A Maximum-Likelihood Estimator for Trial-to-Trial Variations in Noisy MEG/EEG Data Sets

Jan Casper de Munck*, Fetsje Bijma, Pawel Gaura, Cezary Andrzej Sieluzycki, Maria Inês Branco, and R. M. Heethaar

Abstract—The standard procedure to determine the brain response from a multitrial evoked magnetoencephalography (MEG) or electroencephalography (EEG) data set is to average the individual trials of these data, time locked to the stimulus onset. When the brain responses vary from trial-to-trial this approach is false. In this paper, a maximum-likelihood estimator is derived for the case that the recorded data contain amplitude variations. The estimator accounts for spatially and temporally correlated background noise that is superimposed on the brain response.

The model is applied to a series of 17 MEG data sets of normal subjects, obtained during median nerve stimulation. It appears that the amplitude of late component (30–120 ms) shows a systematic negative trend indicating a weakening response during stimulation time. For the early components (20–35 ms) no such a systematic effect was found. The model is furthermore applied on a MEG data set consisting of epileptic spikes of constant spatial distribution but varying polarity. For these data, the advantage of applying the model is that positive and negative spikes can be processed with a single model, thereby reducing the number of degrees of freedom and increasing the signal-to-noise ratio.

Index Terms—Covariance, habituation, maximum-likelihood, MEG noise.

I. INTRODUCTION

S PONTANEOUS brain activity in magnetoencephalography (MEG) and electroencephalography (EEG) signals is caused by dendritic currents of neurons receiving signals from connected neurons. Part of this activity, like the alpha and mu rhythms are enhanced in the absence of visual and motor input, respectively. When a stimulus is presented to a subject several parts of the brain that are involved in processing the stimulus will show increased neuronal activity and will, therefore, act as the generators of the brain response, which can be recorded with MEG or EEG. When the same stimulus is applied repeatedly, the brain response is usually extracted from the recorded data by a simple averaging technique. The question is how the averaged response should be interpreted.

Manuscript received July 24, 2003; revised May 2, 2004. Asterisk indicates corresponding author.

*J. C. de Munck is with the Department of Physics and Medical Technology of University Hospital of the Vrije Universiteit, 1081-HV Amsterdam, The Netherlands (e-mail: jc.munck@vumc.nl).

F. Bijma and R. M. Heethaar are with the Department of Physics and Medical Technology of University Hospital of the Vrije Universiteit, 081-HV Amsterdam, The Netherlands.

P. Gaura and C. A. Sieluzycki are with the Department of Measuring and Medical Electronic Instruments, Faculty of Fundamental Problems of Technology, Wroclaw University of Technology, 50-370 Wroclaw, Poland.

M. I. Branco is with the Institute of Biophysics and Biomedical Engineering, Faculty of Sciences, University of Lisbon, 1749-016 Lisbon, Portugal.

Digital Object Identifier 10.1109/TBME.2004.836515

The (most often implicit) assumption underlying response averaging is the "signal plus noise" model

$$R_{ij}^{(k)} = R_{ij} + \varepsilon_{ij}^{(k)}, \quad k = 0, \dots, K - 1.$$
 (1)

Here, $R_{ij}^{(k)}$ is the recorded signal at channel *i*, time sample *j* and trial *k*. Furthermore, R_{ij} is the spatio-temporal pattern of the brain response, which is assumed to be constant over trials. Finally, $\varepsilon_{ij}^{(k)}$ is the background noise, which is assumed to have a Gaussian distribution and which may be correlated over channels and samples. It is straightforward to show that the simple average of the recorded data over trials is the maximum-likelihood (ML) estimator of the constant brain response R_{ij} .

On the other hand, it is known from experiments that the constant response assumption is false in general. Examples of trial-to-trial variations in human EEG are habituation effects [1], [2], P300 effects and event-related synchronization and de-synchronization effects [3], [4]. However, because these effects are generally small, it is not straightforward to prove that these effects are really present in the data, and to distinguish them from the background noise. This problem becomes even more severe if one realizes that, since the background noise is mainly generated by spontaneous brain activity, the spatial and temporal properties of the noise are very similar to those of the brain response under study.

