Independent Component Analysis of Non-invasively Recorded Cortical Magnetic DC-fields in Humans

Gerd Wübbeler, Andreas Ziehe, Bruno-Marcel Mackert, Klaus-Robert Müller, Lutz Trahms, Gabriel Curio

October 1999
In der Reihe GMD Report werden Forschungs- und Entwicklungs-
ergebnisse aus der GMD zum wissenschaftlichen, nicht-
kommerziellen Gebrauch veröffentlicht. Jegliche Inhaltsänderung
des Dokuments sowie die entgeltliche Weitergabe sind verboten.

The purpose of the GMD Report is the dissemination of research
work for scientific non-commercial use. The commercial
distribution of this document is prohibited, as is any modification
of its content.

Anschriften der Verfasser/Addresses of the authors:
Dr. Klaus-Robert Müller, Andreas Ziehe
Institut für Rechnerarchitektur und Softwaretechnik
GMD – Forschungszentrum Informationstechnik GmbH
Kekuléstraße 7
D-12489 Berlin-Adlershof
E-mail: {Klaus-Robert.Mueller, Andreas.Ziehe}@gmd.de

Dr. Lutz Trahms, Dr. Gerd Wübbeler
Physikalisch-Technische Bundesanstalt (PTB)
Abbéstraße 10-12
D-10587 Berlin

Dr. Gabriel Curio, Dr. Bruno-Marcel Mackert
Klinik für Neurologie, Klinikum „Benjamin Franklin”
Freie Universität Berlin
Hindenburgdamm 30
D-12489 Berlin

ISSN 1435-2702
Zusammenfassung


Schlagwörter

Biosignalanalyse, Biomagnetismus, Magnetoecephalography (MEG), DC-Messtechnik, Signalquellentrennung, Komponentenanalyse.

Keywords

Biomedical data processing, Biomagnetism, Magnetoecephalography (MEG), DC-recordings, Blind source separation, Independent component analysis.

Abstract

We apply a recently developed multi-variate statistical data analysis technique - so called blind source separation by independent component analysis - to process MEG recordings of near-DC fields. The extraction of near-DC fields from MEG recordings has great relevance for medical applications since slowly varying DC-phenomena have been found e.g. in cerebral anoxia and spreading depression in animals.

Comparing several blind source separation approaches, it turns out that an algorithm based on temporal decorrelation successfully extracted a DC-component which was induced in the auditory cortex by presentation of music. The task is challenging because of the limited amount of available data and the corruption by outliers, therefore we propose this application as a real-world testbed for studying the robustness of ICA methods.
1 Introduction

Recently, the feasibility of a non-invasive magnetic registration of near-DC\(^1\) magnetic fields from the human cortex using Superconducting Quantum Interference Devices (SQUIDs) has been shown [1]. Such near-DC phenomena may have importance for metabolic injuries of brain cells in stroke or migraine [2, 3, 4]. Since magnetoencephalography (MEG) records the spatio-temporal neumagnetic field with an array of biomagnetometers one can apply multivariate statistical methods for the data analysis. A popular method is Independent Component Analysis (ICA), where the continuous-valued latent variables of input data are inferred by imposing statistical independence on the outputs. ICA has received great attention in various technical application domains like acoustic source separation (e.g. [5, 6]), or telecommunication [7].

In addition this technique has been successfully applied to reduce artifacts in multi-channel EEG, MEG and MNG (magnetoencephalography) recordings [8, 9, 10, 11] and also to analyze evoked responses [12].

In this work we will concentrate on biomagnetic data analysis and show that ICA provides an efficient, unsupervised tool to extract an interesting physiological phenomenon from near-DC neumagnetic data. A chance for DC-coupled brain monitoring is of high medical relevance because many pathophysiological processes have their main energy in the frequency range below 0.1 Hz. Therefore, it is of utmost importance to further improve the signal extraction from DC-MEG data. The biomagnetic recording technology employed here is based on a mechanical modulation of the head, respectively, body position relative to the sensor [13]. This yields a high sensitivity which is both chance and challenge since it will not only enable physicians to detect minute injury-related fields [14, 15] but also poses interesting problems for data analysis since a multitude of different biological (and ambient or system’s noise) processes will add to the signal of pathophysiological interest. It is a helpful matter of fact that many of these processes vary in intensity independently of each other, e.g. slow fields from the eyes or from respiration, brain state changes related to drifting vigilance or periods of more focused attention, and – perhaps – near-DC phenomena in stroke such as peri-infarct and anoxic depolarizations. When introducing the present paradigm of prolonged auditory (music) stimulation for DC-MEG [1] we sought a physiological DC-source in the brain which we could (i) switch on and off arbitrarily, and (ii) which had a field pattern that could be predicted by comparison to other (plastic) evoked activities from auditory cortices [16]. Thus, this paradigm of externally controlled music-related DC-activations of auditory cortices defines a measurement and analysis scenario with almost complete knowledge about both the spatial pattern and the time course of a cerebral DC-source which on the other hand is fully embedded in the ‘real’ biological and ambient noise background. Hence it may serve as a testbed

