

# Simultaneous Estimation and Testing of Sources in Multiple MEG Data Sets

Fetsje Bijma, Jan C. de Munck, Hilde M. Huizenga, Rob M. Heethaar, and Arye Nehorai, *Fellow, IEEE*

**Abstract**—The proposed Extended Couple Dipole Model (ECDM) is a trilinear component model that can be used to analyze multiple, related MEG data sets simultaneously. Related MEG data sets are data sets that contain activity of the same sources or activity of sources that have proportional source amplitudes. The simultaneous model uses a set of common sources and a set of common source time functions (wave shapes) to model the measured data in each data set. The set of common sources contains all sources that are active in at least one of the data sets to be analyzed. The number of common spatial and temporal components is specified by the user. The model for each data set is a linear combination of these common spatial and temporal components. This linear combination is estimated in a coupling matrix. Unlike the Coupled Dipole Model, where the user selects certain entries of the coupling matrix to be zero, the entire coupling matrix is estimated in the ECDM. This yields a more objective and statistically transparent estimation method, of which the identifiability constraints do not depend on the user's chosen design as in the CDM. Cramèr–Rao Bounds are derived for the ECDM, and the significance of the estimated source activity is computed and illustrated by confidence regions around estimated source time functions.

**Index Terms**—Component model, confidence intervals, constrained Cramèr–Rao Bound, coupled dipole model, MEG, spatiotemporal covariance, trilinear model.

## I. INTRODUCTION

THE idea of describing measured magnetoencephalography/electroencephalography (MEG/EEG) data as a linear combination of common (basic) components has been elaborated in many ways. Main examples are Principal Component Analysis (PCA) [19], Parallel Factor Analysis (PARAFAC) [22], a special case of which is the Topographic Component Model (TCM) [1], [23], [30], and other trilinear models, as in [3], [7], and [34]. These models make use of common spatial and temporal components, which all have a certain level of generality (parametrization). In the TCM, each data set is modeled as a weighted sum of parametric spatiotemporal components of which the weights depend on subject and

Manuscript received November 8, 2004; revised November 8, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Guido Nolte.

F. Bijma and J. C. de Munck are with the MEG Centre, Vrije University Medical Centre, Amsterdam, The Netherlands (e-mail: f.bijma@vumc.nl; jc.de-munck@vumc.nl).

H. M. Huizenga is with the Developmental Psychology Department, University of Amsterdam, Amsterdam, The Netherlands (e-mail: h.m.huizenga@uva.nl).

R. M. Heethaar is with the Physics and Medical Technology Department of the Vrije University Medical Centre, Amsterdam, The Netherlands (e-mail: rm.heethaar@vumc.nl).

A. Nehorai is with the Engineering Department, University of Illinois at Chicago, Chicago, IL 60646 USA (e-mail: nehorai@ece.uic.edu).

Digital Object Identifier 10.1109/TSP.2005.853097

condition. The other models cited decouple the spatial and temporal components by placing a coupling (loading) matrix between the common spatial and common temporal component matrices, specifying for each data set the linear combination of common components.

Recently, the Coupled Dipole Model (CDM) [3] was presented. The CDM is a trilinear model in which multiple, related, MEG data sets (e.g., different conditions) are analyzed simultaneously. It is assumed that certain sources and source time functions are equal or proportional across data sets, which relatively increases the signal-to-noise ratio (SNR). For each data set, the coupling matrix specifies the combination of common sources and common source time functions that models the data in that particular data set. Thus, this matrix couples the common spatial to the common temporal components, hence, the name "coupling matrix." This matrix is designed by the user; the entries corresponding to the chosen combination are estimated from the data, whereas the remaining entries are set to zero. Although the CDM was shown to improve the stability of the Inverse Problem (IP) in MEG, this matrix design is subject to the *a priori* information and assumptions of the user because fewer or more assumptions about the similarity of the data sets are reflected in different coupling matrices. On the one hand, it is this exploitation of prior knowledge that is the most important asset of the CDM, although, on the other hand, this also causes subjectivity.

In order to resolve this subjectivity, the currently proposed Extended Coupled Dipole Model (ECDM) estimates the coupling matrices entirely. At the same time, the exploitation of prior information changes: Only the dimension of the coupling matrix is chosen, whereas all entries are estimated from the data. An important advantage of the (E)CDM over most of the other component models cited is the statistical clarity within the Maximum Likelihood (ML) paradigm (cf. [7]). Moreover, whereas the identifiability constraints on the CDM parameters depend on the precise design of the coupling matrices, the identifiability of the ECDM parameters is guaranteed by constraints that only depend on the number of common components. In addition, this enables computing error bounds of estimated source parameters (cf. [7], [9], [14], [24], [25], and [32]) by means of the Cramèr–Rao Bound for constrained parameters [29].

In Section II, the model is explained, and the maximum likelihood estimators for the ECDM parameters are presented. Furthermore, identifiability constraints for the parameters are given, and the constrained Cramèr–Rao Bound is computed for all parameters. Finally, an expression for the confidence region around the estimated source time functions is given. Computational details of this section are moved to the appendices and may be skimmed over by the reader. In the section

on the results, two experimental applications of the ECDM are considered. In the first section, the CDM and the ECDM are compared, and in the second section, the importance of the confidence regions is investigated. In the last section, the results are discussed, and conclusions are drawn.

## II. METHODS

### A. Model

In the ECDM, multiple MEG data sets are modeled in one simultaneous analysis using a set of common stationary dipole sources and a set of common source time functions (STFs). This integrated model is, like the CDM, applicable when MEG data are measured in different though partly similar experimental conditions such that in the various data sets (conditions), the same sources are activated, and/or proportional STFs occur. An example would consist of different data sets in which the somatosensory cortex is activated using electric pulses on the median nerve in a stimulation frequency that varies across data sets. The thus-obtained MEG data sets will show certain similarities, although they will be different.

The following notational convention is used: Matrices are denoted by capital letters that are either Greek or Latin (e.g.,  $\Omega$ ,  $A$ ,  $\mathbf{I}$ ); vectors are written either in boldface Latin or in lowercase Greek letters (e.g.,  $\mathbf{a}$ ,  $\xi$ ); scalars are, in general, denoted by lowercase Latin letters; however, for indices, the lower and upper cases of one letter are used, e.g.,  $p = 1, \dots, P$ ; finally, the  $(n, m)$ th entry of a matrix  $A$  is written as  $A_{n,m}$ .

Suppose there are  $Q$  data sets  $q = 1, \dots, Q$ , each of which contains  $K_q$  trials. Following the Signal-Plus-Noise model for evoked response data, the measured response  $R_q^k$  in trial  $k$  of data set  $q$  is assumed to consist of two components: a fixed response  $R_q$  (that is, fixed within data set  $q$ ) and additional noise  $E_q^k$  [2], [5], [21]

$$R_q^k = R_q + E_q^k. \quad (1)$$

The ECDM models the fixed responses  $R_q$  as linear combinations of the common spatial and temporal components, which are specified by coupling matrices. In the model equation for the fixed response, this is expressed as

$$R_q = AC_qB \quad (2)$$

where  $A$  is the  $(I \times P)$  common dipole field matrix,  $C_q$  the  $(P \times Z)$  coupling matrix for data set  $q$ , and  $B$  the  $(Z \times J)$  common source time function matrix. Here,  $I$  is the number of sensors,  $J$  the number of time samples,  $P$  the number of source time functions belonging to the common sources, and  $Z$  the number of common source time functions. Note that it may be assumed that  $B$  has full row rank and, in particular, that  $Z \leq J$ . To see this, assume that  $\text{rank}(B) = r$  and that  $r \leq J$  and  $r \leq Z$ . Then, there exists a  $Z \times r$  matrix  $C_0$ , and a  $r \times J$  matrix  $B_0$ , such that  $B = C_0B_0$ , and  $\text{rank}(C_0) = \text{rank}(B_0) = r$ . The model can then be written as  $AC_qB = AC_q(C_0B_0) = A(C_qC_0)B_0$ , where  $C_qC_0$  is now a  $P \times r$  matrix, functioning as a coupling matrix, and  $B_0$  is the new common STF matrix with full row rank. The matrix  $A$  in (2) is parameterized by the

nonlinear source location parameters  $\xi$  and by the linear source orientation parameters  $\eta$ :

$$A = A(\xi, \eta). \quad (3)$$

If all common sources are single dipoles with fixed orientations, then  $P$  equals the number of common sources. This is because the forward field of a single dipole source is represented by one column in  $A$  and is parameterized by one location and one orientation. In the case of symmetric sources, that is, sources consisting of two dipoles having symmetric locations and orientations with respect to the midsagittal plane,  $P$  equals twice the number of common sources. The forward field of a symmetric source is a set of two columns in  $A$  (one for the left and one for the right dipole) because each dipole has its own amplitude function and is parameterized by one location (the other is the mirrored location) and one orientation.

The orientations of the common sources in  $\eta$  and the common times series in  $B$  are normalized, whereas the magnitudes of activation are estimated in the coupling matrices  $C_q$ . For a full overview of all symbols used globally in this paper, see the Appendix.

In the CDM, the coupling matrices are designed by the user: The user specifies the dimension of the coupling matrices (i.e., the number of common components) and selects the zero entries. In the ECDM, the subjectivity of this zero selection is circumvented by estimating the entire coupling matrices, whereas the dimension is still set by the user. The first, immediate, consequence is the nonuniqueness

$$R_q = AC_qB = A(C_qG^{-1})(GB) \quad (4)$$

for any nonsingular matrix  $G$ . Therefore, constraints on the parameters will be necessary to obtain a unique solution. Moreover, whereas in the CDM the common temporal components bear a physiological meaning (e.g., contralateral STF), in the ECDM, the direct physiological meaning of the matrices  $C_q$  and  $B$  has been changed into a more abstract one, depending on the choice of  $G$ , although the product  $C_qB$  is still meaningful. This illustrates the tradeoff between objectivity and interpretability in the ECDM and the CDM. Finally, the ECDM is more general than the CDM because the ECDM solutions include all CDM solutions, although the converse is, in general, not true.

### B. ML Estimators for the Parameters

The model parameters to be estimated are

- the location parameters of the common sources, contained in the vector  $\xi$ ;
- the orientation parameters of the common sources, contained in the vector  $\eta$ ;
- the entire coupling matrices  $C_q$ , for  $q = 1, \dots, Q$ ;
- the common STF matrix  $B$ .