In a recent paper [5], Truccolo *et al.* demonstrate that neglecting trial-to-trial variations results in an estimate of the background noise of which the variance is nonstationary over the time interval of interest. However, that paper does not show how trial-to-trial variations can be estimated from "first principles." In [6] and [7], an ML model is formulated yielding estimators of amplitude and latency jitters of single trials in a multitrial evoked-potential experiment. A restriction of these papers is that they are based on the assumption of uncorrelated background noise, whereas we know that for instance the alpha rhythm is both correlated in time and over channels (e.g., [8] and [9]). Furthermore, it should be realized that estimates of trial-to-trial variations and of background noise are mutually related because the recorded signals partly belong to the noise, and partly to the brain response.

In this paper, an ML model is presented to describe and estimate habituation effects in multichannel multitrial data, without the restricting assumption of uncorrelated background noise. This model is applied to a series of standard experiments wherein the subject was electrically stimulated at the median nerve. Since the model is equally well applicable to the simultaneous analysis of epileptic spikes of varying polarity, an example of such data will be discussed in detail.

II. MATERIAL AND METHODS

A. The Model

In the data model it is assumed that the recorded brain signal $R_{ij}^{(k)}$ consists of a constant spatio-temporal pattern R_{ij} multiplied with a trial dependent amplitude factor $\alpha^{(k)}$, and a Gaussian noise part $\varepsilon_{ij}^{(k)}$, which is statistically independent from trial-to-trial

$$R_{ij}^{(k)} = \alpha^{(k)} R_{ij} + \varepsilon_{ij}^{(k)}, \quad k = 0, \dots, K - 1.$$
 (2)

The spatio-temporal covariance of $\varepsilon_{ij}^{(k)}$ is modeled as a Kronecker product $X \otimes T$ (e.g., [10]) of a spatial covariance matrix X and a temporal covariance matrix T. In a formula, the following is assumed for the noise covariance:

$$\mathbf{E}\left\{\varepsilon_{ij}^{(k)}\varepsilon_{i'j'}^{(k')}\right\} = \delta_{kk'}\{X \otimes T\}_{ij,i'j'}.$$
(3)

In this model, the unknown parameters are contained in $\alpha^{(k)}$, R_{ij} , X, and T and need to be estimated from the data $R_{ij}^{(k)}$.

The underlying idea of the ML estimator is to express the probability density distribution of the noise in terms of the *a priori* unknown parameters, and to derive those parameters for which this probability density reaches its maximum. With the above assumptions it is found that the probability density function is

$$f_{\varepsilon}(\alpha, R, X, T) = \frac{e^{-\frac{1}{2}\sum_{k} \operatorname{Tr}\left\{\left(R^{(k)} - \alpha^{(k)}R\right)^{T} X^{inv}\left(R^{(k)} - \alpha^{(k)}R\right)T^{inv}\right\}}}{(2\pi)^{\frac{IJK}{2}} \det(T)^{\frac{IK}{2}} \det(X)^{\frac{JK}{2}}}.$$
(4)

Here, Tr{} indicates the trace and det() indicates the determinant of the matrix within brackets. Furthermore, I is the number of channels, J is the number of time samples, and K is the number of trials. Finally, R indicates the response matrix, with elements R_{ij} and $R^{(k)}$ indicates the single trial data, with elements $R_{ij}^{(k)}$.

The ML estimators of α , R, X, and T are found by setting the corresponding derivatives of (4) equal to zero and solving the estimators from the resulting equations. Doing so for R it is found that R is the weighted average of the single trial data

$$R = \frac{1}{K} \sum_{k} \alpha^{(k)} R^{(k)} \tag{5}$$

and the weights $\alpha^{(k)}$ appear to be the elements of the eigenvector with the largest eigenvalue of the following system:

$$\sum_{k_2} \operatorname{Tr} \left\{ R^{(k_1)} X^{inv} R^{(k_2)T} T^{inv} \right\} \alpha^{(k_2)} = \lambda_{\max} \alpha^{(k_1)}.$$
 (6)

