## The Effect of Mutual Information on Independent Component Analysis in EEG/MEG Analysis: A Simulation Study

A. Neumann [a]; M. Grosse-Wentrup [b]; M. Buss [b]; K. Gramann [c]

[a] Neuro-Cognitive Psychology Program, Ludwig-Maximilians-University, Munich, Germany [b] Institute of Automatic Control Engineering (LSR), Technische Universität München, Theresienstr, Munich, Germany [c] Swartz Center for Computational Neuroscience, University of California, San Diego, California, USA

Online Publication Date: 01 November 2008

## PLEASE SCROLL DOWN FOR ARTICLE

**informa**
healthcare

# THE EFFECT OF MUTUAL INFORMATION ON INDEPENDENT COMPONENT ANALYSIS IN EEG/MEG ANALYSIS: A SIMULATION STUDY

**A. NEUMANN**

Neuro-Cognitive Psychology Program
Ludwig-Maximilians-University
Munich, Germany

**M. GROSSE-WENTRUP**
**M. BUSS**

Institute of Automatic Control Engineering (LSR)
Technische Universität München
Theresienstr, Munich, Germany

**K. GRAMANN**

Swartz Center for Computational Neuroscience
University of California
San Diego, California, USA

*Objective*: This study investigated the influence of mutual information (MI) on temporal and dipole reconstruction based on independent components (ICs) derived from independent component analysis (ICA). *Method*: Artificial electroencephalogram (EEG) datasets were created by means of a neural mass model simulating cortical activity of two neural sources within a four-shell spherical head model. Mutual information between neural sources was systematicallyvaried. *Results*: Increasing spatial error for reconstructed locations of ICs with increasing MI was observed. By contrast, the reconstruction error for the time course of source activity

1534

was largely independent of MI but varied systematically with Gaussianity of the sources. *Conclusion*: Independent component analysis is a viable tool for analyzing the temporal activity of EEG/MEG (magnetoencephalography) sources even if the underlying neural sources are mutually dependent. However, if ICA is used as a preprocessing algorithm for source localization, mutual information between sources introduces a bias in the reconstructed locations of the sources. *Significance*: Studies using ICA-algorithms based on MI have to be aware of possible errors in the spatial reconstruction of sources if these are coupled with other neural sources.

## INTRODUCTION

Independent Component Analysis (ICA) has become a frequently used tool in the analysis of electroencephalogram/magnetoencephalography (EEG/MEG) data. The general idea underlying the use of ICA is that by unmixing the measured scalp potentials into independent components (ICs), the analysis of a single neural source can be decoupled from all other sources. This is utilized in a variety of applications, such as the single-trial analysis of brain dynamics (Jung et al., 2001), or as a preprocessing step for source localization algorithms (Zhukov et al., 2000). Recently, a concurrent EEG/fMRI (**functional Magnetic Resonance Imaging)** recording demonstrated that ICA can reconstruct spatial sources with high accuracy and that the major temporal characteristics of ICs might be directly coupled to the BOLD-signal in a certain brain region (Debener et al., 2005).

For ICA to be applicable to a given data set several basic requirements have to be fulfilled (Comon, 1994). The postulation of *linearity* assumes the measured electric/magnetic fields on the scalp to be formed by linear superposition of the fields generated by the neural sources. *Stationarity* assumes the probability density function (PDF) of the temporal activity of the neural sources to be time invariant. The assumption of *complete bases* supposes that the number of neural sources in the dataset may not outnumber the sensors used for recording the data. Finally, *statistical independence* assumes that the fields generated by the neural sources have to be mutually independent.