In order to apply the ML method to estimate these parameters, the likelihood of the data has to be derived. The noise is assumed to have a Gaussian distribution with a known spatiotemporal covariance following the Kronecker product model [4], [13]. This can be denoted by

$$\text{vec}(E_q^k) \sim \mathcal{N}(0, T \otimes X) \quad (5)$$

where  $T$  and  $X$  are the temporal and spatial covariance matrices, respectively, which are assumed to be known or estimated, as described in [3]. Furthermore, different trials are assumed to be independent, and different data sets even more so (which are contained in different trials in practice); thus,  $X$  and  $T$  are independent of  $k$  and  $q$ . These matrices are decomposed as

$$T^{-1} = W_T W_T^t \quad (6)$$

$$X^{-1} = W_X W_X^t. \quad (7)$$

The following definitions are convenient:

$$\tilde{R}_q^k := W_X^t R_q^k W_T \quad (8)$$

$$\tilde{R}_q := \frac{1}{K_q} \sum_{k=1}^{K_q} \tilde{R}_q^k \quad (9)$$

$$\tilde{E}_q^k := W_X^t E_q^k W_T \quad (10)$$

$$\tilde{E}_q := \frac{1}{K_q} \sum_{k=1}^{K_q} \tilde{E}_q^k. \quad (11)$$

Using these prewhitened matrices, we define the arrays

$$\begin{aligned} \mathbf{R} &= \begin{pmatrix} \sqrt{K_1} \tilde{R}_1 \\ \vdots \\ \sqrt{K_Q} \tilde{R}_Q \end{pmatrix}, \quad \mathbf{A} = (\mathbf{I}_Q \otimes W_X^t A) \\ \mathbf{C} &= \begin{pmatrix} \sqrt{K_1} C_1 \\ \vdots \\ \sqrt{K_Q} C_Q \end{pmatrix}, \quad \mathbf{B} = B W_T \\ \text{and } \mathbf{E} &= \begin{pmatrix} \sqrt{K_1} \tilde{E}_1 \\ \vdots \\ \sqrt{K_Q} \tilde{E}_Q \end{pmatrix} \end{aligned} \quad (12)$$

so that the joint model of (1) for all  $q$  becomes

$$\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{B} + \mathbf{E}. \quad (13)$$

Matrix  $\mathbf{A}$  is still parameterized by  $(\xi, \eta)$ . Furthermore, since  $W_X$  and  $W_T$  are nonsingular matrices, an equivalent set of model parameters is given by  $(\xi, \eta, \mathbf{C}, \mathbf{B})$ . From (5), the distribution of  $\sqrt{K_q} \text{vec}(\tilde{R}_q)$  is derived:

$$\sqrt{K_q} \text{vec}(\tilde{R}_q) \sim \mathcal{N} \left( \text{vec} \left( W_X^t A \sqrt{K_q} C_q B W_T \right), \mathbf{I}_{IJ} \right) \quad (14)$$

yielding the marginal likelihood for  $\mathbf{R}$

$$L(\xi, \eta, \mathbf{C}, \mathbf{B}) = \frac{1}{(2\pi)^{(IJJ/2)}} e^{-\frac{1}{2} \text{tr}((\mathbf{R} - \mathbf{A} \mathbf{C} \mathbf{B})(\mathbf{R} - \mathbf{A} \mathbf{C} \mathbf{B})^t)} \quad (15)$$

and the log likelihood

$$\begin{aligned} l(\xi, \eta, \mathbf{C}, \mathbf{B}) &= -\frac{IJJ}{2} \log(2\pi) \\ &\quad -\frac{1}{2} \text{tr}((\mathbf{R} - \mathbf{A} \mathbf{C} \mathbf{B})(\mathbf{R} - \mathbf{A} \mathbf{C} \mathbf{B})^t). \end{aligned} \quad (16)$$

The ML estimators for  $(\xi, \eta, \mathbf{C}, \mathbf{B})$  are defined as

$$(\hat{\xi}, \hat{\eta}, \hat{\mathbf{C}}, \hat{\mathbf{B}}) = \arg \max l(\xi, \eta, \mathbf{C}, \mathbf{B}). \quad (17)$$

The estimators are found by equating the corresponding first-order derivatives of (16) to zero and solve for the parameters. Owing to the prewhitening, this is equivalent to the Least Squares method.

The expressions for the estimators are derived in three steps. In the first step,  $\xi$ ,  $\eta$ , and  $\mathbf{B}$  are assumed to be known, and  $\mathbf{C}$  is estimated as

$$\hat{\mathbf{C}} = (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R} \mathbf{B}^t (\mathbf{B} \mathbf{B}^t)^{-1}. \quad (18)$$

$(\mathbf{A}^t \mathbf{A})^{-1} = \mathbf{I}_Q \otimes (A^t X^{-1} A)^{-1}$  requires the inverse of a  $P \times P$  matrix, and  $(\mathbf{B} \mathbf{B}^t)^{-1} = (B T^{-1} B^t)^{-1}$  requires a  $Z \times Z$  inverse. Substituting (18) into (16) yields the concentrated log likelihood function  $l(\xi, \eta, \mathbf{B})$

$$\begin{aligned} l(\xi, \eta, \mathbf{B}) &= -\frac{IJJ}{2} \log 2\pi - \frac{1}{2} \text{tr}(\mathbf{R}^t \mathbf{R}) \\ &\quad + \frac{1}{2} \text{tr}(\mathbf{R}^t \mathbf{A} (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R} \mathbf{B}^t (\mathbf{B} \mathbf{B}^t)^{-1} \mathbf{B}). \end{aligned} \quad (19)$$

In the next step,  $\mathbf{B}$  is estimated, under the assumption that  $(\xi, \eta)$  is known. First, write

$$\mathbf{R}^t \mathbf{A} (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R} = U \Delta U^t \quad (20)$$

with  $U U^t = U^t U = \mathbf{I}_J$ , and  $\Delta = \text{diag}(\mu_1, \mu_2, \dots, \mu_J)$ , with  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_J \geq 0$ . Because  $U$  is orthogonal, there exists a  $(Z \times J)$ -matrix  $B_0 = \mathbf{B} U$  such that

$$\mathbf{B} = B_0 U^t. \quad (21)$$

Now, maximizing the concentrated log likelihood in (19) is equivalent to

$$\max_{B_0} \text{tr} \left( \Delta B_0^t (B_0 B_0^t)^{-1} B_0 \right). \quad (22)$$

In Appendix B, the optimal  $B_0$  is determined:

$$\hat{B}_0 = G (\mathbf{I}_Z \quad \mathbf{0}) \quad (23)$$

where  $G$  is any nonsingular  $(Z \times Z)$  matrix. For  $\mathbf{B}$ , this implies

$$\hat{\mathbf{B}} = G U_Z^t \quad (24)$$

where  $U_Z$  denotes the matrix containing only the first  $Z$  columns of the matrix  $U$  defined in (20). The matrix  $G$  exhibits the nonuniqueness of the model, as mentioned above. The imposed constraint is

$$G = \mathbf{I}_Z \text{ which is equivalent to } \mathbf{B} U_Z = \mathbf{I}_Z. \quad (25)$$

Note that the rows in  $\hat{\mathbf{B}}$  (the prewhitened STFs) are normalized by this constraint. Again, the log likelihood function can be concentrated, yielding  $l(\xi, \eta)$ :

$$l(\xi, \eta) = \frac{1}{2} \sum_{j=1}^Z \lambda_j (\mathbf{R}^t \mathbf{A} (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R}) \quad (26)$$

where  $\lambda_j(A)$  denotes the  $j$ th biggest eigenvalue of matrix  $A$ .

In the last step, the source parameters  $(\xi, \eta)$  are estimated. In order to find these parameters, the expression in (26) has to be maximized using a nonlinear iterative algorithm (e.g., Marquardt).

Finally, the procedure of estimating  $\xi$ ,  $\eta$ ,  $\mathbf{C}$ , and  $\mathbf{B}$  is run in reverse order:

- 1) Estimate the location and orientation parameters  $\hat{\xi}$  and  $\hat{\eta}$  in a nonlinear iterative method maximizing (26).
- 2) Derive  $\hat{\mathbf{B}}$  from (24).
- 3) Calculate  $\hat{\mathbf{C}}$  from (18).

As a remark, see an alternative to the first step, which was presented in [6]. In that study, the linear dependency of  $\mathbf{A}$  on the orientation parameters  $\eta$  is exploited, and only the locations in  $\xi$  are estimated in the (time-consuming) nonlinear search. In this faster approach, the first step is replaced by the following scheme, which is iterated until convergence of the cost function:

- 1) Obtain a starting value or update value for  $\xi$  using the nonlinear algorithm.
- 2) Set the orientations in  $\eta$  to default values, e.g., (1,0,0).
- 3) Iterate until convergence of the maximization of (26):
  - a) Estimate the optimal matrix  $\mathbf{B}$  and  $\mathbf{C}$  using (18) and (24).
  - b) Estimate the optimal orientations by

$$\hat{\eta} = \left( \sum_{j=1}^J (A(j))^t A(j) \right)^{-1} \sum_{j=1}^J (A(j))^t \mathbf{r}_j \quad (27)$$

(see Appendix C for the meaning of  $A(j)$  and  $\mathbf{r}_j$ ).

For this fast approach, the starting values for the location parameters are obtained by a global search.

### C. Parameter Testing

Due to noise in the data array  $\mathbf{R}$ , all estimated entries of  $\hat{\mathbf{C}}$  in (18) will be nonzero with probability 1. Consequently, all common sources will have nonzero estimated activity in all data sets. To test the significance of the estimated activity of a certain source  $p$ , in a certain data set  $q$ , its estimated amplitude over time is tested against the null hypothesis of being zero. This amplitude is the vector-matrix product of the  $p$ th row in  $C_q$  and the matrix  $\mathbf{B}$ ,  $H_0 : ((\hat{\mathbf{C}}\mathbf{B})_{p+(q-1)P;1}, \dots, (\hat{\mathbf{C}}\mathbf{B})_{p+(q-1)P;J}) = (0, \dots, 0) \in \mathbb{R}^{1 \times J}$ . Because the common moments and common STFs are normalized, the amplitude of the source activity over the entire time window is assessed by considering the  $(p + (q - 1)P)$ th row in  $\hat{\mathbf{C}}$ . In other words, the null hypothesis can be written as a test on a subset of the ECDM parameters  $H_0 : (\hat{\mathbf{C}}_{p+(q-1)P;1}, \dots, \hat{\mathbf{C}}_{p+(q-1)P;Z}) = (0, \dots, 0) \in \mathbb{R}^{1 \times Z}$ . Below, the statistical distribution of these amplitudes is derived first, and then, the tests on the null hypotheses are described.

