies were 
generally not satisfactory, although the suppression methods improved them. Still, the 
estimated suppressed topographies contained a sufficient amount of information for a 
successful dipole localization with the prior information on the unknown source that it 
consisted of a single cortical dipole oriented perpendicular to the cortex. In fact, this 
prior information has frequently been used in the EEG analysis. The low values of the 
correlation coefficient, makes it a poor measure of the quality of the source localization by 
the ICA estimated topographies. We suggest that in the future studies a better goodness

16
parameter for the dipole search should be found, e.g., one based on the confidence region for the estimated dipole location.

In summary, our theoretical and experimental results let us conclude that it is possible to study the neuronal independent components found by ICA also in the case of highly artifactual TMS-EEG data. The time courses of the hidden components are not affected by the artifacts, and only the topographies are distorted. However, that distortion can be circumvented by suppression methods, and all of the three suppression methods presented in this paper are good for that, and accordingly, they can be used in improving the source localization accuracy. These suppression methods, used with ICA and applied to TMS-EEG data, open possibilities for studying different sites of the brain, where the stimulation evokes not only muscle artifacts but also other kinds of artifacts that badly distort the ICA separation.

Acknowledgments

The authors want to thank the Academy of Finland for funding. J.C. Hernandez also wants to thank CIMO (Centre for International Mobility) and CONACYT (Consejo Nacional de Ciencia y Tecnologia) Mexico for the scholarships. J. Metsomaa would like to thank the Helsinki Biomedical Graduate School, and the Instrumentarium Foundation for Science for funding.
Appendix A. Preprocessing and obtaining ICA estimates for A and S

This Appendix presents the formulas used in the preprocessing of the data matrix $X_1$ and explains how the estimates for the mixing matrix $A$ and the time course matrix $S$ are formed. The zero potential level is set at each time point by setting the average over channels equal to zero in the $X_1$, yielding the data matrix $X_2$ as

$$X_2(j,t) = X_1(j,t) - \frac{1}{M} \sum_{k=1}^{M} X_1(k,t), \quad (A.1)$$

for all $j = 1, \ldots, M$, $t = 1, \ldots, N$. Next $X_2$ is time centered in each channel by setting the time average equal to zero, yielding the matrix $X_3$, as

$$X_3(j,t) = X_2(j,t) - \frac{1}{N} \sum_{s=1}^{N} X_2(j,s), \quad (A.2)$$

for all $j = 1, \ldots, M$, $t = 1, \ldots, N$. Finally, $X_3$ is compressed by letting the compressed matrix $X$ to consist of the $m$ leading principal components of $X_3$, with $m \leq \text{rank}(X_3) \leq M$. To this end, let

$$X_3 = U D V^T = \sum_{j=1}^{r} d_j u_j v_j^T, \quad (A.3)$$

be the singular value decomposition (SVD) of $X_3$, where $U$ and $V$ are $M \times M$ and $N \times N$ orthogonal matrices, respectively, $D = \text{diag}_{M \times N}(d_1, \ldots, d_r, 0, \ldots, 0)$ with non-zero singular values $d_1 \geq d_2 \geq \ldots \geq d_r > 0$, $r = \text{rank}(X_3)$, and $u_j = U(:,j), v_j = V(:,j)$, $j = 1, \ldots, r$. Then,

$$X = \sum_{j=1}^{m} d_j u_j v_j^T = U(:,1:m) \text{diag}_{m \times m}(d_1, \ldots, d_m) V(:,1:m)^T \quad (A.4)$$

is the desired compressed data matrix with $\text{rank}(X) = m$.