For the spatial and temporal covariances a pair of equations can be derived in which T is expressed in terms of X and vice versa

$$X = \frac{1}{J} \left(\frac{1}{K} \sum_{k} R^{(k)} T^{inv} R^{(k)T} - RT^{inv} R^T \right)$$
(7)

and

$$T = \frac{1}{I} \left(\frac{1}{K} \sum_{k} R^{(k)T} X^{inv} R^{(k)} - R^T X^{inv} R \right).$$
(8)

Note that in (5)–(8), as well as in the sequel, the symbols α , R, X, and T have been used to indicate the estimators of the true values. Equations (5)–(8) have to be solved iteratively. First, $\alpha = (1, 1, 1, ..., 1)^T$ and $T = I_J$ are taken, and then (7) and (8) are solved in iteration [9]. These first estimates of the covariances are substituted into (6) to obtain an update of α , which is substituted back into (5), etc. In practice, it appears that a few of these iterations are sufficient to obtain a stable solution.

A few remarks may shed some light on the meaning of these formulas. First, when the assumption is made that there are no trial-to-trial variations, $\alpha = (1, 1, 1, ..., 1)^{T}$, and the best estimate of the brain response reduces to the simple average over trials. Second, when the noises were white $(X = I_{I}, T = I_{J})$, and when $R_{ij}^{(k)}$ is considered as a matrix with a spatio-temporal index ij and trial index k, it can be understood that the solution of (6) corresponds to the strongest spatio-temporal pattern that is present in all trials. The effect of premultiplication and postmultiplication of the matrices X^{inv} and T^{inv} is a prewhitening of the recorded data. Finally, the iterative solution of (5)–(8) demonstrates the mutual influence of the trial-to-trial variability due to brain response and the trial-to-trial variability due to correlated noise.

B. The Data

To nine subjects (normal volunteers recruited from the lab) an electrical median nerve stimulation was applied, for both left hand and right hand (N = 8), or only left hand (N = 1). The stimulus intensity was individually adjusted such that a twitch of the thumb appeared. A regular inter-stimulus interval was used, which was set at 1 s. MEG data [somatosensory evoked fields (SEF)] were acquired on a whole head 151-channel system of CTF, Inc. using a sample rate of 2083 Hz. Offline, bad epochs and bad channels were marked and removed from the data. The number of remaining good responses varied between 450 and 570. Data were offset corrected using a prestimulus interval of 100 ms, which is optimal to reduce the alpha rhythm noise [11]. No bandpass filtering was used. For each subject two time intervals were chosen to investigate trial-to-trial variations. The early window extended from 23 to 37 ms after the stimulus, the late window extended from 30 to 120 ms. These windows were chosen based on the distribution of the averaged MEG data power as a function of time.

The MEG data of a patient with epilepsy was used to illustrate the applicability of the model for the integrated analysis of multiple spike data of varying polarity. Simple averaging would result in a vanishing of the signal. The MEG data were collected in a period of 1 hr, and stored in six data files of 10 min each. First, the data were motion-corrected [12] and spikes were marked at the maximum signal amplitude by an experienced MEG technician. Then, for each spike marker a symmetric data window was cut, 110 ms left and 110 ms right of the marker. These spikes were subject to cluster analysis [13], after which 254 spikes remained having a similar topography, but a varying amplitude

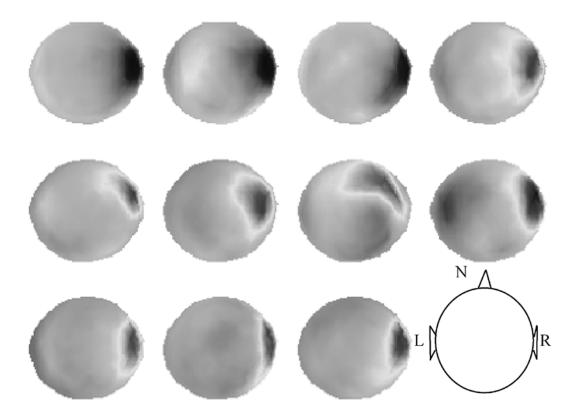


Fig. 1. Selection of maps is shown corresponding to the markers set by a technician on different spikes. The spatial pattern of these spikes are similar, but the polarity changes from spike to spike.

and polarity. The field maps of a selection of these spikes are presented in Fig. 1. One observes that most graphs only show one polarity, the other one false outside the MEG helmet. This situation often occurs with MEG spikes, in particular with temporal lobe epilepsy.