Using the ML paradigm, the estimators are asymptotically efficient [18], [27], i.e., the covariance of the estimated parameters approaches the Cramèr–Rao Lower Bound for infinitely many trials (observations). The calculation of this bound utilizes the Fisher Information Matrix [16]

$$\mathcal{I}(\theta) = \mathcal{E} \left[ \left( \frac{\partial l(\theta)}{\partial \theta} \right)^t \frac{\partial l(\theta)}{\partial \theta} \right] = -\mathcal{E} \left[ \frac{\partial^2 l(\theta)}{\partial \theta^t \partial \theta} \right] \quad (28)$$

where  $\theta$  is the vector containing all model parameters, and  $l(\theta)$  is the log likelihood function. The Cramèr–Rao inequality for the covariance of the estimated parameters  $\hat{\theta}$  is given by [26]

$$\text{Cov}(\hat{\theta}) \geq (\mathcal{I}(\theta^*))^{-1} \quad (29)$$

where  $A \geq B$  means that  $A - B$  is positive semi-definite, and  $\theta^*$  denotes the true value of the parameter vector. Furthermore, the covariance matrix of an ML-estimated  $\hat{\theta}_{\text{ML}}$ , based on finitely many observations, is estimated by [18]

$$\widehat{\text{Cov}}(\hat{\theta}_{\text{ML}}) = (\mathcal{I}(\hat{\theta}_{\text{ML}}))^{-1}. \quad (30)$$

For the ECDM, the parameter vector  $\theta$  is defined as

$$\theta = (\text{vec}(\mathbf{C}), \zeta, \text{vec}(\mathbf{B})) \quad (31)$$

with  $\zeta = (\xi, \eta)$ . The symmetric information matrix can then be partitioned as

$$\mathcal{I}(\theta) = \begin{pmatrix} \mathcal{I}_{CC} & \mathcal{I}_{C\zeta} & \mathcal{I}_{CB} \\ \mathcal{I}_{\zeta C} & \mathcal{I}_{\zeta\zeta} & \mathcal{I}_{\zeta B} \\ \mathcal{I}_{BC} & \mathcal{I}_{B\zeta} & \mathcal{I}_{BB} \end{pmatrix}. \quad (32)$$

From (16) and (28), an arbitrary entry of  $\mathcal{I}(\theta)$  is given by

$$\begin{aligned} \mathcal{I}(\theta)_{l,l'} &= \text{tr} \left( \frac{\partial(\mathbf{ACB})^t}{\partial \theta_l} \frac{\partial(\mathbf{ACB})}{\partial \theta_{l'}} \right) \\ &= \frac{\partial \text{vec}((\mathbf{ACB})^t)}{\partial \theta_l} \frac{\partial \text{vec}(\mathbf{ACB})}{\partial \theta_{l'}}. \end{aligned} \quad (33)$$

Therefore, the following derivatives are sufficient to compute all the submatrices in (32):

$$\frac{\partial \text{vec}(\mathbf{ACB})}{\partial \text{vec}(\mathbf{C})} = \mathbf{B}^t \otimes \mathbf{A} \quad (34)$$

$$\frac{\partial \text{vec}(\mathbf{ACB})}{\partial \zeta} = (\mathbf{B}^t \mathbf{C}^t \otimes \mathbf{I}_{IQ}) \frac{\partial \text{vec}(\mathbf{A})}{\partial \zeta} \quad (35)$$

$$\frac{\partial \text{vec}(\mathbf{ACB})}{\partial \text{vec}(\mathbf{B})} = \mathbf{I}_J \otimes \mathbf{AC}. \quad (36)$$

The information submatrices are

$$\begin{aligned} \mathcal{I}_{CC} &= \mathbf{B}\mathbf{B}^t \otimes \mathbf{A}^t \mathbf{A} \\ \mathcal{I}_{C\zeta} &= (\mathbf{B}\mathbf{B}^t \mathbf{C}^t \otimes \mathbf{A}^t) \frac{\partial \text{vec}(\mathbf{A})}{\partial \zeta} \\ \mathcal{I}_{CB} &= \mathbf{B} \otimes \mathbf{A}^t \mathbf{AC} \\ \mathcal{I}_{\zeta\zeta} &= \left( \frac{\partial \text{vec}(\mathbf{A})}{\partial \zeta} \right)^t (\mathbf{C}\mathbf{B}\mathbf{B}^t \mathbf{C}^t \otimes \mathbf{I}_{IQ}) \frac{\partial \text{vec}(\mathbf{A})}{\partial \zeta} \\ \mathcal{I}_{\zeta B} &= \left( \frac{\partial \text{vec}(\mathbf{A})}{\partial \zeta} \right)^t (\mathbf{C}\mathbf{B} \otimes \mathbf{AC}) \\ \mathcal{I}_{BB} &= \mathbf{I}_J \otimes \mathbf{C}^t \mathbf{A}^t \mathbf{AC}. \end{aligned} \quad (37)$$

To compute the estimated covariance matrix (30), the inverse of the information matrix is required. However, the information matrix given by the submatrices in (37) is singular, which can be easily seen using the determinant formula for the partitioned matrix [31]: For a nonsingular symmetric  $m \times m$  matrix  $\mathbf{M}$ , a symmetric  $n \times n$  matrix  $\mathbf{N}$  and any  $m \times n$  matrix  $\mathbf{K}$ , we have

$$\begin{vmatrix} \mathbf{M} & \mathbf{K} \\ \mathbf{K}^t & \mathbf{N} \end{vmatrix} = |\mathbf{M}| |\mathbf{N} - \mathbf{K}^t \mathbf{M}^{-1} \mathbf{K}|. \quad (38)$$

Setting  $\mathbf{M} = \mathcal{I}_{CC}$  and matrices  $\mathbf{K}$  and  $\mathbf{N}$  correspondingly, the term  $\mathbf{N} - \mathbf{K}^t \mathbf{M}^{-1} \mathbf{K}$  becomes a block diagonal matrix having the following two matrices as blocks on the diagonal:

$$\left( \frac{\partial \text{vec}(\mathbf{A})}{\partial \zeta} \right)^t (\mathbf{C} \mathbf{B} \mathbf{B}^t \mathbf{C}^t \otimes (\mathbf{I}_{PQ} - \mathbf{A}(\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t)) \\ \times \frac{\partial \text{vec}(\mathbf{A})}{\partial \zeta}$$

and

$$(\mathbf{I}_J - \mathbf{B}^t (\mathbf{B} \mathbf{B}^t)^{-1} \mathbf{B}) \otimes \mathbf{C}^t \mathbf{A}^t \mathbf{A} \mathbf{C}$$

of which the second is singular, because  $\mathbf{I}_J - \mathbf{B}^t (\mathbf{B} \mathbf{B}^t)^{-1} \mathbf{B}$  is a projection.

This singularity implies that the CRB cannot be calculated straightforwardly, which, indeed, characterizes the over-parametrization (4). Sufficient identifiability constraints on the parameters for MEG data are as follows [11], [12]:

- 1) normalized source orientations ( $P$  constraints);
- 2) tangential source orientations ( $P$  constraints);
- 3)  $\mathbf{B} \mathbf{U}_Z = \mathbf{I}_Z$ , see (25) ( $Z^2$  constraints).

See Appendix D for the explicit formulas. In [29], the CRB under parametric constraints is derived. All constraints are written in one  $((2P + Z^2) \times 1)$  vector  $\mathbf{c}(\theta) = 0$ , of which the derivative is defined as

$$D(\theta) = \frac{\partial \mathbf{c}(\theta)}{\partial \theta}. \quad (39)$$

$D(\theta)$  (see Appendix D) is a  $((2P + Z^2) \times (PQZ + M + ZJ))$  matrix, where  $M$  is the length of the source parameter vector  $\zeta$ , i.e.,  $M = 3N + 3P$ .  $D(\theta)$  is assumed to be of full row rank, that is, the linear approximations of the constraints are linearly independent. Then, there exists a  $((PQZ + M + ZJ) \times (PQZ + M + ZJ - 2P - Z^2))$  matrix  $V$ , having orthogonal columns, so that  $DV = \mathbf{0}$ , i.e., the columns of  $V$  span the nullspace of  $D$ . The Constrained Cramèr-Rao Bound (CCRB) is now given by [29]

$$\Sigma_\theta = V(V^t \mathcal{I}(\theta) V)^{-1} V^t. \quad (40)$$

This bound is based on the nonsingularity of  $V^t \mathcal{I}(\theta) V$ ; the columns in  $V$  span the same space as the column (or row) space of  $\mathcal{I}(\theta)$ . In other words, the linear approximations to the constraints span the space perpendicular to the row space of the information matrix.

The first  $PQZ$  columns in  $D(\theta)$ , which is the derivative of  $\mathbf{c}(\theta)$  with respect to  $\text{vec}(\mathbf{C})$ , are zero because there are no constraints on  $\mathbf{C}$ . Therefore, we can choose  $V$  to be

$$V = \begin{pmatrix} I_{PQZ} & \mathbf{0} \\ \mathbf{0} & V_0 \end{pmatrix} \quad (41)$$

where the  $\mathbf{0}$ 's present null matrices of appropriate dimensions, and  $V_0$  is an  $((M + ZJ) \times (M + ZJ - 2P - Z^2))$  matrix with orthonormal columns, such that

$$D \begin{pmatrix} \mathbf{0} \\ V_0 \end{pmatrix} = \mathbf{0}. \quad (42)$$

$\Sigma_\theta$  in (40) then becomes

$$\Sigma_\theta = \begin{pmatrix} I_{PQZ} & \mathbf{0} \\ \mathbf{0} & V_0 \end{pmatrix} \\ \times \begin{bmatrix} \mathcal{I}_{CC} & (\mathcal{I}_{C\zeta} \quad \mathcal{I}_{CB}) V_0 \\ V_0^t \begin{pmatrix} \mathcal{I}_{\zeta C} \\ \mathcal{I}_{BC} \end{pmatrix} & V_0^t \begin{pmatrix} \mathcal{I}_{\zeta\zeta} & \mathcal{I}_{\zeta B} \\ \mathcal{I}_{B\zeta} & \mathcal{I}_{BB} \end{pmatrix} V_0 \end{bmatrix}^{-1} \\ \times \begin{pmatrix} I_{PQZ} & \mathbf{0} \\ \mathbf{0} & V_0^t \end{pmatrix}. \quad (43)$$

The derived CCRB is now used to test the significance of the estimated activity, as discussed above. For a certain source  $p$  in a particular data set  $q$ , the null hypothesis is  $H_0 : (\hat{\mathbf{C}}_{p+(q-1)P;1}, \dots, \hat{\mathbf{C}}_{p+(q-1)P;Z}) = (0, \dots, 0) \in \mathbb{R}^{1 \times Z}$ . To test such a subset of the estimated parameters, a Fisher F-test is performed [27], which is based on a linear approximation of the nonlinear model (see Appendix E). Taking into account the constraints on the ECDM parameters, the  $\mathcal{F}$ -statistic  $F_{\text{coupling}}$  in (69) is used to calculate the significance of the activity of the  $p$ th source in the  $q$ th data set over the entire time window, which is indicated by a  $p$ -value.

