Journal of Neuroscience Methods 183 (2009) 63–71

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Three-way matrix analysis, the MUSIC algorithm and the coupled dipole model

J.C. de Munck[a,*], F. Bijma[b]

[a] VU University Medical Centre, Dept. Physics and Medical Technology, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands
[b] VU University, Faculty of Sciences, Dept. Mathematics, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 11 May 2009
Received in revised form 28 June 2009
Accepted 30 June 2009

Keywords:
MEG/EEG
Inverse model
MUSIC
Three-way matrices
PARAFAC
Dipole estimation

ABSTRACT

The inverse problem of multi-channel MEG/EEG data is considered as a parameter estimation problem. The stability of the solution of the inverse problem, which decreases with the number of included dipoles, can be improved by either adding constraints to the model parameters, or by adding more data of related data sets. The latter approach was taken by Bijma et al. [Bijma F, de Munck JC, Böcker KBE, Huizenga HM, Heethaar RM. The coupled dipole model: an integrated model for multiple MEG/EEG data sets. NeuroImage 2004;23(3):890–904; Bijma F, de Munck JC, Huizenga HM, Heethaar RM, Nehorai A. Simultaneous estimation and testing in multiple MEG data sets. IEEE Trans SP 2005;53(9):3449–60] by introducing coupling matrices that link dipole parameters and source time functions of different data sets.

Here, the theoretical foundations of the coupled dipole model are explored and the MUSIC algorithm is generalised to the analysis of multiple related data sets. Similar to the MUSIC algorithm, the number of sources and the number of constraints are derived from the data by considering the minimum possible residual error as a function of the number of sources and constraints. However, contrary to the MUSIC algorithm, where the minimum residual error can be obtained from an SVD analysis of a two-way data matrix, here we deal with multiple data sets and therefore three-way matrix analysis is used.

From a simulation study it appears that the number of sources and constraints can be clearly determined from a generalised SVD analysis. The generalisation of the MUSIC algorithm to three-way data gives reasonable estimates of the dipole parameters. These results can be used in the simultaneous analysis of MEG/EEG data of multiple subjects, multiples data sets of the same subject or models where subsequent trials of data show habituation effects.

1. Introduction

The technological developments leading to fMRI and PET, with their sub-centimetre spatial resolution, have not reduced the clinical and scientific interest in localisation of brain activity using multi-channel MEG and EEG. Due to the superior temporal resolution of MEG and EEG, electric brain activity can be recorded directly, in the form of rhythms, waves and inter-ictal spikes; phenomena that are not accessible with other brain imaging techniques. Also, communication between different parts of the brain takes place on a milli-second time scale, and, therefore, in order to study connectivity in the brain, MEG and EEG remain indispensable (e.g. Grasman et al., 2004). For example, when studying the response of the brain due to a certain stimulus, fMRI and PET are unable to provide information on the order in which the sources are activated and hence, to study these aspects of brain responses, MEG and EEG are the modalities of choice.

The problem of localising sources requires a mathematical model that explicitly describes recorded MEG/EEG signals in terms of the electric generators. Usually, these generators are described with mathematical current dipoles. When a large part of the cortex is (homogeneously) activated, a current dipole well describes its centre of gravity (de Munck et al., 1988). A more challenging problem is to determine how many dipoles should be incorporated into the model (e.g. Waldorp et al., 2002, 2005), and to describe how these dipoles behave over time. The more dipoles are included in the model, the better the goodness of fit, but also, the more unstable the solution, and hence the larger their confidence intervals. The instability can be reduced by adding more constraints, such as the assumption that dipole positions and orientations do not change over time (Scherg and Von Cramon, 1985; de Munck, 1990).

Because of its complexity, several strategies have been presented in the literature to attack the source localisation problem. Amongst them, minimum norm solution methods have been put forward to circumvent the a priori choice of the number of dipoles and the multiple local minima problem that play a role in dipole fitting (Sarvas, 1987). Minimum norm solutions are based on the assumption of a large number of dipoles with unknown amplitudes, located on a fixed regular grid or a geometrical model of the grey matter

* Corresponding author. Tel.: +31 (0) 204 440169.
E-mail address: jc.demunck@vumc.nl (J.C. de Munck).

0165-0270/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2009.06.040

(Dale and Sereno, 1993). Since such models are underdetermined, maximum smoothness constraints are often applied to obtain a unique solution (Pascual-Marqui et al., 1994). Minimum norm solutions have also been combined with the data obtained from similar experiments performed using fMRI. In Daunizeau et al. (2005) this strategy was presented in a formal Bayesian framework.

Another alternative is to use so-called beamformers (Van Veen et al., 1997; Robinson and Vrba, 1999; Sekihara et al., 2005; Hillebrand et al., 2005). Beamformers are originally devised as spatial filters, converting recorded MEG/EEG signals at selected time samples or in a certain frequency band, into a 3D image. This image represents the theoretical filter output at all grid points, assuming that the filter outputs were maximally focussed on that point. However, the image can also be interpreted as the dipole strength estimate at each point, assuming that there is only a single dipole active and that background noise is spatially correlated over sensors. In both models, beamformers and single dipole fitting in correlated noise, the expressions for the estimators are equal. Although different interpretations of beamformers are possible (Mosher et al., 2003; Wipf and Nagarajan, 2007), beamformers have shown to be quite successful in the localisation of changes in rhythmic brain activity, induced by an external stimulus (e.g. Cheyne et al., 2003).

Our preferred strategy for the source localisation problem is to consider it as a "classical" parameter estimation problem, which can be placed in a solid ML framework. The advantage of this approach is that all assumptions, concerning forward model, noise and source model, are made explicit, which facilitates interpretation of the results. Also, the ML framework provides ways to compute confidence intervals of all estimated parameters, which is much less precisely defined with minimum norm or beamformer solution methods. Nevertheless, the dipole fitting approach is not without problems, for instance due to the highly non-linear dependence of the cost function on the source position parameters. This non-linearity creates a multiple minima problem and instabilities of which the occurrences are hard to predict, when the method is applied on a particular data set. One example, that many have encountered, is that a model consisting of two stationary dipoles may give rise to a solution where the two dipoles are very close and almost oppositely oriented.

Slotnick et al. (1999) and, independently Bijma et al. (2004, 2005), have presented a novel way to solve instability problems related to the dipole fitting approach. The idea is to constrain the models, by making the assumption such as that different related data sets share common underlying generators, which may or may not have the same amplitude time functions. In Bijma et al. (2004) the model is formulated in terms of coupling matrices, which connect all of the available dipoles to all of the available source time functions. Each data set has a different coupling matrix, but dipoles and source time functions are constant over data sets. In this sense, the *coupled dipole model* (CDM) can be considered as a model of three-way matrices (with space, time and data set number as indices). When the coupling matrices are chosen as diagonal matrices, the structure of the coupled dipole model is the same as that of a so-called PARAFAC model (Möcks, 1988; Kroonenberg, 1983).