The linearity of the model is fulfilled in EEG/MEG analysis if only spatially stationary sources are present (Nunez, 2005). This is in agreement with the usual modelling of electric/magnetic fields of cortical columns by spatially fixed current dipoles (Baillet et al., 2001). The stationarity of the model can be

achieved by randomly shuffling the data points in time.This leads to a temporal activity having the same PDF over time, which can therefore be referred to as quasi-stationarity. The assumption of having at least as many sensors as neural sources is questionable in many practical applications, but will not be considered further in this study (for a discussion see Grosse-Wentrup & Buss, 2007; Li et al., 2006). The mutual independence assumption of neural sources is questionable and might not be fulfilled in EEG/MEG analysis (Jung et al., 2001) since neural generators might be closely coupled in their activity over time. Whenever activity of two neural sources systematically covaries over time, information about the PDF of the temporal activity of one source can be inferred from knowledge about the PDF of the temporal activity of the other source. This relation can be stated very adequately in the mathematical term of mutual information (MI) (Cover, 1991). Yet, for many neuroscientific problems, it is important to gain a stable and accurate reconstruction even of sources that have a large overlap in their activity with other sources. Therefore, it is of interest to investigate how the reconstruction of EEG/MEG sources using ICA is affected, if the neural sources are not mutually independent.

## METHOD

In the present study, we investigated how certain degrees of MI influence the spatial and temporal reconstruction of neural sources using ICA. For this purpose, a platform for creating artificial EEG datasets was developed, allowing for a direct control of several aspects like the amount of MI between neural sources or the amount of noise introduced to the data. To investigate the effect of MI on the quality of source reconstruction of ICs resulting from Infomax-ICA (Bell & Sejnowski, 1995) the following procedure was applied:

1. Stimulus-related neural activity of two cortical columns was simulated using a neural mass model (David & Friston, 2003). Modelling MI was achieved by coupling the activity of the two columns. The degree of coupling was varied systematically in order to generate sources with different amounts of MI.
2. Artificial scalp potentials (forward solution) were calculated by placing two current dipoles in a four-shell spherical head model, which were assigned the temporal activity of the electric field generated by the two cortical columns (further referred to as the original neural sources). In addition, 126 dipolar sources with Gaussian noise activity were added to the source space. The

resulting scalp potentials were further distorted by adding Gaussian noise to each of the 128 sensors.

3. The forward calculated noisy scalp potentials were decomposed by means of Infomax-ICA (Bell & Sejnowski, 1995) and reconstructed using an equivalent dipole model as implemented in the DIPFIT toolbox (Oostenveld & Oostendorp, 2002). The reconstructions included the spatial location of the two sources as well as their time course of source activity.

*Neural Mass Model.*   The amount of MI was varied through the implemented neural mass model described by David & Friston (see David & Friston, 2003 for a detailed description). Neural mass models simulate biologically plausible signals by reducing the activity of millions of neurons to a few single-state variables, approximating the average activity in certain cortical areas or columns (mean field approximation) (Lopes da Silva et al., 1974). The core parts of the cortical columns are described in Figure 1, consisting of its input, its inner dynamics, and its output. The inner dynamics are determined by the temporal activity of several neural populations within the column.

In this model, three different neuronal populations [excitatory stellate cells (1), inhibitory interneurons (2), and excitatory pyramidal cells (3)] compose a cortical column. This structure aims at approximating real EEG data, since it emphasizes the dynamic interactions between different neural populations. Any present input stimuli alter the baseline activity of the model and therefore also its output activity over time. The output activity of a column represents its average rate of action potentials, computed by an average post-synaptic membrane potential. In addition, each active column produces an electric field, which forms the basis of the EEG potentials. Both, the electric field and the rate of action potentials, are based on the input stimuli and the inner dynamics of the column.