Although the activity of a certain source may be significant over the entire time window in a certain data set, its activity need not be significant at each time sample. Hence, another interesting statistical feature is the confidence interval around the estimated STF for each source in each data set. These STFs are contained in the matrix product  $\mathbf{C} \mathbf{B}$ , which is a nonlinear function of the ECDM parameters. In [27], testing a nonlinear null hypothesis  $H_0 : ((\hat{\mathbf{C}} \mathbf{B})_{p+(q-1)P;j} = 0 \in \mathbb{R}^{1 \times 1})$  is performed through linear approximation. This same strategy is applied in the present study to compute the confidence regions around the STFs (see Appendix F). The resulting  $\mathcal{F}$ -statistic  $F_{\text{hypstest}}$  in (70) determines the confidence region (74).

### III. RESULTS

The ECDM was applied in two different experimental MEG designs in order to examine its performance. In the first application, MEG data obtained in a Visual Evoked Field (VEF) experiment were studied. These VEF data had also been analyzed by the CDM [3], thus allowing for a comparison between the performances of the CDM and the ECDM. Second, the ECDM was applied to Somatosensory Evoked Field (SEF) data, which enabled investigation of the accuracy and importance of the significance tests and confidence regions.

#### A. VEF Data

VEF data were obtained in three different experimental conditions; a 6-ft checkerboard pattern was shown to the subject in either the left or the right visual hemi-field, or full field. Using a 151-channel CTF Omega system, data of five subjects were acquired at a sample rate of 625 Hz. Each data set contained 400 trials, and the time window of analysis was set to 80–112 ms post stimulus.

As outlined in [3], describing the activity evoked by this check size in this time window requires two common sources and two common STFs. Although, in the CDM, the two STFs

TABLE I

ESTIMATED SOURCE PARAMETERS OF THE SEMI-SYMMETRIC SOURCES FOR SUBJECT 5. THE POSITIONS OF THE SOURCES ARE RELATIVE TO THE CENTER OF THE SPHERICAL VOLUME CONDUCTOR. THE DIRECTION OF THE  $x$ -AXIS IS FORWARD, THE  $y$ -AXIS IS TO THE LEFT, AND THE  $z$ -AXIS POINTS UPWARDS

	position left source (cm)			orientation left source			orientation right source		
	x	y	z	x	y	z	x	y	z
ECDM	-6.41	1.35	5.47	0.30	0.72	0.62	0.26	-0.96	0.10
CDM	-6.40	1.42	5.48	0.29	0.76	0.58	0.28	-0.95	0.12

are given physiological meanings (contra- and ipsilateral STF) in the ECDM, the common STFs are orthogonal and, therefore, lack such a direct physiological meaning. As a model for the two sources, a semisymmetric dipole pair was employed, that is, two dipole sources having symmetric locations, although they have arbitrary orientations. This generated three ( $2 \times 2$ ) coupling matrices for each subject. In the CDM, both hemi-field data sets were modeled by an ipsi- and contralateral source (with ipsi- and contralateral STFs, respectively), and the full field data set was modeled by both sources having the contralateral STF.

In general, the estimated ECDM parameters are very close to estimated CDM parameters. Over the five subjects, the average distance between the ECDM and the CDM estimated locations is 0.11 cm ( $\pm 0.09$  cm), and the average rotation between estimated left orientations equals  $1.6^\circ$  ( $\pm 1.9^\circ$ ) and right  $1.3^\circ$  ( $\pm 1.5^\circ$ ). The ECDM solutions, like the CDM solutions, show nearly symmetric orientations. A slight difference was found in the shapes of the estimated STFs, i.e., the rows of **CB**; however, the magnitudes are comparable for subjects 1, 2, 4, and 5. For subject 3, the magnitude in the ECDM solution is approximately 1.9 times the magnitude in the CDM solution, whereas the shapes are similar. The explanation for this distinction lies in the difference between the estimated sources: The ECDM sources are located closer to the medial plane than the CDM sources (ECDM: 0.4 cm; CDM: 0.8 cm), and the orientations (both ECDM and CDM) are mainly lateral (the angle between orientation and pure lateral direction is  $17^\circ$  (left) and  $9^\circ$  (right) for the ECDM). In other words, the ECDM, which is less restrictive than the CDM, yields slightly canceling sources, as was discussed in [3]; two sources that are close together and have opposite orientations are vulnerable to counterbalancing each other. The data of this subject were in any case the most troublesome of all because separate analyses of all three data sets could not yield an interpretable solution (see [3]). Nevertheless, in general, the ECDM and CDM solutions resemble each other reasonably for the uncomplicated data sets.

The results of subject 5 are representative for the type of shape (all subjects) and magnitude (subject 1, 2, 4, and 5) dissimilarities between the ECDM and the CDM. In Table I, the estimated source parameters are presented, and Fig. 1 shows the estimated STFs, together with the 95% confidence intervals. From this figure and table, it is seen that the ECDM magnitude is slightly bigger than the CDM magnitude at most instants, and correspondingly, the CDM sources are slightly more superficial than the ECDM sources, although not significantly.

### B. SEF Data

In circumstances where a certain source is not expected to be active in one (or some) of the data sets, the estimated coupling

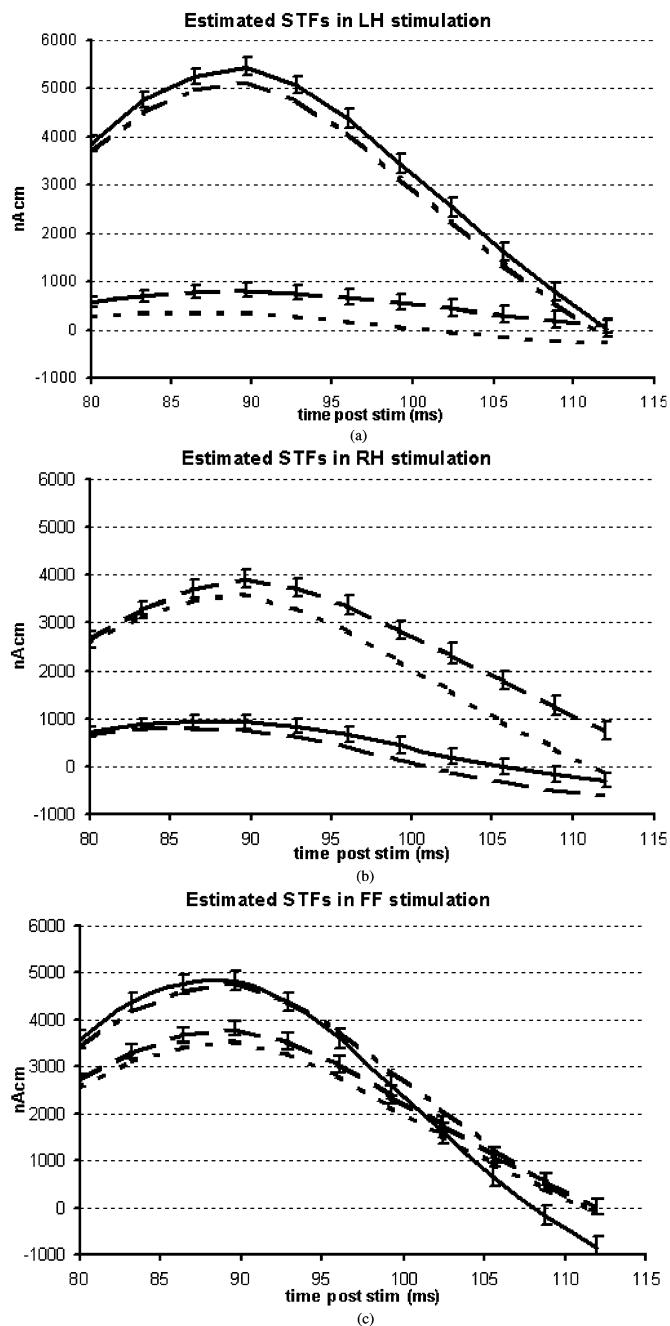


Fig. 1. Estimated STFs of the semi-symmetric sources resulting from both the ECDM and the CDM analyzes of subject 5. The solid (dashed) line presents the ECDM estimated STF of the right (left) source, and the dotted (dotted) line presents the CDM estimated STF of the right (left) source. The vertical error bars indicate a 95% confidence region around the ECDM amplitudes. The horizontal axis represents the post-stimulus time in milliseconds, and the vertical axis shows the amplitude in nano ampere centimeters (nAcm). (a) Left hemi-field stimulation. Right hemi-field stimulation. Full field stimulation.

parameters in the ECDM will, nonetheless, be nonzero with probability 1. This is the case, for example, when the number of common sources is too large for one of the data sets; the entire coupling matrix is estimated, producing nonzero estimated activity for all common sources. To investigate the performance of the ECDM in cases of these expectedly *silent* sources, the model was applied to the N20 response in SEF data of the left and right median nerve. Typically, one would analyze these two data sets

separately, using single source models. However, it is a plausible assumption that the somatosensory cortices on the left and right have symmetric locations. Therefore, the ECDM can be applied, assuming semi-symmetric sources. Nevertheless, the authors want to stress that the purpose of this SEF study is not to find the optimal SEF source localization method but rather to test the robustness and correctness of the proposed model.