With the conventional stationary dipole model, the problem of determining the appropriate number of sources can be addressed by performing a SVD on the data matrix. This is a very important element of the MUSIC algorithm, which separates the available space into signal space and noise space. It appears that with the coupled dipole model, a similar analysis is possible, although the SVD must be replaced by another algorithm because the coupled dipole model describes three-way matrices, instead of ordinary matrices. In the next sections the CDM framework of Bijma is placed in a theoretical perspective and its relation to PARAFAC modelling and the MUSIC algorithm is discussed.

2. Parameter estimation

Here we consider the situation that we are interested in the estimation of parameters $\mathbf{p} = (p_0, p_1, \ldots, p_{Z-1})^{\mathrm{T}}$ that cannot be measured directly. Instead, N measurements have been done, $\tilde{r}_n(\mathbf{p})$, $n = 0, \ldots, N-1$, that have a theoretical dependence on the parameters $\mathbf{p}$. Here N is the total number of repeated and simultaneous measurements of r_n. Therefore, in a practical situation, the index n may represent a multiple index of channel, time, data set or repeated trial. Furthermore, the measurements r_n of $\tilde{r}_n(\boldsymbol{p})$ are embedded in the noise $\eta_n(\mathbf{q})$, with a Gaussian distribution. The mean of $\eta_n(\mathbf{q})$ is $\mathbf{0}$, and its $N \times N$ covariance matrix $C()$ depends on unknown parameters $\mathbf{q} = (q_0, q_1, \ldots, q_{Z'-1})^{\mathrm{T}}$. Both the parameters of interest $\mathbf{p}$ and the noise parameters $\mathbf{q}$ are considered as deterministic, and they have therefore no covariance matrix defined. This situation, which is very general, can be expressed in formulas as

$$r_n = \tilde{r}_n(\boldsymbol{p}) + \eta_n(\boldsymbol{q}), \tag{1}$$

and

$$\boldsymbol{\eta} \propto N(\mathbf{0}, C(\boldsymbol{q})). \tag{2}$$

In the dipole fitting context, $\mathbf{p}$ represents the dipole parameters (positions, orientations, strengths, amplitude time functions, etc.), r represents MEG/EEG recordings as a function of channel, time sample and/or trial and $\mathbf{q}$ represents the parameters describing the distribution of the background activity. The maximum likelihood (ML) estimator of $\mathbf{p}$ and $\mathbf{q}$ can be obtained by solving the following maximisation problem and taking the arguments $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$ for which $Gain_{ML}(\mathbf{p}, \mathbf{q})$ is maximum (e.g. Seber and Wild, 1989).

$$Gain_{ML} = \max_{(\boldsymbol{p},\boldsymbol{q})} \frac{e^{-((\boldsymbol{r}-\tilde{\boldsymbol{r}}(\boldsymbol{p}))^T C(\boldsymbol{q})^{-1}(\boldsymbol{r}-\tilde{\boldsymbol{r}}(\boldsymbol{p})))/2}}{(2\pi \det(C(\boldsymbol{q})))^{1/2}}. \tag{3}$$

The estimates $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$ for which $Gain_{ML}(\mathbf{p}, \mathbf{q})$ yields the maximum value, correspond to the highest likelihood of the noise $\boldsymbol{\eta}$ ($=\mathbf{r} - \mathbf{r}(\mathbf{p})$), given all the model assumptions. This explains the name *maximum likelihood estimator*.

The advantage of an ML-estimator is its solid statistical foundation and the fact that all modelling assumptions are of an explicit mathematical nature, in the form of expressions for the model $\tilde{r}_n(\boldsymbol{p})$ and covariance $C(\mathbf{q})$. However, in a practical situation like dipole fitting, simultaneous estimation of the dipole parameters and noise parameters might be unfeasible, due to computational demands. In such situations, the simultaneous estimation problem can be split into an estimation problem of $\mathbf{p}$ and one for $\mathbf{q}$, that are solved separately (e.g. Waldorp et al., 2001). To clarify this, suppose that the MEG/EEG measurements have the form of multi trial evoked fields/potentials. In this case, index n represents three indices: channel i, sample j and trial k. When the assumptions are made that the stimulus evokes the same spatio-temporal pattern $\bar{r}_{ij}$ in each trial k and that the noise is independent over trials, the noise estimation problem would imply estimation of $\bar{r}_{ij}$ (irrespective of the underlying dipoles) and the covariance model. This approach was followed in Huizenga et al. (2002) and de Munck et al. (2002), where a Kronecker product model was used for the covariance. The ML estimate of the spatio-temporal pattern $\bar{r}_{ij}$ appears to be a simple average of the raw trial data, and an iterative scheme is to be applied to find the spatial and temporal covariances. The averaged data matrix $\bar{r}_{ij}$ is then still contaminated with noise, be it with noise level that scales with the number of trials N_{tr}. This data average is then treated as a new set of measurements, modelled as $\tilde{r}_{ij}(\boldsymbol{p})$ embedded in Gaussian noise with a known covariance matrix. The resulting parameter estimation problem is referred to as the *Generalised Least Squares* approach.

In the following this splitting of parameters will be assumed implicitly, and the measurements will represent averages over

trials. The GLS cost function becomes

$$\begin{aligned}\text{Cost}_{\text{GLS}} &= \min_{\boldsymbol{p}}(\boldsymbol{r}-\tilde{\boldsymbol{r}}(\boldsymbol{p}))^T N_{tr} C^{-1}(\boldsymbol{r}-\tilde{\boldsymbol{r}}(\boldsymbol{p})) \\ &= \min_{\boldsymbol{p}}(W_C\boldsymbol{r}-W_C\tilde{\boldsymbol{r}}(\boldsymbol{p}))^T(W_C\boldsymbol{r}-W_C\tilde{\boldsymbol{r}}(\boldsymbol{p}))\end{aligned}, \tag{4}$$

where W_C is a matrix such that

$$W_C^T W_C = N_{tr} C^{-1}. \tag{5}$$

Within the GLS framework, one can compute the covariance matrix of the parameter estimates using a linearization of the model around the parameter estimates:

$$\text{Cov}(\hat{\boldsymbol{p}}) = \frac{1}{N_{tr}}\left(\left(\frac{\partial\tilde{\boldsymbol{r}}(\hat{\boldsymbol{p}})}{\partial\boldsymbol{p}}\right)^T C^{-1}\left(\frac{\partial\tilde{\boldsymbol{r}}(\hat{\boldsymbol{p}})}{\partial\boldsymbol{p}}\right)\right)^{-1}. \tag{6}$$

There are alternative expressions in the literature (e.g. Huizenga et al., 2002; Seber and Wild, 1989, pp. 88–89), where a noise estimate factor is included. Such factor is required when the covariance C contains an unknown scaling factor that is to be derived from the data.