III. RESULTS

A. SEF Data

The algorithm presented in Section II usually converged in 8-12 iterations. Here, a quite strict stopping rule was used, implying that relative changes were smaller that 10^{-10} . For the late responses (187 samples) this resulted in a computation time of about 2 h (Pentium IV, 1 GHz) or more. For the early responses (29 samples) this time amounted typically 15 min. Fig. 2 shows four typical examples of the trial-to-trial variations in the (late) SEF response. In this figure, the trial multiplication factor $\alpha^{(k)}$ is plotted as a function of the stimulus time in s. The vertical scale is such that 1 implies the traditional constant response model. In all cases, there are relative fast variations from one trial to the next and also slower variations on scales between 50 and 200 s. To extract systematic behavior from all data sets, a straight line was fit to the amplitude data, using the stimulus times of each trial as x coordinate and the $\alpha^{(k)}$ as y coordinate. The quality of these line fits, were in all cases comparable to the examples presented in Fig. 2. In all of the cases presented in Fig. 2, the slope, or trend, of the line fit was negative, showing that the responses are weaker at the end of the measurement session than at the beginning. The other SEF data sets were treated similarly and for all data sets the linear trend was computed.

Fig. 3 summarizes the trends of the early and late responses for all acquired SEF data sets. The numbers on top refer to the subject. The light blocks correspond to the early components, the darker ones to the late ones. For each subject (except #4) first the left and then the right hand side response trends are depicted. The arrows correspond to the series presented in Fig. 2. From these data it appears that for the early responses the trend varies strongly over subjects, whereas for the late responses all trends are negative or slightly positive at most.

B. Spike Data

Fig. 4 shows the results of the same analysis for the spike data. Since the spikes originate from different data files, the multiplication factor $\alpha^{(k)}$ is now plotted as function of spike occurrence instead of time. It appears that 60 of the 254 spikes have a negative polarity with respect to the others. Furthermore, it seems as if the "negative" spikes have slightly smaller amplitude than the "positive" ones.

When the spikes are simply averaged, which is done in the first iteration of the estimation of amplitude modulation factors, it is predicted by Truccolo *et al.* [5], that the variance changes in time with a pattern corresponding to the shape of the averaged event. This effect is demonstrated in Fig. 5(a). After 9 iterations, when the algorithm converges, the variance has decreased and becomes much more constant. Fig. 5(b) and (c) shows that not only the variance, but the whole covariance matrix T becomes much more stationary. The remaining nonstationarity may be explained by a small misalignment in the spike markers, which introduce latency jitters.

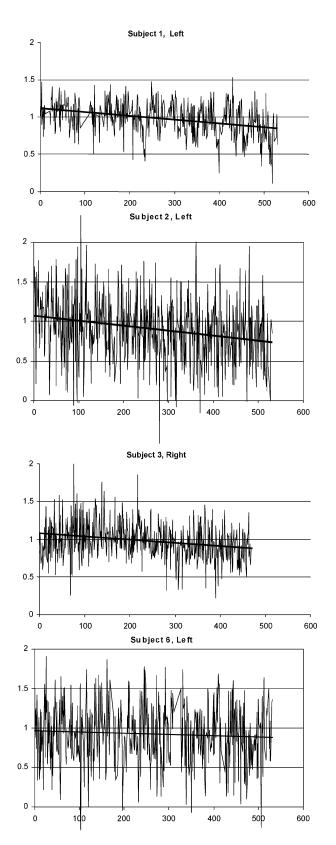


Fig. 2. Trial dependence of the SEF response (between 30 and 120 ms) is shown for four cases (which are indicated by arrows in Fig. 3). The horizontal scales run from 0 to 600 s. The vertical scales are in arbitrary units, where a constant value of 1 would indicate the ideal solution for the simple signal plus noise model, with no trial-to-trial variations. In all cases, there are relative fast variations from one trial to the next and also slower variations on scales between 50 and 200 s. Also, all these cases show a negative trend indicating that the response gets weaker over time.