---

\(^1\)below 0.1 Hz
for a critical comparison of advanced ICA approaches which all will have to cope with the ‘real world’ problems of low signal-to-noise ratio coming along with a limited number of data samples and the presence of outliers.
In section 2 we will first review some common ICA techniques (JADE, FastICA, TDSEP) and then (section 3) describe the neurophysiological experiment for which we analyze the robustness of the above mentioned ICA algorithms. We conclude with a discussion and outlook.

2 Blind Source Separation

2.1 Model

Due to the fact that magnetic fields of different bioelectric current sources superimpose linearly it is reasonable to assume that the measured values of the SQUID-sensor array can be modeled by a linear combination of component vectors

\[ \mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \]  

(1)

where \( \mathbf{x} = [x_1, \ldots, x_m]^T \), \( \mathbf{s} = [s_1, \ldots, s_n]^T \), \( m \geq n \). For independent component analysis we assume that the observed signals \( \mathbf{x}(t) \) consist of \( n \) underlying unknown sources \( \mathbf{s}(t) \), that are mutually statistically independent. Mathematically this means that the joint probability density function factorizes

\[ p(s_1, \ldots, s_n) = \prod_{i=1}^{n} p_i(s_i). \]  

(2)

It is assumed that each component \( s_i \) has zero mean. Within these assumptions one can separate the measurements \( \mathbf{x}(t) \) through the linear transformation

\[ \mathbf{u}(t) = \mathbf{W}\mathbf{x}(t) \]  

(3)

into independent components by imposing statistical independence on the output \( \mathbf{u}(t) \) of this demixing process and recover the original sources \( \mathbf{s}(t) \) from the observed mixtures up to scaling factors and a random permutation of the order. As both the mixing process \( \mathbf{A} \) and the sources \( \mathbf{s}(t) \) are unknown, these techniques are called blind source separation.

2.2 Three algorithms for blind source separation

In the following we will shortly review three representative types of source separation algorithms (JADE, FastICA, TDSEP) that each follow a different line of thought and stand for a number of similar algorithms. A substantial amount of research has been conducted on algorithms using higher-order statistics for estimation of ICA [17, 18, 19, 20, 21].
For off-line (batch) computation, Cardoso et al. [22] developed the JADE algorithm based on the (joint) diagonalization of matrices obtained from ‘parallel slices’ of the fourth-order cumulant tensor. This algorithm often performs very efficiently on low dimensional data if sufficiently many sample points are available. However, for high dimensional problems like MEG the effort for storing and processing the 4-th order cumulants is $O(n^4)$ and computation may become prohibitive. As a remedy for this problem, Oja and Hyvärinen developed an algorithm utilizing the so called fixed-point iteration [23] termed FastICA which uses the kurtosis as a contrast function (see [24, 25] for extensions to generalized contrasts) The FastICA algorithm can either estimate the independent components one-by-one using hierarchical decorrelation (deflation), or it can estimate all the independent components in parallel by symmetric decorrelation. In matrix notation FastICA takes the form

$$W' = W + \Gamma[\text{diag}(-\beta_i) + E\{g(u)u^T\}]W,$$

where $u = Wx$, $\beta_i = E\{u_i g(u_i)\}$ and $\Gamma = \text{diag}(1/(\beta_i - E\{g'(u_i)\}))$ where $g(u_i)$ is a non-linear contrast function. This version of FastICA has been shown to be equivalent [26] to the maximum likelihood approach for ICA given by a stochastic gradient descent [18, 21, 27, 28]

$$\frac{dW}{dt} \sim \eta(I - f(u)u^T)W,$$

where $f(u)$ is a properly chosen component-wise non-linear function and $\eta$ denotes a learning rate.