3. The spatio-temporal dipole model

When the above theory is applied upon the spatio-temporal dipole model, the measurement index n is represented by two independent indices i and j that represent channel and time respectively. Furthermore, the model $\tilde{r}_n(\boldsymbol{p})$ is based on L dipoles $(0, \ldots, L-1)$ with a fixed position and orientation over time ($\mathbf{p}_l$) and a time varying amplitude s_{lj}. With this notation, the generic parameter vector $\mathbf{p}$ above is split into $\mathbf{p}^{\text{T}} = (\mathbf{p}_0^T, s_{00}, s_{01}\ldots s_{0,J-1}, \ldots, \mathbf{p}_{L-1}^T, s_{L-1,0}, s_{L-1,1}\ldots s_{L-1,J-1})^{\text{T}}$ and Eq. (1) changes into

$$r_{ij} = \sum_{l=0}^{L-1}\psi_i(\boldsymbol{p}_l)s_{jl} + \eta_{ij} \quad \text{or} \quad R = \Psi(\boldsymbol{p}_{dip})S^T + \text{E}. \tag{7}$$

In this equation, the position and orientation parameters of all dipoles have been collected in a single vector $\mathbf{p}_{dip}$. In the notation of Eq. (7) and the rest of this chapter, it is implicitly assumed that index i runs from 0 to $I-1$, j from 0 to $J-1$, etc. In particular, R is a $I\times J$ matrix, Ψ is $I\times L$ and S is $J\times L$.

The $IJ\times IJ$ covariance matrix C becomes extremely large when the number of sensors and the number of time points are more than 100, which is typically the case. Therefore, the spatio-temporal covariance matrix is modelled as a Kronecker product of a spatial X and a temporal matrix T,

$$C_{ij,i'j'} = X_{ii'}T_{jj'} \quad \text{or} \quad C = X\otimes T. \tag{8}$$

As sketched above, this covariance represents the covariance of the unaveraged data. For the dipole fitting part the GLS cost function reduces to

$$\begin{aligned}\text{Cost}_{\text{stat}} &= \min_{\boldsymbol{p}_{dip},S}\frac{1}{N_{tr}}Tr\{(R-\Psi(\boldsymbol{p}_{dip})S^T)^T X^{inv}(R-\Psi(\boldsymbol{p}_{dip})S^T)T^{inv}\} \\ &= \min_{\boldsymbol{p}_{dip},S}Tr\{(R_W-\Psi_W(\boldsymbol{p}_{dip})S_W^T)^T(R_W-\Psi_W(\boldsymbol{p}_{dip})S_W^T)\}\end{aligned}. \tag{9}$$

In the second line the data and model have been pre-whitened by the "square roots" of the covariance matrices X and T. Also, the scaling with the number of trials has been absorbed in this pre-whitening

$$\begin{cases} R_W \equiv W_X^T R W_T \\ \Psi_W \equiv W_X^T\Psi \\ (S_W)^T \equiv S^T W_T \\ W_X(W_X)^T \equiv X^{inv} \quad \text{and} \quad W_T(W_T)^T \equiv N_{tr}T^{inv} \end{cases}. \tag{10}$$

In the sequel this pre-whitening will be assumed implicitly and the lower index W will be omitted.

A lower bound for the cost function can be obtained using the SVD of the (pre-whitened) data matrix

$$R = U\,\Delta V^{\text{T}},$$

with

$$\begin{cases} U^TU = I_J, & I\geq J \\ UU^T = I_I, & I\leq J \end{cases}, \quad V^{\text{T}}V = VV^{\text{T}} = I_J \quad \text{and}$$

$$\Delta = \begin{pmatrix}\lambda_0 & 0 & 0 \\ & \ddots & \\ 0 & 0 & \lambda_{J-1}\end{pmatrix}. \tag{11}$$

By ignoring the dipole parameterisation and minimising over all matrices Ψ of dimension $I\times L$, it is found that (de Munck, 1990)

$$\text{Cost}_{\text{stat}} \geq \min_{\Psi,S}Tr(R-\Psi S^T)^T(R-\Psi S^T) = \sum_{l=L}^{J-1}\lambda_l^2. \tag{12}$$

The product ΨS^T for which this minimum is achieved is the best rank L approximation of R. Obviously, the more sources L are included in the model, the lower cost function can be made. In Mosher et al. (1992) the problem of estimating the number of sources is addressed by observing that

$$\frac{1}{I}E\{R^TR\} = \frac{1}{I}S\Psi^T\Psi S^T + I_J. \tag{13}$$

Therefore, the computation of the squared singular values of the (pre-whitened) data matrix, divided by the number of sensors I, will yield a decreasing array of values that have a limit of 1. When the source time functions S are independent, the L-dimensional space corresponding to the L largest singular values corresponds to the signal space and its complement to the noise space (Schmidt, 1986). This separation of spaces can be used to define a projection matrix P_{signal}:

$$P_{signal} = (u_0, \ldots, u_{L-1})(u_0, \ldots, u_{L-1})^T, \tag{14}$$

where u_l represents the lth column of the matrix U in Eq. (11). This projector leads to an alternative cost function that can be minimised,

$$\text{Cost}_{\text{MUSIC}} = \min_{\boldsymbol{p}_{dip}}\left(1-\frac{\left|P_{signal}\Psi(\boldsymbol{p}_{dip})\right|^2}{\Psi^T(\boldsymbol{p}_{dip})\Psi(\boldsymbol{p}_{dip})}\right). \tag{15}$$

The minimisation of this cost function must be performed such that the matrix Ψ has rank L. Because $|.|$ is a Frobenius norm, this cost can be computed columnwise, and the rank L-restriction leads to a scan function of which all L local minima are to be found,

$$\text{Scan}(\boldsymbol{p}_l) = 1-\frac{\left|P_{signal}\Psi_{dip}(\boldsymbol{p}_l)\right|^2}{\Psi_{dip}^T(\boldsymbol{p}_l)\Psi_{dip}(\boldsymbol{p}_l)}. \tag{16}$$

Therefore, the advantage of this alternative cost function is that all local minima can be found by a global search of one dipole after another, over the area of interest, instead of varying all dipoles

simultaneously. More details of this so-called MUSIC scan, including the extension to time varying orientations, can be found in Mosher et al. (1992), Mosher and Leahy (1998) and Mosher and Leahy (1999).

4. The coupled dipole model

In practice the stationary dipole model appears to be unstable, even in the case of two dipoles, when the true dipoles are close to one another. Without any precautions, the ML solution will result in two dipoles with extremely large amplitudes, oppositely oriented and positions that are almost identical. A practical solution to this instability problem would be to add an artificial term to the cost function, which punishes closely located sources, or sources with extreme amplitudes. The problem with such practical solutions is that they are subjective in the sense that the solution will depend on the exact way the additional term is implemented. An alternative would be to apply a regional constraint on the dipole solution, e.g. that mutual distance should be larger than 1 cm. Such an approach would be less subjective, but in practice one would end up with solutions that precisely satisfy the constraint, and the solution would heavily depend on the 1 cm threshold.

The approach followed by Slotnick et al. (1999) and, independently, Bijma et al. (2004, 2005) was based on the notion that MEG/EEG experiments using evoked activity are often done in series with stimuli that are very much related to one another. For instance, the experimenter uses a visual stimulus with a varying contrast, spatial frequency or part of the visual field. In such series, it would be valid to assume that part of the sources or source time functions are identical over data sets. The so-called "coupled dipole model" (CDM) exploits these assumptions by making them explicit and imposing constraints on combinations of data sets. These combinations of data sets are analysed using an integrated model, instead of using separate spatio-temporal models for each individual data set.