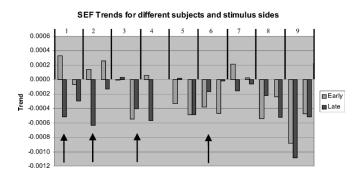


Fig. 3. Overview of the trends in the SEF responses for different data sets. The numbers on top correspond to different subjects. For each subject (except #4) both left and right hand stimulation were performed. The left hand side response trends are given first, then the right hand stimulation. The early response windows (20–35 ms) are represented first, the late responses (30–120 ms) as second collumns. The arrows correspond to the cases, shown in Fig. 2.

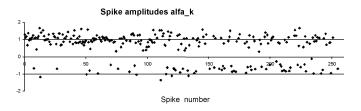


Fig. 4. Spike amplitude factors are shown for different epileptic MEG spikes. It appears that 25% of the spikes have a different polarity than the remaining 75%.

IV. DISCUSSION

This paper presents an ML estimation model to extract trialto-trial amplitude variations in brain response and to distinguish these variations from spatially and temporally correlated background noise. The application of this model to 17 SEF data sets shows that single trial SEF data show a systematic weakening of the SEF amplitudes. This negative trend in the response can be explained as habituation or nerve fatigue effects. Obrig *et al.* [2] recently demonstrated that similar effects are also present during visual stimulation, both in EEG and near infra red spectroscopy data. To extract these effects from the EEG data, the responses were averaged over 15 trials (of .33 s), and trends were computed over 12 of these subsequent averages. In all subjects, (N = 12) negative trends were found in the component from 100 to 135 ms, whereas for the earlier component (from 75 to 110 ms) no systematic negative trend was found. This finding is similar to our results for the SEF data, although the analysis methods are quite different. Pfeiderer et al. [14] have found strong habituation effects in group-averaged auditory fMRI data. Both studies [13], [14], as well as other studies on habituation contain several implicit assumptions that are avoided with the ML method presented here. The price to be paid is the relative long computation time.

The application of the model to spike data can be considered as a validation of the method, because in these data the trial-to-trial variations are so strong that they are already visible in the raw data (Fig. 1). However, the model presented in this paper can be of practical use in the analysis of multispike data, when the number of spikes is limited, when they have a low signal to noise ratio or when they appear in different polarities. Such a change in polarity without a change of pattern

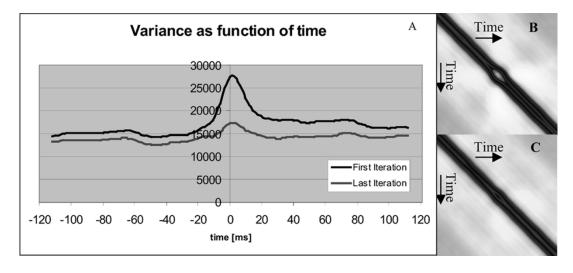


Fig. 5. This figure shows the nonstationarity in the background noise due to the (erroneous) assumption that the amplitude factors are constant, which is used to initialize the algorithm. (A) Variance which is not constant over time in the first iteration, and a much more constant variance function at the end of the iterations. This variance function is equal to the diagonal of the temporal covariance matrix T, which is presented as bitmap in (B) (first iteration) and (C) (last iteration). (B) and (C) Show that not only the main diagonal gets more constant during fitting, but also the sub-diagonals, indicating stationary background noise.

could indicate a small shift of the underlying current dipole from one gyrus to the next. This is not an uncommon situation with MEG spikes, as was shown in a recent paper [13]. In that paper, different spike types were automatically grouped by computing their mutual Euclidean signal distances and applying a clustering algorithm. When in that algorithm instead of Euclidean distances, a negative correlation measure would be used, spikes with the same spatial pattern but different polarity would show up in the same cluster, thereby increasing the number of events per cluster. Applying this alternative distance measure, in combination with the current model, would increase the signal to noise ratio, compared to the case of subgroup averaged spikes.