The number $m$ of the remaining principal components $d_j u_j v_j^T$ is chosen to be the (assumed) number of the latent independent components in the data $X_3$. The compression can also be represented by the compression matrix $C_{\text{comp}}$ as $X = C_{\text{comp}} X_3$ with

$$C_{\text{comp}} = \sum_{k=1}^{m} u_k u_k^T = U(:,1:m)U(:,1:m)^T. \quad (A.5)$$

The compressed data matrix $X$ is still whitened before the FastICA or some other ICA algorithm is applied. Whitening transforms the data so that its covariance matrix becomes the identity matrix $I$. The whitened data matrix $X_{\text{white}}$ can be given by the SVD of $X$ as

$$X_{\text{white}} = \sqrt{N} V(:,1:m)^T = BX, \quad (A.6)$$

where

$$B = \sqrt{N} \text{diag}_{m \times m}(d_1^{-1}, \ldots, d_m^{-1}) U(:,1:m)^T, \quad (A.7)$$

and

$$d_j = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} X_3(j,t)^2$$

for $j = 1, \ldots, m$.
is the whitening matrix, and furthermore,

\[ X = CX_{\text{white}}, \]  

(A.8)

where

\[ C = \frac{1}{\sqrt{N}} U(:, 1 : m) \text{diag}_{m \times m}(d_1, \ldots, d_m), \]

(A.9)

is the dewhitting matrix.

It follows that \( \text{Cov}(X_{\text{white}}) = N^{-1} X_{\text{white}} X_{\text{white}}^T = I \), as the term "whitened" suggests. Because \( X = AS \), we get

\[ X_{\text{white}} = WS, \]  

(A.10)

where \( W = BA \) is the orthogonal \( m \times m \) weight matrix with orthonormal columns \( w_j = W(:, j), j = 1, \ldots, m \), which are called weight vectors.

After these preprocessing steps, FastICA, or any other ICA algorithm, is applied to \( X_{\text{white}} \) to find the weight matrix \( W \) (the whitened mixing matrix), and the algorithm yields an orthogonal matrix \( \hat{W} \) as an estimate for \( W \). An estimate \( \hat{S} \) for \( S \) is then obtained with (A.10) as

\[ \hat{S} = \hat{W}^T X_{\text{white}}. \]  

(A.11)

It follows that

\[ \text{Cov}(\hat{S}) = \frac{1}{N} \hat{S} \hat{S}^T = I, \]  

(A.12)

because \( \text{Cov}(X_{\text{white}}) = I \) and \( \hat{W} \) is orthogonal. An estimate for \( \hat{A} \) is obtained by multiplying equation \( X = AS \) by \( N^{-1} S^T \) from the right, which with (A.12) yields

\[ \hat{A} = \frac{1}{N} X \hat{S} \hat{S}^T. \]  

(A.13)

Appendix B. Estimate \( \hat{S} \) is not affected by large artifacts

We verify the claim that \( \hat{S} \) is not affected by large artifacts, first for FastICA. In this paper we have used the symmetric form of the FastICA algorithm (Hyvärinen and Oja 2000), and we start by summarizing how it works.

When FastICA is applied to a given whitened data \( X_{\text{white}} = WS \), where \( W \) and \( S \) are \( m \times m \) and \( m \times N \) matrices, FastICA searches for the orthonormal weight vectors \( w_1, \ldots, w_m \), i.e., the columns of \( W \), as the maxima of the negentropy function,

\[ J_G(w^T X_{\text{white}}) = \left[ \frac{1}{N} \sum_{j=1}^{N} G(w^T X_{\text{white}}(:, j)) - c \right]^2, \]  

(B.1)
where \( w \) varies over all \( m \)-dimensional unit vectors, \( G(u) \) is the contrast function and \( c \) is the mean of \( G(u) \) with \( u \) being the scalar Gaussian random variable with zero mean and unit variance. The contrast function used in this paper was \( G(u) = \log[\cosh(u)] \).