The generic index n will now vary over channel i, time j and data set k. Over all data sets there are L dipoles available and M source time functions. The user imposes constraints on the data by specifying for each data set a coupling matrix C^k, of which the entries are zero when a certain combination of source time function and dipole is absent in data set k, and non-zero when that combination is present and must be estimated from the data. In a formula, this model can be expressed as

$$r_{ijk} = \sum_{l=0,m=0}^{L-1,M-1} \psi_i(\mathbf{p}_l) c^k_{lm} s_{jm} + \eta_{ijk} \quad \text{or} \quad R^k = \Psi(\boldsymbol{p}_{dip}) C^k S^T + \mathrm{E}. \qquad (17)$$

Here r_{ijk} are the matrix elements of a three-way matrix representing the averaged responses of data set k. When it is assumed that the noise is independent over trials and over data sets, and when it has a spatio-temporal covariance which is a Kronecker product of a spatial and a temporal part, the ML cost function can be expressed as

$$\mathrm{Cost}_{\mathrm{coupled}} = \min_{\boldsymbol{p}_{dip}, C^k, S} \sum_k Tr(R^k - \Psi(\boldsymbol{p}_{dip}) C^k S^T)^T (R^k - \Psi(\boldsymbol{p}_{dip}) C^k S^T) \qquad (18)$$

where the spatio-temporal pre-whitening is again assumed implicitly.

To illustrate the use of the CDM two examples will be given next. Consider three experiments, in which either the left visual half field were stimulated ($k=0$), or the right half field ($k=1$) or both ($k=2$). In an early phase of the responses one could assume that there are two sources active ($L=2$) and two source time functions ($M=2$). When a half field is stimulated, there is a strong contra-lateral source and a weak ipsi lateral source, matrices C_0 and C_1 are designed to link the left and right dipole to these contra and ipsi lateral source time functions. When both hemifields are stimulated, one could assume that the left and right hemifield responses are dominated by the contralateral sources with equal shapes of the sources time function s_{j0}. In this model, the coupling matrices read

$$C_0 = \begin{pmatrix} \alpha_0 & 0 \\ 0 & \alpha 1 \end{pmatrix} \quad C_1 = \begin{pmatrix} 0 & \alpha_2 \\ \alpha_3 & 0 \end{pmatrix} \quad C_2 = \begin{pmatrix} \alpha_4 & 0 \\ \alpha_5 & 0 \end{pmatrix} \qquad (19)$$

In experiment 0, the first row of S^T is the (strong) contra-lateral activity and the second row the (weak) ipsi lateral activity. In experiment 1 the contra/ipsi lateral assignment relative to the side of stimulus presentation remains identical, but relative to the brain the source activities are swapped over hemispheres. In experiment 2, full field stimulation, the ipsi lateral sources are ignored. This explains the non-zeroes in Eq. (19).

Note that the user only has to specify which matrix elements of the coupling matrices are zero, the others, denoted by $\alpha_0, \ldots, \alpha_5$, are estimated by fitting the model to the data. Therefore, the user only indicates which shapes of source time functions are equal, their overall amplitudes (the α parameters in Eq. (19)) are free parameters. This also implies that in Eq. (18) the source time functions need to be normalised to obtain an *identified* model. Note furthermore that if these data sets were analysed separately, the number of free parameters would be $KL(J+5) = 6J+30$, whereas in the coupled case, this number is only $L5+6+M(J-1) = 2J+14$. Bijma et al. (2004) have demonstrated using simulated and real data that the CDM yielded indeed much more stable solutions than the separate analysis of data sets.

Another example of the applicability of the CDM is a case where a series of experiments are done, where two sources are activated with the same source time functions, but different relative weights, depending on the experiment. One could think of a visual stimulus, with varying check-size, where with small checks the response would be dominated by the fovea, and with large cheque sizes the extra-striate areas would dominate. Intermediate cases would result in a mixture of both sources. In this model for e.g. five stimulus conditions, all coupling matrices are diagonal, as represented here:

$$C_0 = \begin{pmatrix} \alpha_0 & 0 \\ 0 & \alpha_1 \end{pmatrix} \quad C_1 = \begin{pmatrix} \alpha_2 & 0 \\ 0 & \alpha_3 \end{pmatrix} \quad C_2 = \begin{pmatrix} \alpha_4 & 0 \\ 0 & \alpha_5 \end{pmatrix}$$
$$C_3 = \begin{pmatrix} \alpha_6 & 0 \\ 0 & \alpha_7 \end{pmatrix} \quad C_4 = \begin{pmatrix} \alpha_8 & 0 \\ 0 & \alpha_9 \end{pmatrix}. \qquad (20)$$

Also in this case the CDM provides much more stability than the separate analysis.

An attractive aspect of the CDM is that it covers different kinds of models in a common framework. The user only has to specify the assumed number of sources, the assumed number of source time functions, and the assumed way in which sources and source time functions are coupled. The ML solution proceeds in all cases in the same way.

However, this assumed structure of the coupling matrices, with fixed zeroes might be considered as a very subjective aspect of this procedure. Also, there are many possible ways for the user to specify the coupling matrices, and each choice may lead to the presence of a new invariant transformation. This will also hamper the computation of confidence intervals of the parameters: if the model contains invariant transformations, the parameter covariance matrix (Eq. (6)) will be singular. For these reasons, Bijma et al. (2005) considered the special case that all matrix elements of the coupling matrix are non-zero and are estimated from the data. In that an case, named *extended* CDM, the invariant transformations are exactly defined by arbitrary regular $M \times M$ matrix G, which transforms C^k to C'^k and S

to S', as follows:

$$\begin{aligned} R^k &= \Psi(\boldsymbol{p}_{dip})C^k GG^{inv}S^T + \mathrm{E} \\ &= \Psi(\boldsymbol{p}_{dip})C'^k S'^T + \mathrm{E} \end{aligned} . \tag{21}$$

In this model, the following restrictions are imposed on the dimensions of the coupling matrices

$$M \le \min(LK, J) \quad \text{and} \quad L \le \min(MK, I). \tag{22}$$

With this variant of the CDM, only the number of sources and the number of source time functions need to be specified, and it is possible to derive the confidence intervals for each of the model parameters, when the indeterminacy caused by G is accounted for. The price to be paid for these advantages is the slight increase of the number of parameters, and the more cautious interpretation of the parameters (implying that the parameters C^k and S should always be considered in the combination $C^k S^T$). Yet it appears that the increased stability of the problem compared to the separate analysis is maintained (Bijma et al., 2005).

5. Relation to MUSIC

The coupled dipole model provides a solution for the instability problems occurring with multiple dipoles, but even the extended CDM suffers from the somewhat arbitrary choice one has to make for the number of dipoles and the number of source time functions. When this choice is made, the dipole estimates result from a multi-dimensional minimisation problem, which can be computationally expensive because of the non-linearity and the presence of multiple local minima. In the case of the stationary dipole model, both these problems can be tackled with the MUSIC-formalism, as sketched above. Here the question is addressed to what extent this formalism can be applied to the case of CDM.