In order to generate neural activity of two sources, which are coupled by a certain degree of MI, two identical cortical columns were used. These were linked by feeding the average rate of action potentials produced by one column to the input of the other column with a time delay of 15 min. The strength of this connection was flexibly determined by a coupling parameter $K \in [0, 1]$. For $K = 0$ no coupling existed, and the dynamics of the two columns were independent. A coupling parameter of $K = 1$ represented full coupling, i.e., the dynamics of one column was completely determined by the activity of the other column.
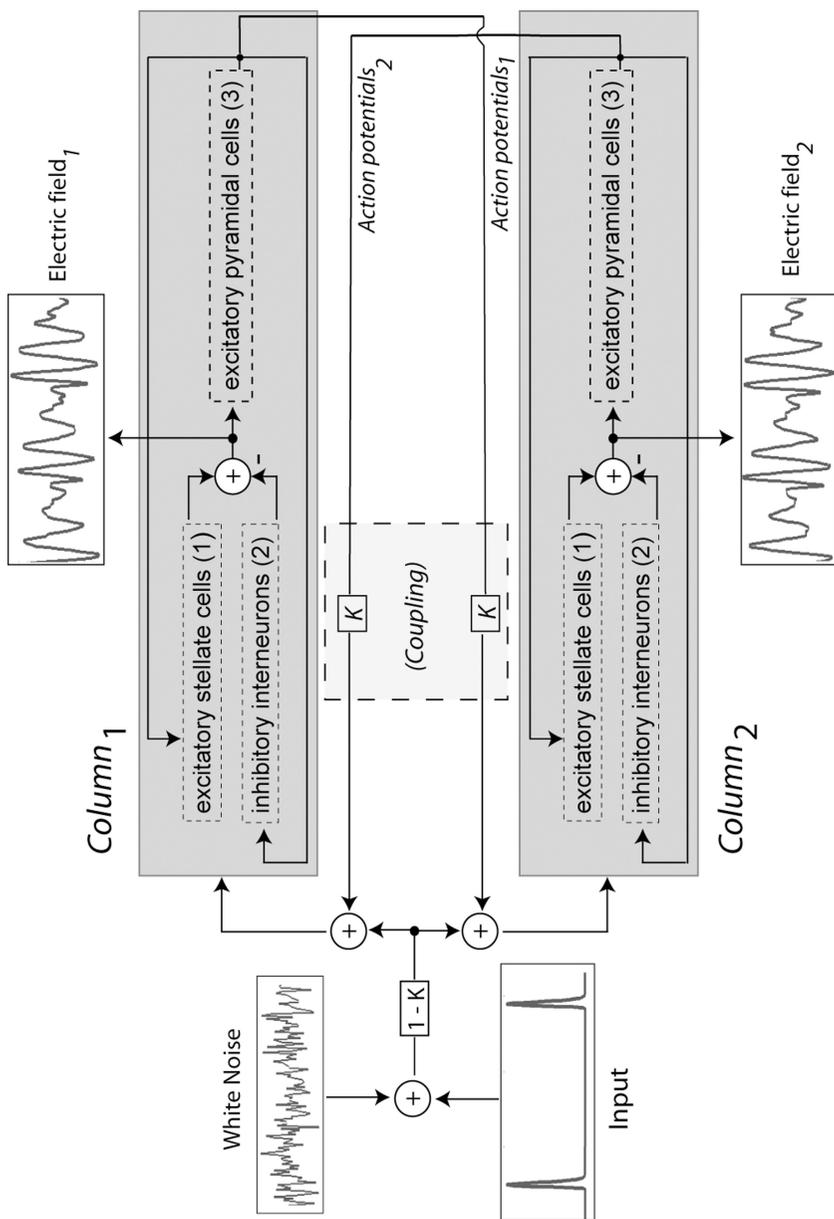
**Figure 1.** Neural populations and connections of the two cortical columns, forming the neural mass model.
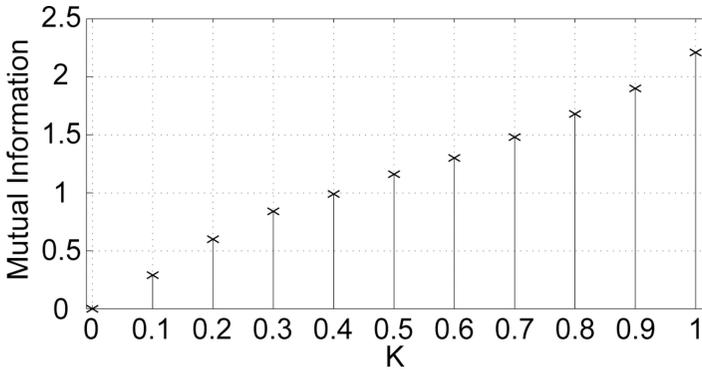
**Figure 2.** Mutual information of original neural sources for all datasets.