The design of the ECDM for the two SEF data sets  $SEF_L$  and  $SEF_R$  used a semi-symmetric source pair. The two sources permit both contra- and ipsilateral activity. The number of STFs was varied from 1 to 4; one STF forces the ipsi- and contralateral amplitudes to be proportional, and two STFs allow for different wave shapes. By extending the number of STFs even further to at most 4 (then  $C$  is square), the difference between the ECDM design and the separate analyses is reduced to only the assumption of symmetric source locations. In the present study, the number of common STFs  $Z$  was varied from 1 to 4; therefore, the dimensions of the coupling matrices varied from  $2 \times 1$  to  $2 \times 4$  for each data set.

Using a 151-channel CTF Omega system, data of five subjects were collected at a sample rate of 2083 Hz. For each subject, the frequency of the median nerve stimulation was set to 2 Hz for both data sets, each of which contained at least 500 trials. For each subject, a narrow interval of 2 ms around the N20 peak was selected as time window of analysis, functioning both for the  $SEF_L$  and the  $SEF_R$  data set.

For all subjects, the estimated contralateral activity clearly dominates the ipsilateral activity. Figs. 2 and 3 show the estimated STFs together with the corresponding confidence intervals for subjects 3 and 5 for  $Z = 2$ . The  $p$  values for both data sets, which are computed according to the above outlined tests, are reported in Table II. As can be seen from the table, the  $p$  values generally increase with  $Z$ , whereas the estimated source parameters (not shown) hardly change with  $Z$ . Moreover, not all  $p$ -values exceed the common 5% limit; in other words, some ipsilateral activity is estimated as significant. Examples of this significant activity are shown in Fig. 3(a) and (b) ( $Z = 2$ ). From these figures, the importance of the confidence regions becomes clear: For all samples, the activity balances on the verge of significance, apparently yielding a  $p$  value just below 5%. Finally, the  $p$  values of the contralateral activity were calculated, yielding values ranging from  $10^{-59}$  to  $10^{-234}$ .

#### IV. DISCUSSION

The ECDM is an extension of the CDM, analyzing multiple MEG data sets simultaneously, in which the subjectivity of the coupling design has been erased. The ECDM can be regarded as an extension of the CDM in the temporal domain: The wave shapes in  $CB$ , i.e., the estimated source time functions for the sources in all data sets, are more general in the ECDM. Spatially, there is no noteworthy difference between the models.

The crucial enhancement by this temporal extension is the statistical transparency; identifiability constraints are independent of the user's design, and the covariance of the parameters is accessible by using the Constrained Cramér–Rao Bound. Another advantage of the ECDM over the CDM is the ease of design: Only the dimension of the coupling matrices is chosen by

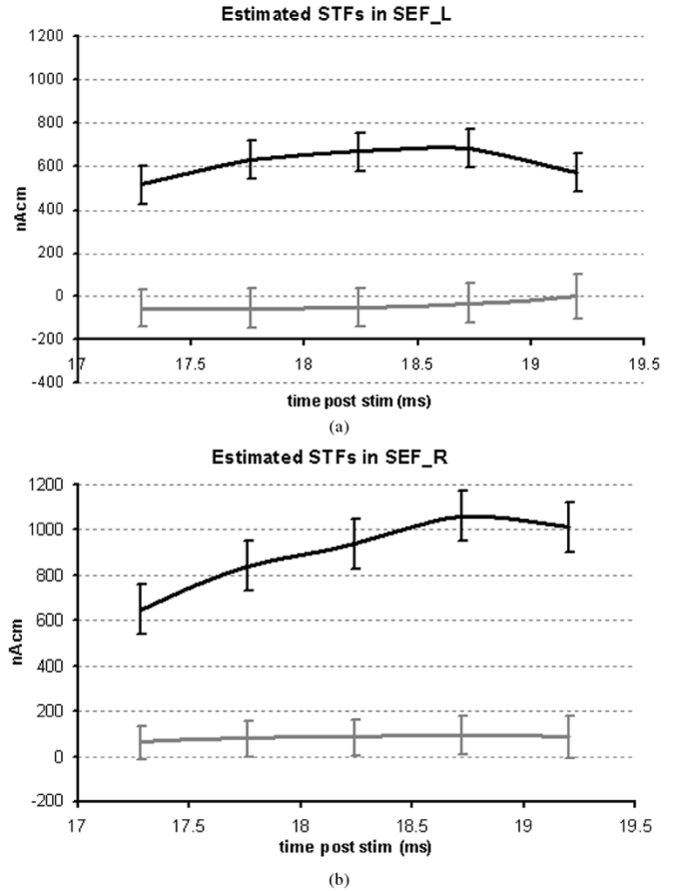


Fig. 2. ECDM estimated STFs of the semi-symmetric sources in the SEF data sets of subject 3. The black (grey) line indicates the estimated amplitude of the left (right) source. The vertical error bars indicate a 95% confidence region around the amplitudes. The horizontal axis represents the post stimulus time in milliseconds, and the vertical axis shows the amplitude in nAcm. (a) SEF left median nerve. (b) SEF right median nerve.

the user. The price to be paid for this benefit is the loss of direct physiological meaning of  $C$  and  $B$ ; however, the product  $CB$  remains meaningful. Furthermore, the prior information cannot be exploited, which can be either a favor or a drawback, depending on the user's desire.

An alternative estimation method for the ECDM is the Generalized Multivariate Analysis of Variance (GMANOVA) model [8], [17], [28], [31]. The GMANOVA (or growth curve) model, which is the trilinear extension ( $Y = AXB$ ) of the common MANOVA model ( $Y = AX$ ), was used to describe a single MEG/EEG data set in [7]. In that study, the spatial noise covariance was estimated simultaneously with the model parameters, having the advantage that the spatial covariance can capture modeling errors. In the more complicated ECDM, this benefit can not be achieved for the general spatiotemporal noise covariance because the GMANOVA model admits this favor only for either the spatial or the temporal noise covariance. Hence, capturing modeling errors in the ECDM should be performed by iterative estimation of the noise covariance and the model parameters.

By applying the ECDM to the same VEF data to which the CDM had been applied, a direct comparison between estimated parameters was made. Spatially, no noteworthy differences were found. Regarding the temporal parameters, the extra freedom of

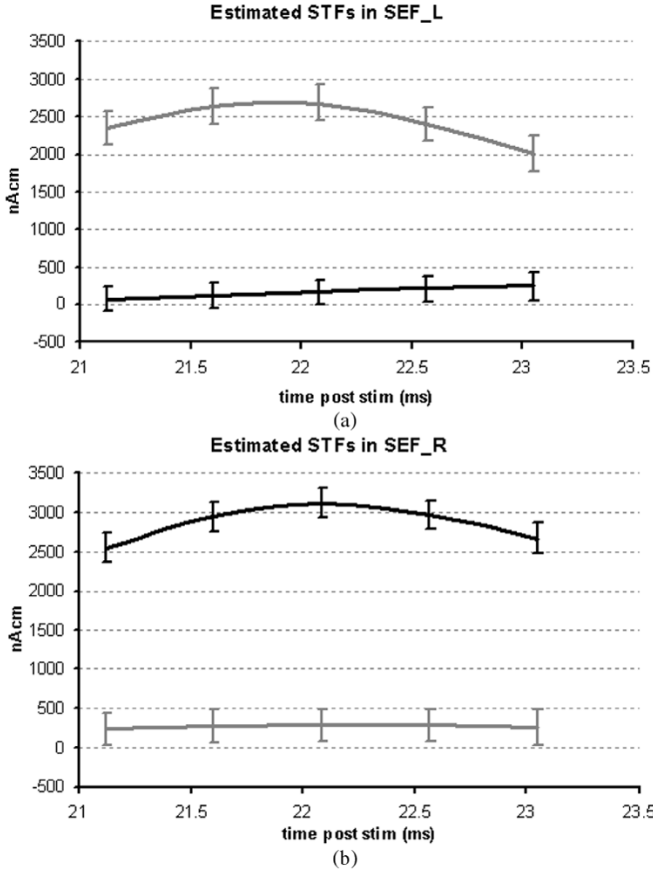


Fig. 3. ECDM estimated STFs of the semi-symmetric sources in the SEF data sets of subject 5. The black (grey) line indicates the estimated amplitude of left (right) source. The vertical error bars indicate a 95% confidence region around the amplitudes. The horizontal axis represents the post stimulus time in milliseconds, and the vertical axis shows the amplitude in nAcm. (a) SEF left median nerve. (b) SEF right median nerve.

TABLE II

P VALUES OF THE ESTIMATED IPSILATERAL ACTIVITY IN BOTH THE SEF DATA SETS OF THE FIVE SUBJECTS FOR THE DIFFERENT VALUES OF  $Z$

subject	Z=1		Z=2		Z=3		Z=4	
	SEF <sub>L</sub>	SEF <sub>R</sub>	SEF <sub>L</sub>	SEF <sub>R</sub>	SEF <sub>L</sub>	SEF <sub>R</sub>	SEF <sub>L</sub>	SEF <sub>R</sub>
1	10.97%	24.23%	38.60%	53.59%	0.00%	0.00%	0.00%	0.02%
2	1.14%	0.82%	1.67%	3.52%	6.46%	17.97%	31.55%	45.79%
3	33.73%	42.36%	62.51%	67.85%	3.98%	8.13%	15.90%	27.99%
4	30.13%	55.34%	55.26%	71.71%	4.78%	16.96%	26.26%	36.42%
5	2.36%	2.25%	5.50%	9.66%	0.34%	1.42%	3.32%	3.88%

the ECDM was manifested by different wave shapes. For one of the five subjects, somewhat different results were reported: The semi-symmetric sources were closer together in the ECDM, and the magnitudes were higher. This leads to the phenomenon of canceling sources, as discussed in [3]. Apparently, the somewhat problematic data of subject 3 need a very strict model, enabling a plausible solution: The separate analyses of the three data sets yield unlikely solutions; the ECDM yields a better parameter estimate, which is still a little doubtful, and the CDM generates the most plausible solution comparable with solutions of other subjects. In these kinds of circumstances, model selection procedures [33] are helpful to decide on the correct model.