In [11], the same SEF data sets were used to study the stationarity of the background noise. It was concluded that, when accounting for nonstationarities caused by the baseline correction interval, the temporal covariance matrix of the background noise can be explained for 99% or more by a stationary noise model. Nevertheless, the present study shows that there are weak trial-to-trial variations in the data sets. It is apparently so that the amount of trial-to-trial variations should exceed a certain threshold in order to become visible in the temporal covariance of the background noise, as is the case with the spike data set.

The central idea behind our method is to consider the determination of brain responses as a parameter estimation problem, wherein a mathematical model is postulated to describe single trial data. Therefore, it is, at least in principle, straightforward to extend our model to physiologically more advanced models, such as suggested in e.g., [15] and [16]. However, one should also realize that when for instance simple latency jitters are included into the model, this would imply the addition of 300 to 1000 nonlinear parameters (one for each trial), which raises the question whether the likelihood function has multiple maxima. In our model, which is also a nonlinear model, the nonlinear amplitude parameters can be solved simultaneously, by solving an eigenvalue problem. Another aspect of more advanced models is that an increase of the number of parameters is accompanied with, on the one hand, an improvement of the goodness of fit, and on the other hand, a larger variance in the estimated parameters. We argue that the ML framework is very well suited to address these questions objectively. However, to perform these ideas in practice, a nonconventional amount of computer power has to be put in action. For that purpose a network of parallel computers could be used. It is our intention to proceed along these lines in the nearby future.

REFERENCES

- D. G. Wastell and D. Kleinman, "Potentiation of the habituation of human brain potentials," *Biol. Psychol.*, vol. 10, pp. 21–29, 1980.
- [2] H. Obrig, H. Isreal, M. Kohl-Bareis, K. Uludag, R. Wenzel, B. Müller, G. Arnold, and A. Villringer, "Habituation of the visually evoked potential and its vascular response: implications for neurovascular coupling in the healthy adult," *NeuroImage*, vol. 17, pp. 1–18, 2002.
- [3] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Res. Rev.*, vol. 29, pp. 169–195, 1999.
- [4] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophys.*, vol. 110, pp. 1842–1857, 1999.
- [5] W. A. Truccolo, M. M. Ding, K. H. Knuth, R. Nakamura, and S. L. Bressler, "Trial to trial variability of cortical evoked responses: implications for the analysis of functional connectivity," *Clin. Neurophys.*, vol. 113, pp. 206–226, 2002.
- [6] D. T. Pham, J. Möcks, W. Köhler, and T. Gasser, "Variable latencies of noisy signals: estimation and testing in brain potential data," *Biometrika*, vol. 74, pp. 525–533, 1987.
- [7] P. Jaškowski and R. Verleger, "Amplitudes and latencies of single trial ERP's estimated by a maximum likelihood method," *IEEE Trans. Biomed. Eng*, vol. 46, pp. 987–993, Aug. 1999.
- [8] J. C. De Munck, P. C. M. Vijn, and F. H. Lopes da Silva, "A random dipole model for spontaneous brain activity," *IEEE Trans. Biomed Eng*, vol. 39, pp. 791–804, Aug. 1992.
- [9] 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 Processing*, vol. 50, pp. 1565–1572, July 2002.
- [10] J. R. Magnus and H. Neudecker, Matrix Differential Calculus With Applications in Statistics and Econometrics. New York: Wiley, 1988.
- [11] F. Bijma, J. C. de Munck, H. Huizenga, and R. M. Heethaar, "A mathematical approach to the temporal stationarity of the background noise in MEG/EEG measurements," *NeuroImage*, vol. 20, no. 1, pp. 233–243, 2003.