The input for the two cortical columns consisted of an impulse-like function with additional white noise activity (David & Friston, 2003). The parameters were adjusted in a way that the function peaked every 1.2 s. Each of these peaks simulated a neural event, whereas the white noise represented unspecific neural background activity. The first 200 min before the onset of each impulse were considered as a baseline consisting only of background noise. A total of 200 consecutive stimuli were generated, resulting in time series of 240 s. The stimulus structure led to an electric field of both columns, approximating event-related potentials (ERPs). These quasi ERPs were taken as the temporal activity of interest of the two original neural sources. This way eleven datasets of 240 s sampled at 200 Hz with K varying from 0 to 1 in steps of 0.1 were created. The resulting MI between sources for different coupling strengths was estimated using a histogram-based method (Moddemeijer, 1989) and is shown in Figure 2. It is important to point out that coupling between sources also altered the PDF of the temporal activity of the sources. More specifically, stronger coupling led to a less Gaussian distribution (Figure 3).

*Forward Solution.* For the calculation of artificial scalp potentials a four-shell spherical head model was used to map the locations, strengths, and orientations of the sources to the positions of 128 EEG channels. It consisted of four concentric and symmetric spheres (cortex, cerebrospinal fluid, skull, and scalp) with the radii of 71, 72, 79, and 85 mm respectively, and conductivity values of 0.33, 1, 0.0042, and 0.33 1/Ω respectively. Each neural source was modelled by a current dipole placed inside the head model. For each dipole, this resulted in a topography vector, defined by a location and
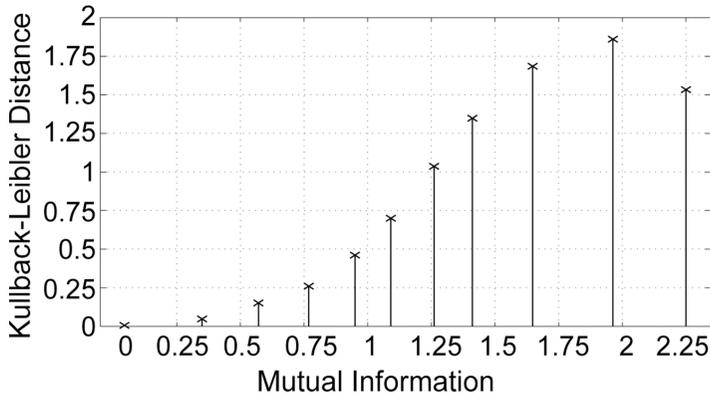
**Figure 3.** Kullback-Leibler distance of the temporal activity of the original neural sources to maximum likelihood Gaussian distribution (zero meaning Gaussian distribution).

an orientation, describing the dipole's projection to each electrode. Two dipoles, corresponding to the original sources, were placed symmetrically in the left and right hemisphere of the most inner sphere representing the cortex. Furthermore, 126 dipoles with Gaussian distributed activity N $(0, \sigma^2)$ were spatially equally distributed on a sphere with the same radius as the two original sources. The variance of the dipoles with Gaussian activity was chosen to equal 5% of the variance of the original neural sources. These dipoles served as noise sources, representing nonstimulus-related activity of the cortex. All dipoles were oriented orthogonal to the scalp. Based on the linearity of volume conduction, the corresponding scalp potentials at 128 electrode sites (equidistant montage) were computed accordingly(Cuffin and Cohen, 1978). The resulting scalp potentials were obscured with sensor noise for which variance was set to 1% of the variance of the noiseless electrode potentials. The scalp potentials were calculated for all simulated time series of neural activity. This resulted in eleven datasets for varying degrees of MI between the two original neural sources.

*ICA Reconstruction.* Due to the popularity in the EEG/MEG community, the extended Infomax-algorithm (Lee et al., 1999), as implemented in EEGLAB (Delorme & Makeig, 2004), was used for estimating the original neural sources. The Infomax-algorithm (Bell & Sejnowski, 1995) was applied to each of the eleven generated datasets, resulting in estimates of the topography and the temporal activity of the underlying sources for each coupling strength. Note