The estimated activity of supposedly *silent* sources was investigated in the SEF data analysis. The N20 response to an SEF stimulus has been reported, arising from the contralateral

SI area only [15], [20]. Activity of both contra- and ipsilateral sources was estimated by fitting a semi-symmetric source model to SEF data of left and right median nerve stimulation of five subjects. Computed  $p$  values of estimated ipsilateral activity were generally high, although some of the values turned out to be significant. A closer look at the confidence interval for each time sample revealed that this significance can change from sample to sample. A few remarks can be made here. In general, the Ordinary Least Squares method underestimates the standard errors of the model parameters, while the Generalized Least Squares method, which is performed through prewhitening in the ECDM, yields a correct estimate of the standard error (and thus of the confidence intervals) (e.g., [13]). Nevertheless, it seems tenable from Fig. 3 that the ECDM confidence intervals are still slightly underestimated. This can be interpreted in line with a feature reported in [4], where it was shown that the source localization suffers from comparable frequencies in source time functions and the temporal covariance. In such a case, not only noise, but signal as well, is removed by prewhitening. In a simulation study, which is not reported here, this phenomenon was studied and confirmed; the ratio between estimated standard errors and sample variances of the estimated amplitudes equalled approx. 0.70 for frequencies below 15 Hz and 0.92 for frequencies above 20 Hz. As a final remark, the liberal way of testing may have produced confidence intervals that are too optimistic; a correction (e.g., Bonferroni) might indicate nonsignificance. Bearing these remarks in mind, the existence of ipsilateral activity is neither approved nor disapproved by this ECDM study.

Summarizing, the ECDM is an alternative to the CDM, which is statistically more convenient, but does not allow for exploitation of prior information. It is up to the user's careful consideration whether to use the CDM or the ECDM.

## APPENDIX

### A. Dimensions and Variables

The dimensions are defined as

- $I$ —# sensors;
- $J$ —# time samples;
- $K_q$ —# trials data set  $q$ ;
- $K$ —# trials all data sets;
- $M$ —# source parameters;
- $P$ —# source time functions of common sources;
- $Q$ —# data sets;
- $Z$ —# common STFs;

and the variables as

- $\zeta$ — $(M \times 1)$  vector with all source parameters;
- $\eta$ — $(3P \times 1)$  vector with source orientation parameters;
- $\xi$ — $((M - 3P) \times 1)$  vector with source location parameters;
- $\theta$ — $((PQZ + M + ZJ) \times 1)$  vector with all parameters;
- $A(\xi, \eta)$ — $(I \times P)$  dipole field matrix of common sources;
- $\mathbf{A}$ — $(IQ \times PQ)$  matrix  $\mathbf{I}_Q \otimes \mathbf{A}$ ;
- $\mathbf{B}$ — $(Z \times J)$  common STF matrix, (STFs as rows);
- $\mathbf{B}$ — $(Z \times J)$  prewhitened STF matrix;
- $\mathbf{c}(\theta)$ — $((2P + Z^2) \times 1)$  constraint vector;
- $C_q$ — $(P \times Z)$  coupling matrix data set  $q$ ;

- $\mathbf{C}$ —( $QP \times Z$ ) stacked coupling matrices;
- $D(\theta)$ —( $(2P + Z^2) \times (PQZ + M + ZJ)$ ) derivative matrix of  $\mathbf{c}(\theta)$ ;
- $E_q^k$ —( $I \times J$ ) noise matrix trial  $k$  data set  $q$ ;
- $\mathbf{I}_n$ —( $n \times n$ ) identity matrix;
- $R_q^k$ —( $I \times J$ ) data matrix trial  $k$  data set  $q$ ;
- $R_q$ —( $I \times J$ ) model matrix data set  $q$ ;
- $\bar{R}_q$ —( $I \times J$ ) average data matrix data set  $q$ ;
- $T$ —( $J \times J$ ) temporal covariance matrix;
- $U$ —( $J \times J$ ) orthogonal matrix with singular vectors of  $\mathbf{R}^t \mathbf{A} (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R}$ ;
- $U_Z$ —( $J \times Z$ ) first  $Z$  columns of  $U$ ;
- $X$ —( $I \times I$ ) spatial covariance matrix;
- $\Sigma_\theta$ —( $(PQZ + M + ZJ) \times (PQZ + M + ZJ)$ ) constrained CRB matrix.

B. Estimator for  $B_0$

In this Appendix, the ML-estimator for  $B_0$  is derived from (22). Define the matrix  $P := B_0^t (B_0 B_0^t)^{-1} B_0$ , which is a projection with  $\text{rank}(P) = \text{rank}(B_0) = \text{rank}(\mathbf{B}) = Z$ . Therefore,  $P$  can be decomposed as  $P = HH^t$ , with  $H$  a ( $J \times Z$ ) matrix having  $Z$  orthonormal columns. If the  $j$ th ( $1 \times Z$ ) row vector of  $H$  is written as  $\mathbf{h}_j^t$ , the trace in (22) becomes

$$\begin{aligned} \text{tr}(\Delta P) &= \text{tr}(\Delta HH^t) \\ &= \text{tr} \left[ \begin{pmatrix} \mu_1 \mathbf{h}_1^t \\ \vdots \\ \mu_J \mathbf{h}_J^t \end{pmatrix} (\mathbf{h}_1 \ \cdots \ \mathbf{h}_J) \right] \\ &= \sum_{j=1}^J \mu_j \mathbf{h}_j^t \mathbf{h}_j \end{aligned} \tag{44}$$

with  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_J \geq 0$ . Now, the trace of a projection matrix equals its rank, and therefore

$$\text{tr}(P) = \text{tr}(HH^t) = \sum_{j=1}^J \mathbf{h}_j^t \mathbf{h}_j = Z. \tag{45}$$

Defining the vectors  $\mathbf{a} = (a_1, \dots, a_J)$  with  $a_j = \mathbf{h}_j^t \mathbf{h}_j$  and  $\mu = (\mu_1, \dots, \mu_J)$ , the maximization problem in (22) is reduced to

$$\begin{aligned} &\max_{\mathbf{a}} \mathbf{a}^t \mu \\ &\text{subject to} \begin{cases} 0 \leq a_j \leq 1 \\ \sum_j a_j = Z \\ \mu_j \geq \mu_{j+1}, \text{ for } j = 1, \dots, J-1 \\ \mu_j \geq 0 \ \forall j. \end{cases} \end{aligned} \tag{46}$$

Obviously, the optimal vector is given by  $\hat{\mathbf{a}} = (1, \dots, 1, 0, \dots, 0)$ , of which only the first  $Z$  elements equal 1. In case of less than  $Z$  nonzero  $\mu_j$ , this optimal solution for  $\mathbf{a}$  is not unique. Yet,  $Z$  exceeding the number of nonzero  $\mu_j$ , in (20), indicates overparametrization; hence, it may be assumed that  $\hat{\mathbf{a}}$  is unique. This implies that for  $H$

$$\mathbf{h}_j \mathbf{h}_j^t = \begin{cases} 1, & \text{for } 1 \leq j \leq Z \\ 0, & \text{for } Z + 1 \leq j \leq J. \end{cases} \tag{47}$$

Writing

$$H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} \tag{48}$$

with  $H_1$ , a ( $Z \times Z$ ) matrix and  $H_2$  a ( $(J - Z) \times Z$ ) matrix, (47) implies that  $H_1$  is orthogonal, and  $H_2$  equals the null matrix. The projection  $P$  then becomes

$$P = HH^t = \begin{pmatrix} \mathbf{I}_Z & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{49}$$

Recalling  $P = B_0^t (B_0 B_0^t)^{-1} B_0$ , the estimator for  $B_0$  is

$$\hat{B}_0 = G (\mathbf{I}_Z \ \mathbf{0}) \tag{50}$$

$\mathbf{0}$  being the  $Z \times (J - Z)$  null matrix and  $G$  any nonsingular ( $Z \times Z$ ) matrix.

C. Linear Estimator for the Source Orientations

In this Appendix, we derive an estimator for the source orientations, which exploits the linear dependency of the model on the orientations. First,  $A$  is decomposed using the linear dependency in  $\eta : W_X^t A(\xi, \eta) = \Phi(\xi) \Omega(\eta)$ , where  $\Phi(\xi)$  is the ( $I \times 3P$ ) prewhitened unit dipole field matrix in the three Euclidean moment directions, and  $\Omega(\eta)$  is the ( $3P \times P$ ) block diagonal matrix containing  $P$  ( $3 \times 1$ ) moment blocks on the diagonal. Next, this decomposition is substituted in (15), which then is maximized with respect to  $\eta$ , or equivalently

$$\min_{\eta} \text{tr}((\mathbf{R} - (\mathbf{I}_Q \otimes \Phi \Omega) \mathbf{C} \mathbf{B})(\mathbf{R} - (\mathbf{I}_Q \otimes \Phi \Omega) \mathbf{C} \mathbf{B})^t) \tag{51}$$

for known  $\xi$ ,  $\mathbf{B}$ , and  $\mathbf{C}$ . This quadratic function of  $\eta$  is minimized when the derivatives with respect to all the orientation parameters  $\eta_m, m = 1, \dots, 3P$ , are zero, i.e.,

$$\begin{aligned} &\text{tr} \left( \left( \left( \mathbf{I}_Q \otimes \Phi \frac{\partial \Omega}{\partial \eta_m} \right) \mathbf{C} \mathbf{B} \right)^t (\mathbf{R} - (\mathbf{I}_Q \otimes \Phi \Omega) \mathbf{C} \mathbf{B}) \right) \\ &= \sum_{j=1}^J \left( \left( \mathbf{I}_Q \otimes \Phi \frac{\partial \Omega}{\partial \eta_m} \right) \mathbf{C} \mathbf{b}_j \right)^t (\mathbf{r}_j - (\mathbf{I}_Q \otimes \Phi \Omega) \mathbf{C} \mathbf{b}_j) \\ &= 0 \end{aligned} \tag{52}$$

for all  $m$ , where  $\mathbf{r}_j$  and  $\mathbf{b}_j$  denote the  $j$ th column in  $\mathbf{R}$  and  $\mathbf{B}$ , respectively. Defining

$$A_m(j) = \left( \mathbf{I}_Q \otimes \Phi \frac{\partial \Omega}{\partial \eta_m} \right) \mathbf{C} \mathbf{b}_j, \text{ for } m = 1, \dots, 3P \tag{53}$$

$$A(j) = (A_1(j) \ A_2(j) \ \cdots \ A_{3P}(j)) \tag{54}$$

and noting that  $\Omega = \sum_{m=1}^{3P} (\partial \Omega / \partial \eta_m) \eta_m$ , (52) can be rewritten as

$$\sum_{j=1}^J (A_m(j))^t \mathbf{r}_j = \sum_{m'=1}^{3P} \sum_{j=1}^J (A_m(j))^t A_{m'}(j) \eta_{m'} \tag{55}$$

for all  $m$ . Finally, stacking the scalar (55) for all  $m$  as entries in one column, the following equation is obtained:

$$\sum_{j=1}^J (A(j))^t \mathbf{r}_j = \sum_{j=1}^J (A(j))^t A(j) \eta \tag{56}$$

and the estimator in (27), which requires a  $3P \times 3P$  inverse, follows. An alternative expression for the linear system in (56) can be found in [3].