The algorithm starts with random initial orthonormal vectors \( w_1, \ldots, w_m \) and upgrades them, step by step, retaining the orthonormality and trying to push each \( w_j \) as close to one of the maxima of (B.1) as possible. After no changes in the upgrading have happened up to a given tolerance, the algorithm stops, and the estimate \( \hat{W} = [w_1, \ldots, w_m] \) for \( W \) has been formed. The estimate for \( S \) is obtained by

\[
\hat{S} = \hat{W}^T \mathbf{x}_{\text{white}}. 
\]  

(B.2)

Because of the random initial vectors, FastICA returns slightly varying \( \hat{W} \) and \( \hat{S} \) from one run of the algorithm to another.

Next we examine how FastICA treats two different data matrices \( \mathbf{X} = \mathbf{A} \mathbf{S} \) and \( \tilde{\mathbf{X}} = \tilde{\mathbf{A}} \mathbf{S} \) with the same time course matrix \( \mathbf{S} \). Let \( \mathbf{x}_{\text{white}} = \mathbf{W} \mathbf{S} \) and \( \tilde{\mathbf{x}}_{\text{white}} = \mathbf{W} \mathbf{S} \) be the corresponding whitened data. Because \( \mathbf{W} \) and \( \mathbf{W} \) are orthogonal matrices, we get that \( \tilde{\mathbf{x}}_{\text{white}} = \mathbf{Z} \mathbf{x}_{\text{white}}, \) where \( \mathbf{Z} \) is an orthogonal matrix.

We now use the above result to verify the main claim of this appendix. Let \( \mathbf{X} = \mathbf{A} \mathbf{S} \) be given and we choose \( \tilde{\mathbf{X}} \) as

\[
\tilde{\mathbf{X}} = \mathbf{W}^T \mathbf{B} \mathbf{X} = \mathbf{S} = \mathbf{I} \mathbf{S}, \quad \text{20}
\]
where $B$ is the whitening matrix for $X$ and $BX = X_{\text{white}} = WS$. Then $\hat{X} = \hat{A}S$ with $\hat{A} = I$. It follows that if FastICA is applied to $X = AS$, it finds an estimate $\hat{S}$ for $S$ with the same accuracy as when it is applied to $X = IS$. Therefore, the estimation errors in $\hat{S}$ are independent of the mixing matrix $A$, and thus independent of the artifactual topographies (columns) of $A$.

The above reasoning remains valid if FastICA is replaced by any other ICA algorithm which for different orthogonal matrices $W$ returns essentially equal estimates for the time course matrix $S$ when applied to the whitened data $X_{\text{white}} = WS$ with the same $S$. Such algorithms are, e.g., Infomax and the (fast fixed point) maximum likelihood algorithm (Hyvärinen et al., 2001).

Appendix C. Relative Error $\rho_j$

In this appendix we derive the Equation (4) for the relative error $\rho_j$ of the estimate $\hat{A}(::j)$, where $\hat{A}$ is the estimate for the $M \times m$ mixing matrix $A$. Assume that $\hat{W}$ is an estimate for the true weight matrix $W_0$ and

$$\hat{W} = W_0 + \sigma \varepsilon,$$

where $\sigma > 0$ is the noise level and the elements $\varepsilon(j,t)$ of the error matrix $\varepsilon$ are uncorrelating with zero mean and unit variance. On the other hand, $\hat{A} = CW$, where $C$ is the $M \times m$ dewhittening matrix, because $\hat{A} = N^{-1}X\hat{S}^T = N^{-1}CX_{\text{white}}\hat{S}^T = CW$. Similarly, $A = CW_0$. Then the error matrix $\tau$ of $\hat{A}$ gets the form,

$$\tau = \hat{A} - A = C(W - W_0) = \sigma C\varepsilon.$$  

Next we consider the mean square error of the column $\hat{A}(::j)$ which is $E\{\|\tau(::j)\|^2\}$, where $E$ stands for the mean (expectation value). We get,

$$E\{\|\tau(::j)\|^2\} = E\{\tau(::j)^T\tau(::j)\} = \sigma^2 E\{\varepsilon(::j)^TC^T\varepsilon(::j)\} = \sigma^2 \sum_{k,l=1}^{m} \Gamma(k,l) E\{\varepsilon(k,j)\varepsilon(l,j)\} = \sigma^2 \sum_{k=1}^{m} \Gamma(k,k),$$