that the Infomax-algorithm aims to maximize temporal independence, even if the underlying sources are not independent and returns as many source and topography estimates as sensors being used (Bell & Sejnowski, 1995). Consequently, it was necessary to identify the two ICs representing the original sources from the 128 estimated ICs. This was done by utilizing the fact that the Infomax-algorithm is blind to Gaussian sources. If applied repeatedly to the same dataset with randomized initial conditions, only the non-Gaussian sources are reconstructed consistently. All sources with Gaussian distribution on the other hand are randomly mixed together, and thus cannot be consistently reconstructed (Grosse-Wentrup & Buss, 2006). Since the two original sources were the only sources with a non-Gaussian PDF, the two ICs that could be reconstructed most consistently represented the two original neural sources. The actual procedure used for measuring the consistency of each IC is described in (Grosse-Wentrup et al., 2005). A total of 20 repititions of the Infomax-algorithm were applied to each of the datasets to identify the original neural sources. The topographies of the two ICs representing the original neural sources were used to independently estimate the origin of each original neural source in the four-shell spherical head model with a single current dipole as described in (Delorme & Makeig, 2004).

## RESULTS

The major results of the present study can be outlined as follows: Higher MI led to a monotonic increase of the spatial reconstruction error. This was primarily due to a biased reconstruction of source locations toward the centre of the head. Reconstruction of the temporal activity of the neural sources, on the other hand, was not strongly influenced by MI between sources, but was primarily determined by the PDF of the neural sources. Better temporal reconstruction was achieved for less Gaussian sources. Consequently, good temporal reconstructions were obtained in spite of high MI.

The resulting positions of the reconstructed original neural sources for the different coupling strengths and the corresponding mean spatial reconstruction errors are shown in Figure 4.

As can be seen in Figure 4a, the spatial reconstruction error increased with MI between the sources. The lowest spatial reconstruction error of 4.5 mm was attained for no coupling ($K = 0$), and the highest reconstruction error of 32.8 mm for full coupling ($K = 1$). Figure 4b shows the reconstructed and the true locations of the sources in the spherical head model. Note that the reconstructed locations moved toward the centre of the head with increasing
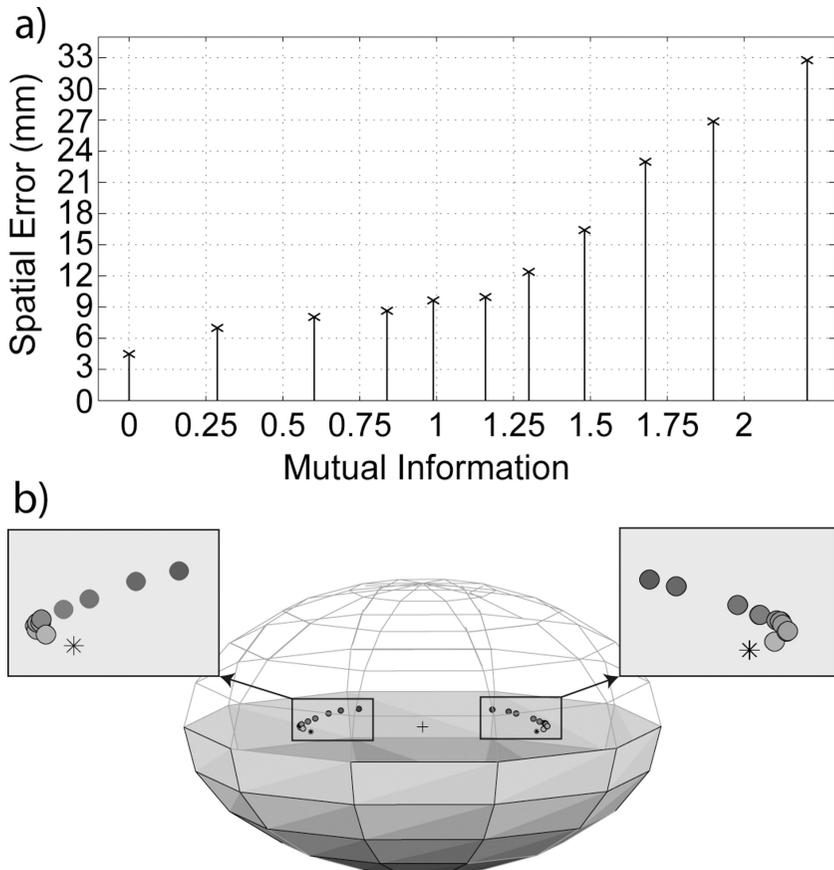
**Figure 4.** (a) Spatial reconstruction error. (b) Original (stars) and reconstructed (circles) positions of the two original neural sources. The darker the circles the higher the mutual information.