- [12] J. C. De Munck, J. P. A. Verbunt, D. Van't Ent, and B. W. Van Dijk, "The use of an MEG device as 3-D digitizer and motion detection system," *Phys. Med. Biol*, vol. 46, pp. 2041–2052, 2001.
- [13] D. Van't Ent, I. Manshanden, P. Ossenblok, J. P. A. Verbunt, J. C. De Munck, D. N. Velis, and F. H. Lopes da Silva, "Spike cluster analysis in neocortical localization related epilepsy achieves clinically significant source localization results in MEG," *J. Clin. Neurophysiol.*, vol. 114, pp. 1948–1962, 2003.
- [14] B. Pfeiderer, J. Ostermann, N. Michael, and W. Heindel, "Visualization of auditory habituation by fMRI," *Neuro Image*, vol. 17, pp. 1705–1710, 2002.
- [15] S. Makeig, M. Westerfield, T. P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Dynamic brain sources of visual evoked responses," *Science*, vol. 295, pp. 690–694, 2002.
- [16] W. D. Penny, J. Stephan, J. Kiebel, J. M. Kilner, and M. D. Rugg, "Event related brain dynamics," *Trends Neurosci.*, vol. 25, pp. 387–390, 2002.



Jan Casper de Munck was born in 1962. He received the M.Sc. degree in experimental physics (*cum laude*) and the Ph.D. degree (*cum laude*) from the University of 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, and MEG). Currently, he is a Staff Member at the University Hospital Vrije Universiteit, Amsterdam, The Netherlands.



Fetsje Bijma received the M.Sc. degree in pure mathematics (*cum laude*) from the RijksUniversiteit Groningen, Groningen, The Netherlands, in 2000. She is currently working towards the Ph.D. degree at the Vrije Universitiet Medical Center, Amsterdam, The Netherlands.

Her research interests include inverse modeling of MEG/EEG data as well as statistical analysis of the data. Furthermore, she is working on the simultaneous analysis of multiple MEG/EEG data sets.



Pawel Gaura was born in 1975. He received the M.Sc. degree in technical physics from the Wroclaw University of Technology, Wroclaw, Poland. He is currently working towards the Ph.D. degree at this university.

His research interests include mathematical modeling of brain-related phenomena in MEG/EEG, habituation in the human brain, and statisical analysis of noise in MEG.



Cezary Andrzej Sieluzycki was born in 1975. He received the M.Sc. degree in technical physics from the Wroclaw University of Technology, Wroclaw, Poland, in 1999, and the Ph.D. degree in biocybernetics and biomedical engineering from Warsaw University of Technology, Warsaw, Poland, in 2003.

His research interests include mathematical modeling of brain related phenomena in MEG/EEG, learning and habituation in human brain, and brain imaging.



Maria Inês Branco was born in Lisbon, Portugal, in 1976. She received the M.Sc. degree in technological physics engineering from the Technical University of Lisbon, in 1999. She is currently working towards the Ph.D. degree on biophysics by the University of Lisbon.

Her research interests include the integration and combination of data from different functional brain imaging techniques, in particular functional magnetic resonance imaging and magnetoencephalography.



R. M. Heethaar was born in The Netherlands in 1946. He studied physics and mathematics from 1963–1968 at the Utrecht University, Utrecht, The Netherlands, He completed the Ph.D. degree study on atrio-ventricular conduction in the rat heart in 1972.

He held joint appointments with the departments of Cardiology and Medical Physics at the Utrecht University from 1969–1985. and worked for the two years following his Ph.D. degree on artificial electrical stimulation of the human heart. From

1974–1976, he was a research fellow at the MAYO Clinic, USA in the fields of cardiac mechanics and computer tomography. In 1979, he became Professor of bioengineering at the University of Technology Twente, Twente, The Netherlands, and in 1987 Professor of Medical Physics at the Utrecht University. He founded two Dutch companies in bioengineering. Since 1993, he is Professor of Medical Physics at the Vrije Universiteit, Amsterdam, and head of the department. His main interests are physics of heart and circulation and the brain. He is author of numerous papers on electrophysiology, cardiac mechanics, bioengineering, and hemorheology. He holds 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 and circulation noninvasively.

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