where $\Gamma = CC^T$ and $E\{\varepsilon(k,j)\varepsilon(l,j)\} = \delta_{k,l}$ (Kronecker delta), because $\varepsilon(k,j)$ and $\varepsilon(l,j)$ are uncorrelating. On the other hand,

$$\sum_{k=1}^{m} \Gamma(k,k) = \sum_{j=1}^{M} \sum_{k=1}^{m} C(j,k)(C^T)(k,j).$$

Furthermore, $X = CX_{\text{white}}$ and so

$$XX^T = CX_{\text{white}}X_{\text{white}}^TC^T = NCC^T,$$
because $N^{-1} \mathbf{X}_{\text{white}} \mathbf{X}_{\text{white}}^T = \mathbf{I}$. Therefore, (C.4) implies,

$$
\sum_{k=1}^{m} \Gamma(k, k) = \frac{1}{N} \sum_{j=1}^{M} \sum_{k=1}^{m} \mathbf{X}(j, k)(\mathbf{X}^T)(k, j)
= \frac{1}{N} \sum_{j=1}^{M} \sum_{k=1}^{m} \mathbf{X}(j, k)^2 = \frac{1}{N} \text{trace}(\mathbf{X} \mathbf{X}^T).
$$

(C.6)

This with (C.3) implies,

$$
E\{\|\tau(:, j)\|^2\} = \frac{\sigma^2}{N} \text{trace}(\mathbf{X} \mathbf{X}^T).
$$

(C.7)

Finally, the relative error $\rho_j$ of $\hat{\mathbf{A}}(:, j)$ is the ratio

$$
\rho_j = \sqrt{\frac{E\{\|\tau(:, j)\|^2\}}{\|\hat{\mathbf{A}}(:, j)\|^2}},
$$

(C.8)

and so with (C.7) and (C.8) we get (4).

It can be checked that (C.8) also yields the relative error of the estimated suppressed topographies; we only need to replace $\mathbf{X}$ and $\hat{\mathbf{A}}$ by $P \mathbf{X}$ and $P \hat{\mathbf{A}}$. 
References


Figure captions

Fig. 1 3D display of the subject’s MRI generated by the Nexstim eXimia NBS, and typical artifact waveforms after the stimulation of Broca’s area. (A) The stimulus was delivered over left Broca’s area. The arrows indicate the direction of the induced current. (B) Absolute value of the average over channels of the evoked EEG signal. (C) Signal recorded by electrode FT9, which is close to the stimulation target. (D) Signal recorded by electrode C6, which is far from the stimulation target. In (C) and (D) the electrodes were labeled according to the international 10–20 system.

Fig. 2 Top: The three different simulated topographies. The neural source was placed in the left Broca’s (a_sim1) or Wernicke’s (a_sim2) area, or in the contralateral hemisphere corresponding to Broca’s area (a_sim3). The white cross indicates the stimulation site. Bottom, the Mexican hat waveform used for the simulated source.

Fig. 3 The search for the best match between the rows of $\hat{S}$ and $s_{sim}$. (A) ICA was applied on the data which constituted both the measured data and one simulated neural component. The ICA-estimated time courses $\hat{S}$ are shown in black, whereas the simulated time course with green. Among the estimated components, the best matching estimate $\hat{s}_{sim}$ compared to the simulated neural time course $s_{sim}$ was identified (red line, upper right corner). The y axis represents the amplitude of the ICs in arbitrary units (A.U.), whereas the x axis indicates the time in milliseconds (ms). (B) The best matching estimated time course as compared to the simulated one. (C) The worst matching estimated time course. Note, the time interval is from 15 to 250 ms.