A potential problem of higher-order statistics is that they rely mainly on the tails of a distribution which makes these methods susceptible to outliers [17], as it is difficult to distinguish without knowledge of the distribution whether a data point contributes properly to the tail of the distribution or whether it is an outlier.

While the previously described set of methods utilizes higher-order moments to exploit the non-Gaussian distribution of the sources to achieve a separation, the temporal decorrelation algorithm TDSEP [29] relies on second-order statistics only, exploiting distinctive spectral/temporal characteristics of the sources (see also [30, 31, 32, 33]). Such inherent time structure of signals can be found particularly in neurophysiological recordings. The advantage of second-order methods is their computational simplicity and efficiency. Furthermore they are more robust against outliers and for a reliable estimate of covariances only comparably few samples are needed.

Let us first remember that for mutual independent signals the cross-correlation function vanishes. If the signals have a temporal structure resulting in a ‘non-delta’ auto-correlation function we can define so called time-delayed correlation matrices $R_{\tau(m)}$, which should be in diagonal form. This knowledge is used to calculate the unknown mixing matrix in eq. (1) as follows. Let us consider
time-lagged correlation matrices of the form

\[
R_{\tau}(\mathbf{x}) = \langle \mathbf{x}(t) \mathbf{x}^T(t - \tau) \rangle = \begin{bmatrix}
\phi_{x_1,x_1}(\tau) & \cdots & \phi_{x_1,x_n}(\tau) \\
\phi_{x_2,x_1}(\tau) & \cdots & \phi_{x_2,x_n}(\tau) \\
& \ddots & \vdots \\
\phi_{x_n,x_1}(\tau) & \cdots & \phi_{x_n,x_n}(\tau)
\end{bmatrix}
\]

where \(\phi_{x_i,x_j}(\tau) = \langle x_i(t) x_j(t-\tau) \rangle\) denotes the respective auto- or cross-correlation functions.

Since the mixing model in eq. (1) is just a linear transformation we can substitute \(\mathbf{x}(t)\) by \(\mathbf{A} \mathbf{s}(t)\) and get:

\[
R_{\tau}(\mathbf{A}) = \langle \mathbf{A} \mathbf{x}(t) \mathbf{A}^T(t - \tau) \rangle = \langle \mathbf{A} \mathbf{s}(t) \mathbf{A}^T(t - \tau) \rangle = \mathbf{A} R_{\tau}(\mathbf{s}) \mathbf{A}^T.
\]

Equation (4) shows that the matrix \(\mathbf{A}\) is a transformation which simultaneously diagonalizes all correlation matrices \(R_{\tau}(\mathbf{x})\). For two lagged correlation matrices, e.g. \(R_{\tau}(\mathbf{x})\) for \(\tau = 0\) and \(\tau \neq 0\), one can achieve a joint diagonalization (and hence find the mixing matrix \(\mathbf{A}\)) by solving the general eigenvalue problem \((R_{\tau=0}, R_{\tau})^{-1}) \mathbf{A} = \mathbf{A} \Lambda [30]\). It should be clear that the temporal decorrelation algorithm can be used successfully only if the signals have non-identical spectra i.e. distinctive auto-correlation functions, since otherwise the eigenvalues would be degenerate. Hence the quality of the signal separation depends strongly on the very choice of \(\tau\) [29], therefore it is better to try to diagonalize a larger set \(\{R_{\tau}(\mathbf{x})\}\) of delayed correlation matrices simultaneously. To achieve an approximate simultaneous diagonalization of several matrices one proceeds in two steps: (1) whitening and (2) a number of Jacobi rotations [22, 34]. First a whitening transformation \(\mathbf{W} = R_{\tau=0}^{-1/2} = (\mathbf{V} \Lambda \mathbf{V}^T)^{-1/2} = \mathbf{V} \Lambda^{-1/2} \mathbf{V}^T\) achieves a white basis \(\mathbf{z}(t) = \mathbf{W} \mathbf{x}(t)\) on a unit sphere [35]. The remaining set of time delayed correlation matrices \(R_{\tau}(\mathbf{x})\) can be diagonalized subsequently by a unique orthogonal transformation \(\mathbf{Q}\), since in the white basis all degrees of freedom left are rotations [22].