The main difference between the stationary dipole model and the extended CDM is that in the latter model the data matrix is a three-way matrix. This is at the same time the main problem. To estimate the number of dipoles L, one could e.g. compute the sensor by sensor $(I \times I)$ covariance matrix COV^I as defined below:

$$\{COV^I\}_{i,i'} \equiv \frac{1}{JK} E\left\{\sum_{k,j} R^k_{ij} R^k_{i'j}\right\} = \frac{1}{JK}\left\{\Psi \sum_k (C^k S^T S C^{k,T})\Psi^T\right\}_{i,i'} + \delta_{i,i'}. \tag{23}$$

This equation shows that, similarly to Eq. (13), all the expected eigenvalues equal 1, except the first L eigenvalues, which are increased due to the presence of activated sources. A plot of the eigenvalues of COV^I would show a jump at the Lth eigenvalue, beyond which the eigenvalues would tend to a constant value of 1.

Alternatively, one could compute COV^J to determine the number of source time functions M in an analogous way:

$$\{COV^J\}_{j,j'} \equiv \frac{1}{IK} E\left\{\sum_{k,i} R^k_{ij} R^k_{ij'}\right\} = \frac{1}{IK}\left\{S \sum_k (C^{T,k}\Psi^T \Psi C^k) S^T\right\}_{j,j'} + \delta_{j,j'}. \tag{24}$$

The problem is that, contrary to the stationary dipole case, this jump detection does not naturally lead to a separation into a signal and a noise space. The reason is that using the best rank L approximation of the data matrix in the two-way case, does not maintain a straightforward generalisation to the three-way case.

To derive these spaces it seems better to derive a lower bound of the cost function, by ignoring the parameterisation of the forward model Ψ:

$$\mathrm{Cost}_{\mathrm{coupled}} \ge \min_{\Psi, C_k, S} \sum_k Tr(R^k - \Psi C^k S^T)^T (R^k - \Psi C^k S^T). \tag{25}$$

This lower bound depends on L and M, which are estimated from the eigenvalue plots of Eqs. (23) and (24). Since this lower bound cannot be obtained from a singular value decomposition, the lower bounds have to be computed iteratively, thereby making use of the fact that the cost function is quadratically in each of the matrices Ψ, C^k and S:

$$\begin{cases} \Psi_{\min} = \sum_k R^k S C^{k,T} \left(\sum_k C^k S^T S C^{k,T}\right)^{inv} \\ C^k_{\min} = (\Psi^T\Psi)^{inv}\Psi^T R^k S (S^T S)^{inv} \\ S_{\min} = \sum_k R^{k,T}\Psi C^k \left(\sum_k C^{k,T}\Psi^T\Psi C^k\right)^{inv} \end{cases} . \tag{26}$$

These three update formulas have to be applied sequentially, starting from an appropriate initial guess. Then, similar to the MUSIC formalism in the stationary model, what matters is the range space of $\Psi_{\min}$ and that can be obtained by an SVD on the estimated $\Psi_{\min}$:

$$\Psi_{\min} = U_\Psi \Delta_\Psi V^{\mathrm{T}}_\Psi, \tag{27}$$

yielding an analogues *ThreeWayMUSIC* scan as follows:

$$\mathrm{Scan}_{ThreeWayMusic}(\boldsymbol{p}_{dip}) = 1 - \frac{\left|P_\Psi \Psi(\boldsymbol{p}_{dip})\right|^2}{\Psi^T(\boldsymbol{p}_{dip})\Psi(\boldsymbol{p}_{dip})}, \tag{28}$$

where the projector P_Ψ is now derived from the first L columns of U_Ψ instead of from the SVD of the data matrix R, as was the case in Eq. (14). Similar to the MUSIC scan, the dipole estimation would now proceed by finding the first L local minima of the scan function.

6. Relation to PARAFAC

Another interesting special case of the CDM occurs when all coupling matrices are modelled as diagonal matrices, as in the example of Eq. (20). To relate that case to so-called PARAFAC (PARAllel profiles FACtor analysis) modelling of three-way matrices, the following notation is used:

$$C^k_{lm} = D_{kl}\delta_{lm}, \tag{29}$$

where δ_{lm} is the Kronecker delta. The CDM can now be expressed as

$$r_{ijk} = \sum_l \psi_i(\boldsymbol{p}_l) D_{kl} S_{jl} + \eta_{ijk}. \tag{30}$$

Apart from the dipole parameterisation of Ψ, the solution of this model is identical to the problem of finding the best possible approximation of a three-way matrix r_{ijk} in terms of a triple product of two-way matrices. This approximation of three-way matrices is usually called PARAFAC modelling (e.g. Kroonenberg, 1983; Kiers and Krijnen, 1991; Bro, 1997) and the problem appears in many disciplines where data reduction and interpretation is important. Also, interesting theoretical results have been obtained on PARAFAC modelling (Kruskal, 1989), in particular regarding the uniqueness of the solution when the noise term is vanishing. Contrary to the two-way case, where there are infinitely many ways to express a rank L matrix as a product of two matrices having one dimension L, in the three-way case there is essentially one way, when certain conditions are met (Kruskal, 1977; Ten Berge, 2002). Also, for least squares fitting, such as in PARAFAC modelling, the question of uniqueness in the three-way case is much more involved than in the two-way case.