#### D. Derivative Matrix of the Constraints

In this Appendix, the derivative of the constraint vector  $\mathbf{c}(\theta)$  with respect to all parameters is computed. Define the number of common sources to be  $N$ . Then, the number of source location parameters is  $3N$  and that of source orientations is  $3P$ , together yielding  $M = 3N + 3P$ . The constraints are

$$\begin{aligned} \eta_{3p-2}^2 + \eta_{3p-1}^2 + \eta_{3p}^2 \\ &= 1 \text{ for } p = 1, \dots, P \text{ (normality constraint)} \\ \xi_{3n-2}\eta_{3p-2} + \xi_{3n-1}\eta_{3p-1} + \xi_{3n}\eta_{3p} \\ &= 0 \text{ for } p = 1, \dots, P \text{ (tangentiality constraint)} \\ (\mathbf{B}^t \mathbf{U})_{z_1, z_2} \\ &= \delta_{z_1, z_2} \text{ for } z_1, z_2 = 1, \dots, Z \end{aligned}$$

where  $n$  in the tangentiality constraints is chosen such that the  $p$ th STF corresponds to the  $n$ th source. Furthermore,  $\delta_{i,j}$  denotes the Kronecker Delta function. The dimension of  $D(\theta)$  in (39) is  $(2P + Z^2) \times (PQZ + M + ZJ)$ . The first  $2P$  rows of  $D(\theta)$  are straightforward. For the last  $Z^2$  rows, we have

$$\frac{\partial (\mathbf{B}^t \mathbf{U})_{z_1, z_2}}{\partial B_{z,j}} = \delta_{z, z_1} U_{z_2, j} \quad (57)$$

$$\frac{\partial (\mathbf{B}^t \mathbf{U})_{z_1, z_2}}{\partial \zeta_m} = \mathbf{B}^t_{z_1} \frac{\partial U_{z_2}}{\partial \zeta_m}. \quad (58)$$

In [18], the eigenvector  $u_0$  of matrix  $X$  corresponding to eigenvalue  $\lambda_0$  is differentiated with respect to the corresponding matrix  $X$ :

$$du = (\lambda_0 \mathbf{I} - X)^+ (dX)u_0. \quad (59)$$

$(\cdot)^+$  stands for the Moore–Penrose inverse. Applying (59) and the chain rule to (58) substituting  $X = \mathbf{R}^t \mathbf{A} (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R}$ , which is a function of  $\zeta_m$  and  $u_0 = U_{z_2}$  with corresponding eigenvalue  $\lambda_{z_2}$ , yields

$$\begin{aligned} \frac{\partial U_{z_2}}{\partial \zeta_m} &= (\lambda_{z_2} \mathbf{I}_J - \mathbf{R}^t \mathbf{A} (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R})^+ \\ &\quad \times \frac{\partial (\mathbf{R}^t \mathbf{A} (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{R})}{\partial \zeta_m} U_{z_2}. \quad (60) \end{aligned}$$

Employing the product rule,  $d_X(X^{-1}) = -X^{-1}dXX^{-1}$ , and (60), the matrix  $D(\theta)$  can be calculated.

#### E. F-Test on Estimated Coupling Parameters

In this Appendix, the F-test on a subset of the estimated parameters under constraints is derived, following the derivation in [27]. First, the test of unconstrained parameters is recalled, and after that, the case of constrained parameters is described.

The following  $IJQ \times 1$  vectors are defined:

$$\mathbf{r} = \text{vec}(\mathbf{R}) \quad (61)$$

$$\mathbf{f}(\theta) = \text{vec}(\mathbf{A}\mathbf{C}\mathbf{B}) \quad (62)$$

$$\boldsymbol{\varepsilon} = \mathbf{r} - \mathbf{f}(\theta^*) \quad (63)$$

with  $\theta^*$  indicating the true value of  $\theta$ . For the unconstrained parameter vector  $\theta$ , we have the  $\mathcal{F}$  statistic

$$F_{\text{test}} = \frac{(\hat{\theta} - \theta^*)^t F^t F (\hat{\theta} - \theta^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\theta}{n_\theta} \sim \mathcal{F}_{n_\theta, IJQ - n_\theta} \quad (64)$$

where  $F = F(\theta) = (\partial \mathbf{f}(\theta) / \partial \theta)(\theta)$ ,  $n_\theta$  indicates the length of  $\theta$ , and  $\|\cdot\|$  indicates the Euclidean vector norm. To test a subset of the estimated parameters, write

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \quad (65)$$

$$F = (F_1 \ F_2). \quad (66)$$

As outlined in [27], testing the subset of  $n_{\theta_2}$  parameters in  $\theta_2$  is performed by the test statistic

$$\begin{aligned} F_{\text{subset}} &= \frac{(\hat{\theta}_2 - \theta_2^*)^t F_2^t (I - P_{F_1}) F_2 (\hat{\theta}_2 - \theta_2^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\theta}{n_{\theta_2}} \\ &\sim \mathcal{F}_{n_{\theta_2}, IJQ - n_\theta} \quad (67) \end{aligned}$$

where  $P_{F_1} = F_1 (F_1^t F_1)^{-1} F_1^t$ .

In the case of the constrained ECDM parameters, a parameter transformation

$$\beta = V^t \theta \quad (68)$$

of length  $PQZ + M + ZJ - 2P - Z^2$  is used. This new parameter  $\beta$  is unconstrained, because  $\mathcal{I}(\beta) = V^t \mathcal{I}V$  is nonsingular. Therefore, the above test statistics in (64) and (67) can be applied to  $\beta$ . Choosing the parameter subset consisting of the entries of a certain row in  $\mathbf{C}$ , which are left unchanged by (68), the final test statistic becomes

$$\begin{aligned} F_{\text{coupling}} &= \frac{(\hat{\theta}_2 - \theta_2^*)^t ([\Sigma_\theta]_{\theta_2})^{-1} (\hat{\theta}_2 - \theta_2^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\beta}{n_{\theta_2}} \\ &\sim \mathcal{F}_{n_{\theta_2}, IJQ - n_\beta} \quad (69) \end{aligned}$$

where  $[\cdot]_{\theta_2}$  denotes the submatrix corresponding to the parameters in  $\theta_2$ .

#### F. Confidence Regions Around Estimated STFs

In this Appendix, the confidence region around the estimated STFs for each source in each data set is computed. The same notation is used as in Appendix E. In [27], testing the  $l$ -dimensional nonlinear hypothesis  $\mathbf{h}(\theta) = \mathbf{h}^*$  is outlined using a linear approximation. Defining  $H(\theta) = (\partial \mathbf{h}(\theta) / \partial \theta)$ , this yields the following test statistic:

$$\begin{aligned} F_{\text{hypstest}} &= \frac{(\mathbf{h}(\hat{\theta}) - \mathbf{h}(\theta^*))^t (H \Sigma_{\hat{\theta}} H^t)^{-1} (\mathbf{h}(\hat{\theta}) - \mathbf{h}(\theta^*))}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \\ &\quad \times \frac{IJQ - (PQZ + M + ZJ - 2P - Z^2)}{l} \\ &\sim \mathcal{F}_{l, IJQ - (PQZ + M + ZJ - 2P - Z^2)}. \quad (70) \end{aligned}$$

The estimated STF  $p$  at a certain time sample  $j$  in a certain data set  $q$  is a  $(1 \times 1)$  hypothesis vector  $h_{p,q,j}$ . The derivative matrix  $H$  is now a  $(1 \times (PQZ + M + ZJ))$  vector. Furthermore,  $h_{p,q,j}$  is the product of the  $p$ th row in the  $q$ th coupling matrix and the

$j$ th column of the common STF matrix. Hence, for a given  $q, p$ , and  $j$ ,  $H(\theta)$  is determined by

$$h_{q,p,j}(\theta) = (C_q B)_{p,j} = (C_q \mathbf{B} W_T^{-1})_{p,j} \quad (71)$$

$$\frac{\partial h_{q,p,j}(\theta)}{\partial (C_q)_{p,z}} = (\mathbf{B} W_T^{-1})_{z,j} \quad (72)$$

$$\frac{\partial h_{q,p,j}(\theta)}{\partial \mathbf{B}_{z,j'}} = (C_q)_{p,z} (W_T^{-1})_{j',j} \quad (73)$$

with  $z = 1, \dots, Z$ , and  $j' = 1, \dots, J$ . Using the threshold value  $F_{\text{thres}} = \mathcal{F}_{1, IJQ - (PQZ + M + ZJ - 2P - Z^2)}^\alpha$ , the  $(1-\alpha)$  confidence region for a certain  $h_{q,p,j}(\theta)$  is given by

$$\left\{ h \mid |h_{q,p,j}(\hat{\theta}) - h| \leq \sqrt{\frac{F_{\text{thres}} \|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2 (H \Sigma_{\hat{\theta}} H^t)}{IJQ - (PQZ + M + ZJ - 2P - Z^2)}} \right\}. \quad (74)$$

REFERENCES

[1] A. Achim and S. Bouchard, "Toward a dynamic topographic components model," *Electroenc. Clin. Neur.*, vol. 103, pp. 381–385, 1997.

[2] F. Bijma, J. C. de Munck, H. M. Huizenga, and R. M. Heethaar, "A mathematical approach to the temporal stationarity of background noise in MEG/EEG measurements," *NeuroImag.*, vol. 20, no. 1, pp. 233–243, 2003.

[3] F. Bijma, J. C. de Munck, K. B. E. Böcker, H. M. Huizenga, and R. M. Heethaar, "The coupled dipole model: An integrated model for multiple MEG/EEG data sets," *NeuroImag.*, vol. 23, no. 3, pp. 890–904, 2004.