For several matrices, that share a common eigen-structure, a Jacobi-like algorithm proposed by Cardoso can be used to find a satisfying solution [34, 36]. The basic idea is that one can approximate the rotation matrix \(\mathbf{Q}\) by a sequence of elementary rotations \(Q_k(\phi_k)\) in a two dimensional subspace each trying to minimize the off-diagonal elements

\[
\min \sum_{\tau, \nu \neq \xi, j} |(R_{\tau})_{i,j}|
\]

of the respective \(R_{\tau}(\mathbf{x})\) matrices, where the rotation angle \(\phi_k\) can be calculated in closed form (see Cardoso [36] for details). The final rotation, which diagonalizes \(R_{\tau}(\mathbf{x})\) up to a certain level of accuracy, is then obtained by \(\mathbf{Q} = \prod_k Q_k(\phi_k)\).
Concatenation of both transforms (whitening $W$ and rotation $Q$) yields an estimate of the mixing matrix $A = W^{-1}Q$, which has to be inverted to get the demixing matrix $W = A^{-1}$.

As a side remark: one can carry the thought of simultaneous diagonalization of matrices even further. In principle any two (or more) matrices that are diagonal for the sources are sufficient to find a proper demixing transform $W$. Matrices that could be used apart from the time-delayed correlations introduced before, are e.g. correlation matrices of filtered signals [32, 37, 33] or slices of a higher-order cumulant tensor [22, 38].

2.3 Limits and Problems for Source Separation Algorithms

While using ICA (or other projection) algorithms one has to be aware of their assumptions (see above), general limits and difficulties and we will give a checklist of those possible problems in the following.

(a) A particularly hard practical problem is the availability of only few data points in combination with a high-dimensional sensor input, the latter being a problem of computational complexity that can be overcome by e.g. TDSEP or Fast ICA algorithms, while the former is a ubiquitous systematic statistical problem ("curse of dimensionality").

(b) Channel noise is potentially a rather serious harm to ICA algorithms as it effectively doubles the number of independent sources. Often, however, the application problem allows to construct an approximate noise model and projections to signal spaces orthogonal to the noise space can be performed [38, 39].

(c) A further difficulty comes from the independence assumption in Eq. (2): any projection algorithm can only retrieve and denoise signals within the subspace of the linear space of all components that we define by certain a priori assumptions. Generally speaking in data analysis we are always interested in finding a proper basis that is describing the relevant characteristics of the data. So we aim for a linear component analysis (generative model) where the components (latent variables) are meaningful with respect to the application in mind [40]. An orthogonality assumption leads to principal component analysis (PCA), orthogonality in some feature space gives rise to non-linear PCA (cf. [41]) and enforcing mutual independence of the components defines ICA.

(d) The number of sources that can be unmixed has to be assumed to be smaller or equal than the number of sensors. However, in biomedical measurements a multitude of microscopic sources contributes to the recorded signal. How these sources can be collapsed into fewer macroscopic sources depends on the particular biological system under study.

(e) The mixing model as defined in Eq.(1) can be too simple-minded and models that include noise terms (see discussion above) or cope with convolutive or even non-linear mixtures would be more appropriate. For MEG/EEG recordings a linear model is sufficient, due to the linearly superimposing magnetic/electric fields.
(f) Outliers can strongly decrease the performance of ICA algorithms involving higher-order statistics, whereas second-order algorithms are more robust against outliers.

3 Application

This section first gives some medical background on DC-recordings, then the experimental set-up and preprocessing is described and finally we apply different ICA techniques to the data and discuss the differences that we find in our results.

3.1 Clinical Background of DC-Recordings

Near-DC phenomena are expected in metabolic injuries to brain cells in stroke or migraine, e.g. in anoxic depolarization, peri-infarct depolarization or spreading depression [2, 3, 4]. Non-invasive electrical near-DC recordings are prone to large drift artifacts due to electrochemical instabilities at the electrode-skin interface. Up to now this limitation could be overcome only by invasive approaches [42, 43]. In contrast, Superconducting Quantum Interference Devices (SQUIDs) allow for a non-invasive magnetic registration of near-DC low frequency magnetic fields. Using this technology biomagnetic fields below 0.1 Hz (near-DC) arising from “injury currents” of traumatized tissue, e.g. muscle and nerve, have been measured non-invasively in vitro [15, 14]. Fields in this frequency and amplitude range were detected, quantified and continuously monitored non-invasively also from the human brain by employing an acoustical stimulation paradigm to induce a prolonged auditory cortex activation (for detailed physiological background see [1]).