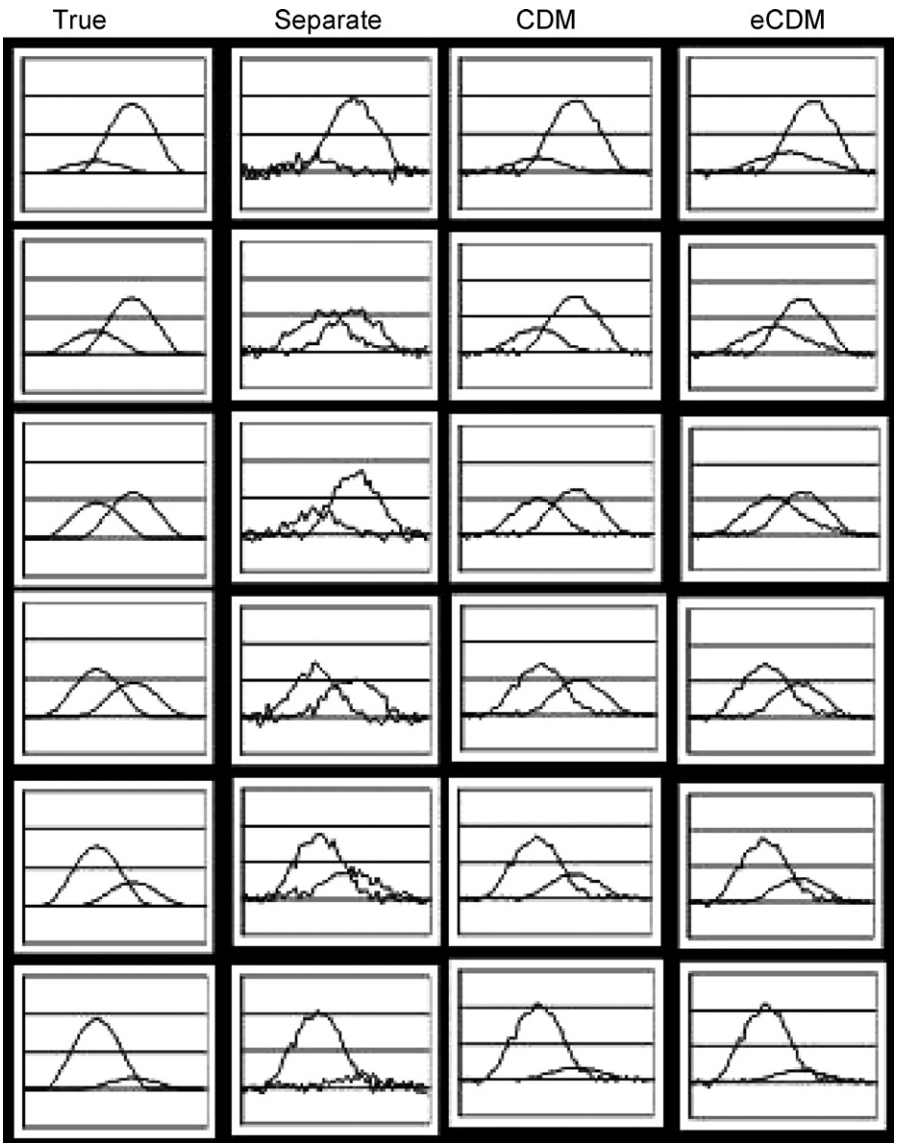

Fig. 1. The first column represents the true (simulated) source time functions for the six conditions. The second column represents the source time functions, estimated using the standard separate analysis, when the SNR is 1. The third and the fourth columns denote the estimated source time functions when the CDM and the ECDM are used. Note that the estimated source time functions of the separate analysis are much noisier than those of CDM and ECDM.

When a lower bound of the cost function is sought in case of diagonal coupling matrices, such bound can be found by determining the PARAFAC-fit:

$$\mathrm{Cost}_{\mathrm{Parafac}} \geq \min_{\Psi,D,S} \sum_{ijk} \left(r_{ijk} - \sum_{l} \{\Psi\}_{il} \{S\}_{kl} \{D\}_{jl} \right)^2, \tag{31}$$

where the factor matrices have to be solved iteratively:

$$\begin{cases} \{\Psi_{\min}\}_{il} = \sum_{jk,l'} r_{ijk} s_{jl'} d_{kl'} (S^T S \circ D^T D)^{inv}_{l'l} \\ \{S_{\min}\}_{jl} = \sum_{ik,l'} r_{ijk} d_{kl'} \psi_{il'} (D^T D \circ \Psi^T \Psi)^{inv}_{l'l} \\ \{D_{\min}\}_{kl} = \sum_{ij,l'} r_{ijk} \psi_{il'} s_{jl'} (\Psi^T \Psi \circ S^T S)^{inv}_{l'l} \end{cases}. \tag{32}$$

In this equation, the symbol $\circ$ represents the Hadamard product, i.e. the element wise product of the enclosing matrices.

The lower bound in Eq. (31) and its solution in Eq. (32) depend on the number of sources. Similar to the extended CDM we propose to compute the following covariance matrices from the data:

$$\begin{cases} \{COV^I\}_{i,i'} \equiv \frac{1}{JK} E\left\{ \sum_{jk} r_{ijk} r_{i'jk} \right\} = \frac{1}{JK} \{\Psi (S^T S \circ D^T D) \Psi^T\}_{i,i'} + \delta_{ii'} \\ \{COV^J\}_{j,j'} \equiv \frac{1}{KI} E\left\{ \sum_{ki} r_{ijk} r_{ij'k} \right\} = \frac{1}{KI} \{S (D^T D \circ \Psi^T \Psi) S^T\}_{j,j'} + \delta_{jj'} \\ \{COV^K\}_{k,k'} \equiv \frac{1}{IJ} E\left\{ \sum_{ij} r_{ijk} r_{ijk'} \right\} = \frac{1}{IJ} \{D (\Psi^T \Psi \circ S^T S) D^T\}_{k,k'} + \delta_{kk'} \end{cases}. \tag{33}$$

Contrary to the extended CDM (Eqs. (23) and (24)), this equation predicts that for PARAFAC models, the eigenvalues of *all three* covariances, will jump at the same index L. Beyond that jump the eigenvalues will tend to unity.

To estimate the underlying dipoles in the CDM with diagonal matrices, one could proceed by fitting a dipole on each of the columns of $\Psi_{\min}$, resulting from the iterative solution of Eq. (32). In the two-way case, such an approach would be unfruitful, because of the non-uniqueness of the rank L approximation. In the three-way case, this approach could be more successful, because the expansion of a three-way matrix as a triple product of two-way matrices is unique in many cases. An alternative is to perform a MUSIC scan, similar to the *extended* CDM (Eq. (28)), but now based on the $\Psi_{\min}$, resulting from Eq. (32). This approach will here be named *PARAFAC MUSIC*.

7. Example data set

The theory sketched above is illustrated with an example data set, wherein the underlying sources are known and the noise level is varied. Two dipoles are assumed, where one dipole has position $(-7.0, 1.5, 2.5)$ cm and orientation (0., 1., 0.) and the other has position $(-6.0, 4.0, 2.5)$ cm and orientation (0., 0., 1.). Co-ordinates are expressed in Nasion-Ear system, where the positive x-axis points to nasion, the positive y-axis points to the left ear and the positive z-axis points to the vertex. The given example dipoles correspond to points in the visual cortex. Dipoles have amplitude time functions defined by

$$\begin{cases} STF^{(1)} = !C_k^{(1)}(1+\sin(\pi(j-12.5)/15)); \quad C_k^{(1)} = k+1; \quad j=0,\ldots,49 \\ STF^{(2)} = C_k^{(2)}(1+\sin(\pi(j+7.5)/15)); \quad C_k^{(2)} = 6-k; \quad k=0,\ldots,5 \end{cases}, \tag{33}$$

see Fig. 1, first column. Here k refers to different data sets, with the same underlying dipoles and source time functions, but different relative strengths. This source configuration satisfies the CDM with diagonal coupling matrices. The MEG signals resulting from these time varying sources were simulated on the sensor positions of a 151-channel CTF MEG system. The MEG signals were embedded in uncorrelated Gaussian noise of different SNR levels: 1, 4 and 10. The SNR is defined as the power of the simulated signals (summed over channels and samples of the first data set) divided by the total power of the added noise. The same amount of noise was applied in the other data sets.