MI. Even though it is a general characteristic of all dipole source reconstruction algorithms to tend to model noisy data deeper inside the head than they actually are, this offers an additional explanation for the observed tendency of ICA to reconstruct sources too deep inside the head (Makeig et al., 2004): if sources are not independent, the topography of an IC will be slightly mixed with the topography of another source. This leads to less focussed, broader source topographies, which, due to volume conduction, can be better explained by dipoles deep inside the cortex (Nunez, 2005).

The accuracy of the reconstructed-activity time course was evaluated by calculating the signal-to-noise-ratio (SNR) between the true and the
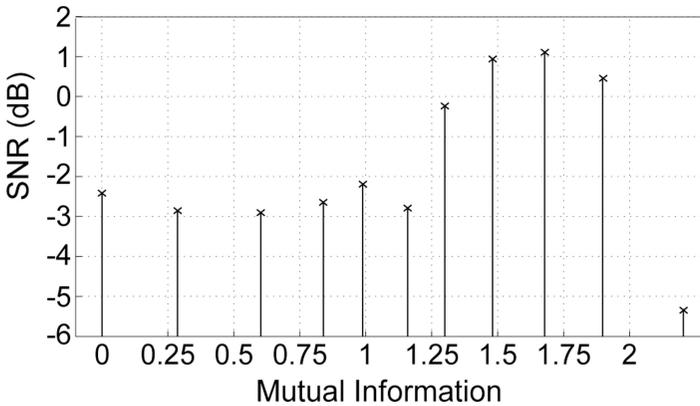
**Figure 5.** Signal-to-noise ratios of the temporal reconstructions of the original neural sources for all datasets.

reconstructed temporal activity. The SNR, measured in dB, was calculated as 10 times the logarithm of the ratio of the variance of the true temporal activity and the variance of the error of the true and the reconstructed temporal activity. Due to scale indeterminacy of ICA the temporal activity of the original and the reconstructed sources was normalized before computing the SNR. The resulting SNR for each of the coupling strengths is shown in Figure 5.

Regarding the quality of the temporal reconstruction, it can be seen in Figure 5 that the SNR does not primarily depend on the coupling strength. More specifically, the highest SNR is achieved for a large MI and strong coupling ($K = 0.8$). For full coupling ($K = 1$), however, the sources cannot be reconstructed, as can be seen by a steep drop in SNR for the highest amount of MI.

These observations can be explained by the following considerations. Theoretically, ICA requires mutually statistically independent sources. Yet, the Infomax-algorithm separates sources by only minimizing the MI between them, regardless of whether it can actually be reduced to zero or not. If the original sources are indeed not statistically independent, this procedure can still lead to good estimates of the temporal source activity. On the other hand, the steep drop in SNR for full coupling ($K = 1$) can be explained by the fact that in this case the relationship between the temporal activity of the two sources becomes deterministic and therefore both have precisely the same temporal activity. It is furthermore interesting to point out that the quality of the temporal reconstruction correlates with the PDF of the original neural sources for the

different coupling strengths. The less Gaussian the original neural sources, the better the temporal reconstruction by ICA (as long as the relationship between the sources is not deterministic). This is explained by the fact that minimization of MI is strongly related to maximization of non-Gaussianity (Cardoso, 2003). Consequently, the less Gaussian the original neural sources, the more easily these sources can be reconstructed by ICA algorithms based on MI such as the Infomax-algorithm.

## DISCUSSION

Summarizing the results, a strong correlation between MI and spatial reconstruction error was found. This reconstruction error was manifested by a shift of the reconstructed source locations toward the centre of the head with increasing MI between sources. The temporal reconstruction, on the other side, was largely independent of MI between sources as long as their relationship was not deterministic. Instead, the quality of the temporal reconstruction was correlated with the non-Gaussianity of the original sources. Higher non-Gaussianity of the original sources led to a better reconstruction of their temporal activity.