3.2 Data acquisition and validation

The neuromagnetic field data were recorded in a standard magnetically shielded room (AK3b), operated at the Klinikum Benjamin Franklin by the PTB, using 49 low noise first order SQUID gradiometers (70 mm baseline) covering a planar area of 210 mm diameter [44]. The sensor was centered tangentially approximately over the left auditory cortex. The acoustic stimulation was achieved by presenting alternating periods of music and silence, each of 30 s length, to the subjects right ear during 30 min. of total recording time. The DC magnetic field values were acquired by using a mechanical horizontal modulation of the body position with a frequency of 0.4 Hz and an amplitude of 75 mm. This modulation transposed the DC magnetic field of the subject to the modulation frequency, which is less contaminated by magnetic noise. The recorded magnetic field data were processed by digital lock-in techniques in order to extract the modulation induced frequency components [45]. Then the DC-field of the subject was reconstructed from these frequency components by using a
transformation technique based on a virtual magnetic field generator [1]. These reconstructed DC magnetic field values, sampled at the modulation frequency of 0.4 Hz, gave a total number of 720 sample points per channel for the 30 minutes recording time and were used as input for the ICA-algorithms.

Figure 1: Input data used for ICA after DC preprocessing (demodulation and reconstruction); arranged according to sensor positions.

Let us examine the time courses of 30 minutes for all 49 channels (cf. Fig.1). At the first glance, the signals have an obvious trend behavior (slow drift) while possible components of interest are covered by other strong signals of unknown origin, i.e. the response to the stimulus is completely hidden in the data. To apply ICA algorithms to this data we have to make sure that the criteria of the checklist from section 2.3 are fulfilled. The hardest problem is probably posed by (a) since we have 49 channels and only 720 data points per channel. Additive channel noise (b) is a minor problem due to the experimental set-up, but internal very low frequency signals (drifts) are certainly present. As we are looking for a signal that is time-locked with the stimulus and due to
linear superposition of biomagnetic fields our assumption of temporal decorrelation/independence and a linear mixing model (c) holds. Also the number of sources (d) has to be less than the number of sensors. Even though the exact number of sources is unknown, at least the eigenvalue spectrum of the covariance matrix decayed rapidly. Finally, as we see from the occasional spikes in various channels of Fig. 1, outliers (f) pose a problem in this data set.

3.3 Results

We now apply TDSEP [29] to the data, reduced to a 23-dimensional subspace by PCA, using 50 time-lagged correlation matrices ($r = 1...50$ sample points) for simultaneous diagonalization. The 10 strongest ICA components are shown in Fig.3. Not surprisingly, the first component (ICA1) mainly captured the slow drift, that was already visible in the data in Fig.1. While most other components show irregular time courses reflecting the dynamics of undetermined processes it is noteworthy that their field maps feature spatially coherent field patterns which clearly distinguish them from random channel noise patterns. Remarkably, one component (ICA10) shows a (noisy) rectangular waveform. Its time course and frequency (see Fig.2) clearly displays the $\frac{1}{2}$ “on/off” characteristics of the stimulus. The spatial field distribution of ICA10 shows a bipolar pattern, located at the expected position of cortical activity [1]. Both findings give direct evidence that ICA10 represents the response to the acoustical stimulus. Although we do not expect that the cortical response resembles the

![ICA10 L2: 9.235](image)

Figure 2: Spatial field pattern, frequency content and time course of ICA10.

stimulus completely, computing the correlation coefficient between the stimulus and the ICA time courses provides a useful measure to evaluate and compare the performance of different separation algorithms. Applying the three algorithms
from section 2.2, we find that only the TDSEP algorithm is able to recover a
signal that is highly correlated to the stimulus, while Fast ICA and JADE fail
for this specific task (for correlation coefficients see also figure 4). There might
be a number of reasons for this finding. On the one hand the limited number of
sample points is a serious problem for algorithms based on higher-order statis-
tics, as they have to estimate a larger amount of parameters from the same
amount of data. On the other hand, the signal to noise ratio is problematic as
well and makes the distinction between different sources solely relying on the
probability density very difficult. Furthermore we note a number of outliers in
Fig. 1 that may harm the estimation of higher-order moments. Unfortunately
simply removing potential outliers did not improve the results, as one might
erroneously remove also data points which are important for a proper estimate
of the higher-order statistics.