To compare the performances of the CDM with the traditional separate analysis, for each of the six data sets the stationary dipole solution was computed and for all data sets together the CDM solution was computed. Table 1 shows the mean dipole position errors and the mean residual errors of the separate analysis, and also the dipole position errors of the CDM. This table shows that the CDM has much smaller localisation errors. Although these data are based on a single noise realisation, other simulations confirm the superior performance of the CDM. Furthermore, it appears that the final costs are very close to their theoretical minima. Finally, it appears from other noise realisations that with CDM, amplitude outliers, consisting of two closely located dipoles with opposite orientations, are much more rare.

Table 1
A comparison of the dipole position errors of the CDM, ECDM and separate analysis. Each noise level is applied with a single noise realisation, but other realisations confirm that with CDM the dipole estimation errors are much smaller than for the separate analysis. The residual errors of the separate analyses are the mean over the 6 simulated data sets. The residual errors of CDM and ECDM are quite close to their theoretical minima (last column). Finally, also the theoretical lower bounds of CDM and ECDM are very close.

SNR	Analysis	Pos error 1 [cm]	Pos error 2 [cm]	Res error [%]	Min res error [%]
1	Separate	0.410	0.561	52.52	
	CDM	0.098	0.187	52.77	52.41
	ECDM	0.153	0.319	52.76	52.38
4	Separate	0.342	0.369	22.00	
	CDM	0.038	0.018	22.02	21.86
	ECDM	0.149	0.404	22.02	21.85
10	Separate	0.182	0.292	10.13	
	CDM	0.038	0.018	10.10	10.03
	ECDM	0.019	0.034	10.10	10.03

In the CDM analysis the correct source model was used: all coupling matrices were diagonal. In a real application, the correct model may be unknown and therefore we also analysed the simulated data with extended CDM, where the coupling matrices are unknown completely and fitted to the data. Table 1 shows that, despite the increased number of degrees of freedom, the performance of the ECDM is comparable to the CDM. Also, the residuals errors are comparable to those of CDM, and very close to their theoretical minima. The last column of Fig. 1 shows the reconstructed source time functions using ECDM, at an SNR of 1. It appears that with CDM and ECDM the estimated source time functions are much less noisy than with the separate analysis.

In the inverse part of the above simulations, the number of dipoles was assumed to be known. In the above theory, it was suggested to obtain this number of a plot of the singular values of COV^I, COV^J and COV^K (Eq. (33)) and to identify where all curves make a jump to a constant value of 1. In Fig. 2 these singular value plots were made for all three SNRs, on a double logarithmic scale. This figure shows that for our simulated case, the singular values behave as expected. Therefore, the use of these plots seems promising in practice. For the extended dipole model, these figures imply that both the number of sources and the number of time functions should be equal to 2.

A *PARAFAC MUSIC* scan was also performed for the simulated data at all three SNRs. The distances between the local minima of the *PARAFAC MUSIC* scan and the true dipole positions were generally larger than 2 cm. When a *ThreeWayMUSIC* scan was performed, using Eqs. (28), (27) and (14), the results look more promising. In the first place, the scan had much more contrast than the *PARAFAC*

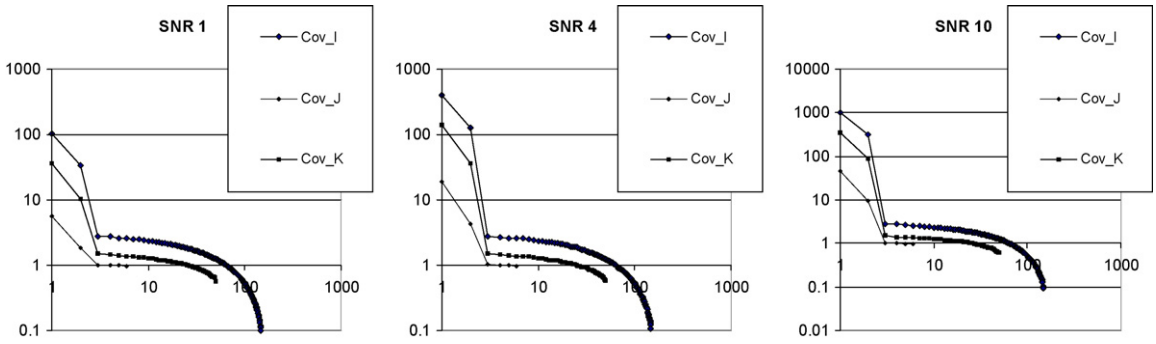

Fig. 2. The singular values of COV^I, COV^J and COV^K (Eq. (33)) are plotted on a double logarithmic scale. Obviously, for all SNRs, the singular values jump at $L=2$, for all SNRs.

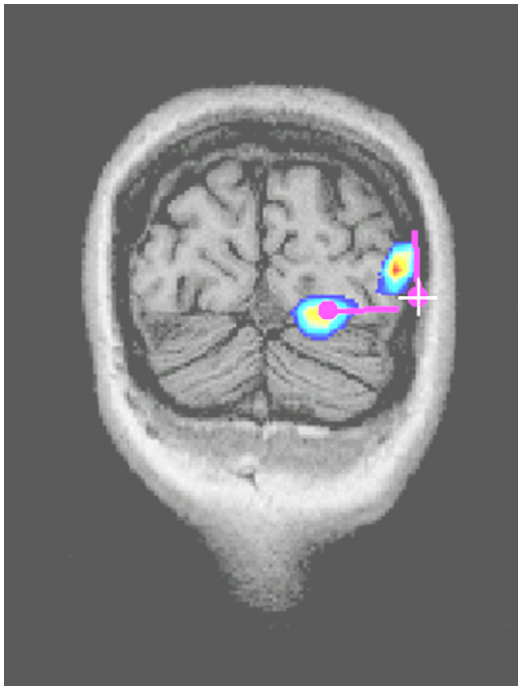

Fig. 3. A comparison of the true dipoles and the *ThreeWayMUSIC* scan. This scan clearly shows two local minima near the true dipoles.

MUSIC scan and the two local minima were much better separated, see Fig. 3.

8. Discussion and conclusions

Dipole fitting algorithms can be stabilized substantially by combining data from different data sets and using one of both coupled dipole models. The extra free parameters in the extended CDM hardly diminish the stability of the dipole parameters (Table 1). It turns out that the estimation of the number of sources and the derivation of the lower bound for the cost function is possible using the singular value plots based on COV^I, COV^J and COV^K of Eq. (33). Two ways of extending the MUSIC algorithm to three-way matrices were presented, *PARAFAC MUSIC* and *ThreeWayMUSIC*. Of these variants, *ThreeWayMUSIC* looks most promising.

One of the reasons why the jump in singular values, as presented in Fig. 2, is so apparent is that simulated noise was Gaussian and uncorrelated over channels, time and conditions. Although this situation is not realistic when dealing with real MEG or EEG data, when a sufficient number of trials is available, an accurate noise covariance matrix can be obtained using ML estimation (de Munck et al., 2002) and after applying pre-whitening (Eq. (10)), data analysis is equivalent to the case of uncorrelated noise. In most VEP/VEF studies presented in the literature, the pre-whitening step is omitted and that might explain why in such studies the jump in singular values is usually not so clear. Finally, we remark that, because of the pre-whitening, not only the jump itself is predicted but also the jump in singular values occurs at unity level. This is true for the two-way case (Eqs. (23) and (24)) as well as for the three-way case (Eq. (33)).