[4] J. C. de Munck, H. M. Huizenga, L. J. Waldorp, and R. M. Heethaar, "Estimating stationary dipoles from MEG/EEG data contaminated with spatially and temporally correlated background noise," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1565–1572, Jul. 2002.

[5] J. C. de Munck, F. Bijma, P. Gaura, C. Sieluzycycki, M. I. Branco, and R. M. Heethaar, "A mathematical description of habituation effects in multi channel MEG/EEG data," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 12, pp. 2123–2128, Dec. 2004.

[6] A. Dogandžić and A. Nehorai, "Localization of evoked electric sources and design of EEG/MEG sensor arrays," in *Proc. 9th IEEE SP Workshop Stat. Signal Array Process.*, Portland, OR, Sep. 1998, pp. 228–231.

[7] —, "Estimated evoked dipole responses in unknown spatially correlated noise with EEG/MEG arrays," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 13–25, Jan. 2000.

[8] —, "Generalized multivariate analysis of variance," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 39–54, Sep. 2003.

[9] M. Fuchs, M. Wagner, and J. Kastner, "Confidence limits of dipole source reconstruction results," *Clin. Neurophys.*, vol. 115, pp. 1442–1451, 2004.

[10] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: John Hopkins Univ. Press, 1990.

[11] H. M. Huizenga and P. C. M. Molenaar, "Estimating and testing the sources of evoked-potentials in the brain," *Multivar. Behav. Res.*, vol. 29, no. 3, pp. 237–262, 1994.

[12] H. M. Huizenga, T. L. van Zuijlen, and P. C. M. Molenaar, "Simultaneous MEG and EEG source analysis," *Psys. Med. Biol.*, vol. 46, no. 7, pp. 1737–1751, 2001.

[13] H. M. Huizenga, J. C. de Munck, L. J. Waldorp, and R. P. P. P. Grasman, "Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 6, pp. 533–539, Jun. 2002.

[14] H. M. Huizenga, D. J. Heslenfeld, and P. C. M. Molenaar, "Optimal measurement conditions for spatiotemporal EEG/MEG source analysis," *Psychometrika*, vol. 67, no. 2, pp. 299–313, 2002.

[15] J. Huttunen, R. Hari, and L. Leinonen, "Cerebral magnetic responses to stimulation of ulnar and median nerves," *Electroenc. Clin. Neur.*, vol. 66, pp. 391–400, 1987.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[17] C. G. Khatri, *A Note on a MANOVA Model Applied to Problems in Growth Curve*: Ann. Inst. Statist. Math., 1966, vol. 18, pp. 75–86.

[18] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*, revised ed. Chichester, U.K.: Wiley, 1995.

[19] J. Maier, G. Dagnelie, H. Spekreijse, and B. W. van Dijk, "Principal components analysis for source localization of veps in man," *Vision Res.*, vol. 27, no. 2, pp. 165–177, 1987.

[20] F. Mauguière, I. Merlet, N. Forss, S. Vanni, V. Jousmäki, P. Adeleine, and R. Hari, "Activation of a distributed somatosensory cortical network in the human brain. A dipole modeling study of magnetic fields evoked by median nerve stimulation. Part I: Location and activation timing of sef sources," *Electroenc. Clin. Neur.*, vol. 104, pp. 281–289, 1997.

[21] C. D. McGillem and J. I. Aunon, "Analysis of event-related potentials," in *Methods of Analysis of Brain Electrical and Magnetic Signals, EEG Handbook*, A. S. Gevins and A. Remond, Eds. Amsterdam, The Netherlands: Elsevier, 1987, ch. 5.

[22] F. Miwakeichi, E. Martínez-Montes, P. A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, and Y. Yamaguchi, "Decomposing EEG data into space-time-frequency components using parallel factor analysis," *NeuroImag.*, vol. 22, no. 3, pp. 1035–1045, 2004.

[23] J. Möcks, "Topographic components models for event related potentials and some biophysical considerations," *IEEE Biomed. Eng.*, vol. 35, no. 6, pp. 482–484, June 1988.

[24] J. C. Mosher, M. E. Spencer, R. M. Leahy, and P. S. Lewis, "Error bounds for EEG and MEG dipole source localization," *Electroenc. Clin. Neur.*, vol. 86, pp. 303–321, 1993.

[25] C. H. Muravchik and A. Nehorai, "EEG/MEG error bounds for a static dipole source with a realistic head model," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 470–484, Mar. 2001.

[26] C. R. Rao, *Linear Statistical Inference and its Applications*, 2nd ed. New York: Wiley, 1973.

[27] G. A. F. Seber and C. J. Wild, *Nonlinear Regression*. New York: Wiley, 1989.

[28] M. S. Srivastava and C. G. Khatri, *An Introduction to Multivariate Statistics*. New York, North-Holland, 1979.

[29] P. Stoica and B. C. Ng, "On the Cramér-Rao Bound under parametric constraints," *IEEE Signal Process. Lett.*, vol. 5, no. 7, pp. 177–179, Jul. 1998.

[30] B. Turetsky, J. Raz, and G. Fein, "Representation of multi-channel evoked potential data using a dipole component model of intracranial generators: Application to the auditory," *Electroenc. Clin. Neur.*, vol. 76, pp. 540–556, 1990.

[31] E. F. Vonesh and V. M. Chinchilli, *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. New York: Marcel Dekker, 1997.

[32] L. J. Waldorp, H. M. Huizenga, C. V. Dolan, and P. C. M. Molenaar, "Estimated generalized least squares electromagnetic source analysis based on a parametric noise covariance model," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 6, pp. 737–741, Jun. 2001.

[33] L. J. Waldorp, H. M. Huizenga, R. P. P. P. Grasman, K. B. E. Böcker, J. C. de Munck, and P. C. M. Molenaar, "Model selection in electromagnetic source analysis with an application to veps," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1121–1129, Oct. 2002.

[34] K. M. Wang, H. Begleiter, and B. Porjesz, "Trilinear modeling of event-related potentials," *Brain Topograph.*, vol. 12, no. 4, pp. 263–271, 2000.



**Fetsje Bijma** received the M.Sc. degree in pure mathematics (*cum laude*) from the RijksUniversiteit Groningen, Groningen, The Netherlands, in 2000. She received the Ph.D. degree (*cum laude*) from Vrije Universiteit, Amsterdam, The Netherlands, in 2005.

Her research interests include inverse modeling of MEG/EEG data as well as the simultaneous analysis of multiple data sets. She is also working on the statistical features of spatiotemporal MEG residuals.



**Jan C. de Munck** was born 1962. He received the degree in experimental physics (*cum laude*) and the Ph.D. degree (*cum laude*) from the University of Amsterdam, Amsterdam, The Netherlands, in 1985 and 1989, respectively.

His research interest is in forward and inverse modeling applied in medicine (EEG/MEG/EIT) and functional brain imaging (PET/fMRI/MEG). Currently, he is a state member with the Vrije Universiteit Medical Center, Amsterdam.



**Rob M. Heethaar** was born in the Netherlands in 1946. He studied physics and mathematics from 1963 to 1968 at Utrecht University, Utrecht, The Netherlands.

At Utrecht University, he held joint appointments with the Departments of Cardiology and Medical Physics from 1969 to 1985. He completed his Ph.D. study on atrioventricular conduction in the rat heart in 1972 and worked for the next two years on the physical principles of artificial electrical stimulation of the human heart. From 1974 to 1976, he was a research fellow in the fields of cardiac mechanics and computer tomography at the Mayo Clinic, Rochester, MI. In 1979, he became professor of Bioengineering at the University of Technology, Twente, The Netherlands, and in 1987, he became Professor of medical physics and informatics at Utrecht University. He founded two Dutch companies in bioengineering. Since 1993, he has been Professor of medical physics at the Vrije Universiteit (VU), Amsterdam, The Netherlands, and since 2002, he has been head of the Department of Physics and Medical Technology of the VU Medical Centre. His main interests are physics of heart, circulation, lungs, and the brain. He is the author of numerous papers on electrophysiology, cardiac mechanics, bioengineering, and hemorheology. He holds some patents in the field of bioengineering, and he is referee of several scientific journals. His current topics of research are related to the development of bioelectrical and MRI-tagging techniques to assess the function of heart, circulation, and lungs noninvasively.

Dr. Heethaar has received several awards for his contributions to cardiology.



**Hilde M. Huizenga** was born in 1965. She received the M.D. degree in psychology from the University of Groningen, Groningen, the Netherlands, in 1990 and the Ph.D. degree (*cum laude*) from the University of Amsterdam, Amsterdam, the Netherlands, in 1995.

Since 1999, she has directed the statistical EEG/MEG source analysis research program. Since 2003, she has been an associate professor at the University of Amsterdam. Her main research interest is the statistical analysis of neuroscientific (EEG/MEG/fMRI) data. In particular, (non)linear

regression, mean and covariance structure analysis, (non)linear mixed models, multivariate statistics, and optimal design.

Dr. Huizenga was awarded a fellowship of the Royal Netherlands Academy of Arts and Sciences from 1996 to 2002. For her dissertation "The statistical approach to electromagnetic source localization in the brain," she received the International Psychometric Society Dissertation Award.



**Arye Nehorai** (S'80-M'83-SM'90-F'94) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion, Haifa, Israel, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

From 1985 to 1995, he was a faculty member with the Department of Electrical Engineering, Yale University, New Haven, CT. In 1995, he joined the Department of Electrical Engineering and Computer Science, The University of Illinois at Chicago (UIC), as a Full Professor. From 2000 to 2001, he was Chair

of the department's Electrical and Computer Engineering (ECE) Division, which is now a new department.

Dr. Nehorai is Vice President (Publications) of the IEEE Signal Processing Society. He is also Chair of the Publications Board and a member of the Board of Governors and the Executive Committee of this Society. He was Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2000 to 2002 and is the founding editor of the special columns on Leadership Reactions in the IEEE SIGNAL PROCESSING MAGAZINE. He was corecipient, with P. Stoica, of the 1989 IEEE Signal Processing Society's Senior Award for Best Paper and co-author of the 2003 Young Author Best Paper Award of this Society with A. Dogandzic. He was elected Distinguished Lecturer of the IEEE Signal Processing Society for the term 2004 to 2005. He has been a Fellow of the Royal Statistical Society since 1996. In 2001, he was named University Scholar of the University of Illinois.