Fig. 4 Flowchart with the analysis steps. First, we created the combined data. Second, we preprocessed the data and run ICA. Third, we identified the estimated component, that was best matching the simulated neural one, based on the waveform. Then, we suppressed the corresponding estimated topography. Fourth, we compute the quality parameters, i.e., the correlation coefficient between the simulated and estimated topographies, and estimate the dipole localization. The parameters were computed for the suppressed topography.

Fig. 5 The effect of PC suppression on the topographies when different number of PCs are suppressed. The bars indicate the magnitudes of the estimated topographies with and without using suppression. When no suppression is used, there is a great variability between the magnitudes of the components, which represents the high proportion of artifacts in the large topographies. However, when applying the suppression, the magnitudes of the largest topographies decline rapidly as compared to the other topographies. This shows the effect of suppression, which is to reduce the artifactual parts of the components. Note that, just based on the topographies, it is impossible to make the distinction between the neural and artifact components. The identification of the neural components is based on inspecting the time courses subjectively.
Fig. 6 The relative errors of the topographies after PC suppression as computed based on equation (4). The red bars show the errors for each topography before and the green ones after the suppression. The blue line indicates the relative error after whitening. All the errors are scaled such that, for whitening, the error is equal to 1. The topographies were ordered according to their magnitudes before the suppression (largest ones on the left). After the suppression, on average, the errors of the topographies decline significantly.

Fig. 7 An example of the topography estimation with the wavelet suppression for a position in the right hemisphere corresponding to Broca’s area. Upper left: The simulated neural topography which corresponds to the simulated waveform in Fig. 2. The simulated component was added to the measured data producing the combined data. ICA was then performed on the combined data, and the best matching estimated time course was identified as shown in Fig. 3 (A) (in red). In the lower left, the estimated topography is shown, as identified according to the time course (Fig. 3 (B)). Due to the artifacts in the left channels, the topography is largely distorted towards that direction. Alternatively, both the simulated and the estimated topographies can be suppressed using the same suppression matrix $\mathbf{P}$, which was computed based on the combined data. The suppressed simulated (upper right) and the suppressed estimated (lower right) topographies are seen to have strong resemblance with each other.
Figure C.3:
Figure C.4:

-Measured EEG data contain several unknown, neural and artefactual components:
  - Measured data $X_1$

-Simulated data contain one known neural component:
  - Simulated data $X_{\text{sim}} = \lambda a_{\text{sim}}$
  - Create the "Mexican hat" waveform $s_{\text{sim}}$
  - Choose the dipole location $r_{\text{sim}} = M(i, t)$ and compute the topography $a_{\text{sim}}$

-Combined data $X_{\text{comb}} = X_1 + X_{\text{sim}}$

-Perform the analysis workflow on $X_{\text{comb}}$ as explained in section (2.1.5), and get the estimate for the suppressed simulated component: $\hat{a}_{\text{sim}}$ and $\hat{a}_{\text{sim}}$

-Compute the quality parameters for the analysis:
  - Correlation coefficient between the suppressed simulated and suppressed estimated topographies: $P_{\text{sim}}$ and $P_{\text{est}}$
  - The error between the simulated and estimated dipole locations: $r_{\text{sim}}$ and $r_{\text{est}}$
Figure C.5:

Amplitude
Figure C.7: Diagram illustrating the process of adding simulated data to measured data, followed by ICA, estimation, suppression, and obtaining a good match.
Table 1. Localization error of the dipole localization and correlation coefficient (CC) computed with three suppression methods for the difficult and easy case in left Broca’s area. In the case of no suppression the CC was computed between the simulated $a_{sim}$ and the estimated simulated $\hat{a}_{sim}$ topographies. When the suppression was performed the CC was determined between the suppressed simulated and suppressed estimated topographies: $Pa_{sim}$ and $P\hat{a}_{sim}$.