The coupled dipole model and its MUSIC variants can be applied in cases where multiple MEG/EEG experiments were carried out on the same subject. However, one could also think of the same (or related) experiments on multiple subjects. One of the difficulties with the analysis of such data sets is that dipole strength and dipole depth are (negatively) correlated hampering, e.g. the comparison of left and right dipole strengths. Application of the coupled dipole model would reduce the variance of dipole strengths (now appearing in the model as the non-zeroes of the coupling matrices) because dipoles of all subjects are automatically forced to stay at the same depth, dictated by the data and the OLS principle. However, direct application of the coupled dipole model would require the assumption that the volume conductors of all subjects are identical.

Other applications could be the analysis of groups of different inter-ictal spikes, or the dipole analysis of single trial data. In the latter case, the index k would refer to different trials instead of different data sets.

References

Bijma F, de Munck JC, Böcker KBE, Huizenga HM, Heethaar RM. The coupled dipole model: an integrated model for multiple MEG/EEG data sets. NeuroImage 2004;23(3):890–904.

Bijma F, de Munck JC, Huizenga HM, Heethaar RM, Nehorai A. Simultaneous estimation and testing in multiple MEG data sets. IEEE Trans SP 2005;53(9):3449–60.

Bro R. PARAFAC. Tutorial and applications. Chemomet Intelligent Lab Sys 1997;38:149–71.

Cheyne D, Gaetz W, Garnero L, Lachaux JP, Ducorps A, Schwartz D, et al. Neuromagnetic imaging of cortical oscillations accompanying tactile stimulation. Cog Brain Res 2003;17:599–611.

Dale AM, Sereno M. Improved localisation of cortical activity by combining EEG and MEG with MRI surface reconstruction: a linear approach. J Cog Neurosci 1993;5:162–76.

Daunizeau J, Grova C, Mattout J, Marrelec G, Clonda D, Goulard B, et al. Assessing the relevance of fMRI-based prior in the EEG inverse problem: a Bayesian model comparison approach. IEEE Trans SP 2005;53:3461–72.

de Munck JC. The estimation of time varying dipoles on the basis of VEPs. Electroencephalogr Clin Neurophysiol 1990;77:156–60.

de Munck JC, van Dijk BW, Spekreijse H. Mathematical dipoles are adequate to describe realistic generators of human brain activity. IEEE Trans Biomed Eng 1988;35(11):960–6.

de Munck JC, Huizenga HM, Waldorp LJ, Heethaar RM. Estimating stationary dipoles from MEG/EEG data contaminated with spatially and temporally correlated background noise IEEE Trans. Signal Process 2002;50(7):1565–72.

Grasman RPPP, Huizenga HM, Waldorp LJ, Bocker KBE, Molenaar PCM. Frequency domain simultaneous source and source coherence estimation with an application to MEG. IEEE Trans Biomed Eng 2004;51(1):45–55.

Hillebrand A, Singh KD, Holliday IE, Furlong PL, Barnes GR. A new approach to neuroimaging with magnetoencephalography. Human Brain Mapping 2005;25:199–211.

Huizenga HM, de Munck JC, Waldorp LJ, Grasman RPPP. Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model. IEEE Trans Biomed Eng 2002;49(6):533–9.

Kiers HAL, Krijnen WP. An efficient algorithm for PARAFAC of Three-Way data with large numbers of observation units. Psychometrika 1991;57:147–52.

Kroonenberg PM. Three-mode principle component analysis. Leiden: DSWO Press; 1983.

Kruskal JB. Three-way arrays: rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics. Linear Algebra Appl 1977;18:95–138.

Kruskal JB. Rank, uniqueness and decomposition for 3-way and N-way arrays. In: Coppi R, Bolasco S, editors. Multiway data analysis. Amsterdam, The Netherlands: North-Holland; 1989. p. 7–18.

Möcks J. Topographic components model for event related potentials and some biophysical considerations IEEE Trans. Biomed Eng 1988;35(6):482–4.

Mosher JC, Lewis PS, Leahy RM. Multiple dipole modeling and localization from spatio-temporal MEG data. IEEE Trans Biomed Eng 1992;39(6):541–57.

Mosher JC, Leahy RM. Recursively applied MUSIC: a framework for EEG and MEG source localisation. IEEE Trans Biomed Eng 1998;45(11):1342–54.

Mosher JC, Leahy RM. Source localization using recursively applied and projected (RAP) MUSIC. IEEE Trans Signal Process 1999;47(2):332–40.

Mosher JC, Baillet S, Leahy RM. Equivalence of linear approaches in bioelectromagnetic inverse solutions. IEEE workshop on, statistical Signal processing page. 2003; 294–7.

Pascual-Marqui RD, Michel CM, Lehman D. Low resolution electromagnetic tomography: a new method to localize electrical activity in the brain. Int J Psychophysiol 1994;18:49–65.

Robinson SE, Vrba J. Functional neuroimaging by synthetic aperture magnetometry. In: Yoshimoto T, Kotani M, Kuriki S, Karibe H, Nakasato N, editors. Recent advances in biomagnetism. Sendai: Tohoku Univ. Press; 1999. p. 302–5.

Sarvas J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys Med Biol 1987;32:11–22.

Scherg M, Von Cramon D. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. Electroencephalogr Clin Neurophysiol 1985;62:32–44.

Schmidt RO. Multiple emitter location and signal parameter estimation. IEEE Trans Antennas Propagat 1986;AP-34:276–80.

Seber GAF, Wild CJ. Nonlinear regression. New York: Wiley; 1989.

Sekihara K, Sahani M, Nagarajan SS. Localisation bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. NeuroImage 2005;25:1056–67.

Slotnick SD, Klein SA, Carney T, Sutter EE, Dastmalchi S. Using multi-stimulus VEP source localization to obtain a retinotopic map of human primary visual cortex. Clin Neurophysiol 1999;110:1793–800.

Ten Berge JMF. On the uniqueness in CANDECOMP/PARAFAC. Psychonometrika 2002;67:1–11.

Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans Biomed Eng 1997;44(9):867–80.

Waldorp LJ, Huizenga HM, Dolan CV, Molenaar PCM. Estimated generalized least squares electromagnetic source analysis based on a parametric noise covariance model [EEG/MEG]. IEEE Trans Biomed Eng 2001;48(6):737–41.

Waldorp LJ, Huizenga HM, Grasman RPPP, Bocker KBE, de Munck JC, Molenaar PCM. Model selection in electromagnetic source analysis with an application to VEFs. IEEE Trans Biomed Eng 2002;49(10):1121–9.

Waldorp LJ, Huizenga HM, Nehorai A, Grasman RPPP, Molenaar PCM. Model selection in spatio-temporal electromagnetic source analysis. IEEE Trans Biomed Eng 2005;52(3):414–20.

Wipf D, Nagarajan S. Beamforming using the relevance vector machine. In: Proceedings of the 24-th international conference on machine learning; 2007.