In Fig. 4 we show the performance of the three algorithms as a different number
of PCA components was used for subspace projection. Clearly TDSEP is the
only algorithm which reliably extracts a component which is highly correlated
to the stimulus, given a sufficient amount of components (i.e. $> 20$). Fig. 5
shows the dependency of the separation result for TDSEP as a function of the
sample size. Already for 300 samples we observe an enhanced correlation, which
is even higher than the respective correlation coefficient obtained by the JADE
or FastICA algorithm for all 720 data points.

4 Discussion and Outlook

The present results provide deeper insights into strengths and limitations of ICA
approaches to process DC-magnetoencephalography data. Here, one physiological
and one methodological point shall be discussed:

(1) Under a general physiological point of view it is of primary interest to note
that when employing ICA it became possible on the single subject level (i.e.
without reverting to group statistics) to derive a faithful estimate for the time
course of the DC-activation level in a relatively circumscribed brain area (i.e.
the auditory cortex in the temporal lobe). Most importantly, this analysis pro-
ceeded fully blind to our a priori experimental background knowledge on both
the spatial signature of the music-related DC-fields (field map characteristic for
auditory cortex activations) and its time course (30 sec on and 30 sec off).
Both the spatial and the temporal source aspects were adequately captured in
one ICA component (ICA10) using TDSEP. It is noteworthy that in contrast
to earlier paradigms which identified cortical sources of short-term (2 - 9 sec
"sustained" fields [16] or potentials [46] by averaging at least dozens of such
repeated activations the present DC-MEG plus ICA approach allows to moni-
tor the time course of cerebral DC-activations without any need for averaging
(Fig.2). In principal this is a first step towards "on-line" brain monitoring pro-
viding a chance for single trial, resp. single event analysis.
It shall be emphasized that the component of interest in the present paradigm had only rank 10 in a list ordered according to the L2-norm of component power. Since many of the ICA components with larger power show up with maps featuring spatially coherent fields (i.e. they did not resemble random sensor noise patterns) a further physiological analysis of possibly underlying biological sources (cf. section 1) can be reasonably based on such ICA-based component definitions.

(2) When considering this DC-MEG scenario as a testbed for evaluating the robustness and validity of ICA in general and several ICA implementations in particular, the TDSEP approach appears remarkable in its performance under two test-the-limits conditions: (A) TDSEP maintains a high correlation between the time course of the signal (here known a priori) and of the best matching ICA component (Fig. 4) even for substantial reductions in the number of PCA components which were retained after subspace projection; the latter is intended to remove the noise subspace but inadvertently could remove also components of the signal under study. (B) When keeping the number of channels fixed (49) but reducing the number of data samples entering the ICA, TDSEP showed a graceful degradation for the correlation of its best matching component with the target signal time course (Fig.5).

Concluding, the two most pressing problems which arise when sending algorithms developed under mathematically ‘clean’ assumptions to ‘dirty’ real-world scenarios (spoiled by noise and simultaneously limited also in observation time) were sufficiently handled by the TDSEP version of ICA for the case of DC-magnetoencephalography. Hence the conjunction of DC-MEG and ICA holds a promising potential for assessing slowly varying neuroelectric brain processes both in health and disease, in particular concerning stroke patients.

Acknowledgments

We thank Guido Nolte and Martin Burghoff for many fruitful discussions. A.Z. was partly funded by DFG under contracts JA 379/52 and JA 379/71. G.W., B.-M. M. and G.C. were partly funded by DFG grants Cu 36/1-1 and 1-2.
References


17
Figure 3: Spatial field patterns, waveforms and frequency contents of the first ten components obtained by TDSEP sorted according to the L2-norms. For units and details of ICA10 cf. Fig. 2.
Figure 4: A PCA projection to a given number of components is performed prior to ICA in this subspace. We show the correlation coefficient between stimulus and the best matching ICA component vs number of components. The correlation to the best matching PCA component is shown as a baseline.

Figure 5: Correlation coefficient between stimulus and the best matching ICA component vs number of samples used for TDSEP applied on the full 49-dimensional sensor space.