These results should be considered for the utilization of ICA in the analysis of EEG/MEG data. If ICA-algorithms based on MI are used, the quality of the temporal reconstruction is only marginally affected by coupling between neural sources. Regarding the spatial reconstruction, an error is likely to be introduced into the estimated location of sources if these are coupled with other neural sources. However, it has to be noted that the results presented here depend at least partly on the parameter settings used like the ratio between sensor noise variance and noiseless electrode potentials. This ratio might change over recording time and between channels. Further, the simulation of only two bilateral sources is not reflecting the high number of cortical and artifactual EEG sources with non-Gaussian activity distributions observed in real EEG-recordings. Nonetheless, the basic approach used in the present study focused on a minimal model to systematically analyse the influence of mutual information on the accuracy of the time course and spatial reconstruction of ICs derived from ICA. Further simulation studies using varying model parameters and numbers of sources to systematically analyze the influence of mutual information in ICA decompositions are necessary.

## REFERENCES

Baillet, S., Mosher, J. C., & Leahy, R. M. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, *18*, 14–30.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7*, 1129–1159.

Cardoso, J.-F. (2003). Dependence, correlation and Gaussianity in independent component analysis. *Journal of Machine Learning Research*, *4*, 1177–1203.

Comon, P. (1994). Independent component analysis, a new concept. *Signal Processing*, *36*, 287–314.

Cover, T. M., & Thomas, J. A. (1991). *Elements of Information Theory*. New York: Wiley Series in Telecommunications.

Cuffin, B. N., & Cohen, D. (1978). Comparison of the magnetoencephalogram and electroencephalogram. *Electroencephalography and Clinical Neurophysiology*, *47*, 132–146.

David, O., & Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage*, *20*, 1743–1755.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram ad functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience*, *25*, 11730–11737.

Delorme, A., & Makeig, S. (2004). EEGLab: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9–21.

Grosse-Wentrup, M., & Buss, M. (2007). Overcomplete independent component analysis via linearly constrained minimum variance filtering. *Journal of VLSI Signal Processing*, *48*, 161–171.

Grosse-Wentrup, M., & Buss, M. (2006). Signal subspace identification through blind source separation. *IEEE Signal Processing Letters*, *13*, 100–103.

Grosse-Wentrup, M., Gramann, K., Wascher, E., & Buss, M. (2005). EEG source localization for brain-computer interfaces. *Proceedings of the 2nd International IEEE EMBS Conference on Neural Engineering*, Arlington, Virginia, 128–131.

Jung, T. P., Makeig, S., McKeown, M. J., Bell, A. J., Lee, T.-W., & Sejnowski, T. J. (2001). Imaging brain dynamics using independent component analysis. *Proceeding of the IEEE*, *89*, 1107–1122.

Lee, T.-W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax agorithm for mixed sub-Gaussion and super-Gaussion sources. *Neural Computation*, *11*, 606–633.

Li, Y., Cichocki, A., & Amari, S.-I. (2006). Blind estimation of chanel parameters and source components for EEG-signals: a sparse factorization approach. *IEEE Transaction on Neural Networks*, *17*, 419–431.

Lopes da Silva, F. H., Hoeks, A., Smits, H., & Zetterberg, L. H. Model of brain rhythmic activity. The alpha-rhythm of the thalamus. *Kybernetik*, *15*, 27–37.

Makeig, S., Delorme, A., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2004). Electroencephalographic brain dynamics following visual targets requiring manual responses. *PLoS Biology*, *2*, 747–762.

Moddemeijer, R. (1989). On estimation of entropy and mutual information of continuous distributions. *Signal Processing*, *16*, 233–246.

Nunez, P. L. (2005). *Electric Fields of the Brain*. New York: Oxford University Press.

Oostenveld, R., & Oostendorp, T. F. (2002). Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull. *Human Brain Mapping*, *17*, 179–192.

Zhukov, L., Weinstein, D., & Johnson, C. (2000). Independent component analysis for EEG source localization–An algorithm that reduces the complexity of localizing multiple neural sources. *IEEE Engineering in Medicine and Biology Magazine*, *19*, 87–96.