<table>
<thead>
<tr>
<th>Suppression Method</th>
<th>Localization Error (cm)</th>
<th>CC</th>
<th>Difficult case $s_{sim1}$</th>
<th>Suppression Method</th>
<th>Localization Error (cm)</th>
<th>CC</th>
<th>Easy case $s_{sim2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Suppression</td>
<td>1.4</td>
<td>0.17</td>
<td></td>
<td>No Suppression</td>
<td>1.4</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Whitening</td>
<td>1.2</td>
<td>0.54</td>
<td></td>
<td>Whitening</td>
<td>0</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>4.3</td>
<td>0.50</td>
<td></td>
<td>PCA</td>
<td>0.67</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Wavelet</td>
<td>6.2</td>
<td>0.43</td>
<td></td>
<td>Wavelet</td>
<td>0</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Localization error of the dipole localization and correlation coefficient (CC) computed with three suppression methods for the difficult and easy case in left Wernicke’s area. In the case of no suppression the CC was computed between the simulated $a_{\text{sim}}$ and the estimated simulated $\hat{a}_{\text{sim}}$ topographies. When the suppression was performed the CC was determined between the suppressed simulated and suppressed estimated topographies: $Pa_{\text{sim}}$ and $P\hat{a}_{\text{sim}}$.

<table>
<thead>
<tr>
<th>Suppression Method</th>
<th>Localization Error (cm)</th>
<th>CC</th>
<th>Difficult case $s_{\text{sim1}}$</th>
<th>Easy case $s_{\text{sim2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Suppression</td>
<td>6.8</td>
<td>0.35</td>
<td>No Suppression</td>
<td>6.8</td>
</tr>
<tr>
<td>Whitening</td>
<td>0.61</td>
<td>0.59</td>
<td>Whitening</td>
<td>0</td>
</tr>
<tr>
<td>PCA</td>
<td>0.53</td>
<td>0.61</td>
<td>PCA</td>
<td>0.85</td>
</tr>
<tr>
<td>Wavelet</td>
<td>6.2</td>
<td>0.88</td>
<td>Wavelet</td>
<td>0</td>
</tr>
</tbody>
</table>

CC: Correlation Coefficient
Table 3. Localization error of the dipole localization and correlation coefficient (CC) computed with three suppression methods for the difficult and easy case in right hemisphere location corresponding to Broca’s area. In the case of no suppression the CC was computed between the simulated $\mathbf{a}_{\text{sim}}$ and the estimated simulated $\mathbf{\hat{a}}_{\text{sim}}$ topographies. When the suppression was performed the CC was determined between the suppressed simulated and suppressed estimated topographies: $\mathbf{P}_{\text{a}_{\text{sim}}}$ and $\mathbf{\hat{P}_{\text{a}_{\text{sim}}}}$.

<table>
<thead>
<tr>
<th>Suppression Method</th>
<th>Error (cm)</th>
<th>Difficult case $s_{\text{sim}1}$</th>
<th>Easy case $s_{\text{sim}2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Suppression</td>
<td>7.6</td>
<td>0.24</td>
<td>No Suppression</td>
</tr>
<tr>
<td>Whitening</td>
<td>0</td>
<td>0.55</td>
<td>Whitening</td>
</tr>
<tr>
<td>PCA</td>
<td>1.5</td>
<td>0.67</td>
<td>PCA</td>
</tr>
<tr>
<td>Wavelet</td>
<td>4.1</td>
<td>0.71</td>
<td>Wavelet</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Suppression Method</th>
<th>Error (cm)</th>
<th>Difficult case $s_{\text{sim}1}$</th>
<th>Easy case $s_{\text{sim}2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Suppression</td>
<td>7.6</td>
<td>0.25</td>
<td>7.6</td>
</tr>
<tr>
<td>Whitening</td>
<td>0</td>
<td>0.60</td>
<td>0</td>
</tr>
<tr>
<td>PCA</td>
<td>0.77</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>Wavelet</td>
<td>0</td>
<td>0.89</td>
<td>0</td>
</tr>
</tbody>
</table>