

# Distinguishing Between Moving and Stationary Sources Using EEG/MEG Measurements With an Application to Epilepsy

İmam Şamil Yetik\*, Arye Nehorai, *Fellow, IEEE*, Jeffrey David Lewine, and Carlos H. Muravchik, *Senior Member, IEEE*

**Abstract**—Performances of electroencephalography (EEG) and magnetoencephalography (MEG) source estimation methods depend on the validity of the assumed model. In many cases, the model structure is related to physical information. We discuss a number of statistical selection methods to distinguish between two possible models using least-squares estimation and assuming a spherical head model. The first model has a single moving source whereas the second has two stationary sources; these may result in similar EEG/MEG measurements. The need to decide between such models occurs for example in Jacksonian seizures (e.g., epilepsy) or in intralobular activities, where a model with either two stationary dipole sources or a single moving dipole source may be possible. We also show that all of the selection methods discussed choose the correct model with probability one when the number of trials goes to infinity. Finally we present numerical examples and compare the performances of the methods by varying parameters such as the signal-to-noise ratio, source depth, and separation of sources, and also apply the methods to real MEG data for epilepsy.

**Index Terms**—Electroencephalography, epilepsy, magnetoencephalography, model selection.

## I. INTRODUCTION

LOCATING electrical sources in the brain has broad applications, ranging from clinical (for instance finding the location of the abnormality before a surgery in epilepsy) to neuroscientific (such as locating the parts of the brain that control certain physical activities).

In neuroscience, typically a stimulus is applied and EEG/MEG measurements are recorded during the course of activation. In clinical applications, measurements of spontaneous activities are made. Then, this data is used to infer some properties of the source. For some applications a model with a single dipole source is shown to be sufficiently good [1], whereas for others models with multiple sources [2], moving

sources [3], or distributed sources [4] are needed. Often a model is selected using past experience or the physics of the problem and the data is analyzed based on the assumed model. However, choosing a model using the measured data is the most direct and effective approach. In this case, past experience or physics of the problem is used only to decide on the possible models to select from.

There has been an extensive research on selecting the number of stationary dipole sources [5]–[8]. Selecting the number of sources is a nested model selection problem, i.e., the candidate models can be hierarchically classified from complex to simple, with each one being a special case of its predecessor. However, in certain problems there is a need to distinguish between sources of different types, making the problem nonnested. Some model selection methods that are suitable for nested problems cannot be applied in this case. For instance for a nested problem we can construct a hypothesis test to find out whether an extra parameter added has the value zero.

In this paper, we focus on distinguishing between two nonnested models, one with two stationary sources and an alternative with a single moving source. These two models may result in EEG/MEG measurements that are difficult to distinguish in certain cases (e.g., the magnetic fields and electric potentials induced by two dipoles when the strength of one of the dipoles is increasing while the other is decreasing may be similar to those of a single moving dipole.) For example, in Jacksonian seizures [9], there is a wave of epileptiform activity that marches along the cortex of the post-central gyrus. In contrast, when an epileptic spike is seen first over one temporal lobe and then 10–20 milliseconds later over the other temporal lobe, it is likely that this reflects transcallosal propagation of activity between the hemispheres, and not a spread of activation in the aforementioned sense. Also, in the case of intralobular activities both marching dipole and propagation models may be valid. For instance, the M100 response elicited by auditory stimuli appears to have early and late components. Some have modeled this as being generated by two discrete sources with temporally delayed activation profiles [10], while others have suggested that activity reflects a wave of cortical excitation spreading across the temporal lobe [11]. We present a number of statistical methods to decide on one of the models mentioned, analyze their limiting behavior, and test them using numerical examples. We employ a spherical head model to simplify the solution, however the same model selection methods can

Manuscript received August 30, 2003; revised August 23, 2004. This work was supported in part by the National Science Foundation under Grant CCR-0105334. Asterisk indicates corresponding author.

\*İ. Ş. Yetik is with the Department of Electrical and Computer Engineering (ECE), University of Illinois at Chicago, 851 S. Morgan, Room 1020 SEO, Chicago, IL 60607 USA (e-mail: syetik@ece.uic.edu).

A. Nehorai is with the Department of Electrical and Computer Engineering (ECE), University of Illinois at Chicago, Chicago, IL 60607 USA.

J. D. Lewine is with the University of Kansas Medical Center, Kansas City, KS 66160 USA.

C. H. Muravchik is with the Universidad Nacional de La Plata, 1900 La Plata, Argentina.

Digital Object Identifier 10.1109/TBME.2004.843289

be applied to a realistic head model [12] for improving the performance at the price of increased computations.

In Section II, we present the two possible data models. We give a brief overview of the parameter estimation for these models in Section III. Section IV is devoted to a number of model selection methods with comments on their advantages and disadvantages. In Section V, we show that all the methods that we use in this paper select the correct model with probability one when number of trials goes to infinity and one of the candidate models is actually correct. We present numerical examples in Section VI to compare the methods and evaluate their performances for different values of signal-to-noise ratio (SNR), source depths, and separation of the sources, and apply the methods to real MEG data for epilepsy in Section VII. A conclusion and a summary of possible directions of future research are given in Section VIII.

## II. MEASUREMENT AND SOURCE MODELS

Consider a system of  $m_B$  MEG sensors and  $m_E$  EEG sensors to measure the magnetic fields and electric potentials induced by  $l$  current dipoles inside the brain responding to a stimulus. The measurements at the  $m = m_B + m_E$  sensors can be modeled as

$$\mathbf{y} = A(\boldsymbol{\theta})\mathbf{s} + \mathbf{n}, \quad (2.1)$$

where  $\boldsymbol{\theta}$  is the vector of dipole location parameters,  $\mathbf{n}$  is an  $m \times 1$  vector representing additive noise (assumed to be zero-mean normally distributed, independent of the source, and uncorrelated in space and time with a variance  $\sigma^2$ ),  $\mathbf{s}$  is the  $3l \times 1$  vector of dipole moments, and  $A(\boldsymbol{\theta})$  the  $m \times 3l$  array response matrix. This matrix has the form  $A(\boldsymbol{\theta}) = [A_B^T(\boldsymbol{\theta}), A_E^T(\boldsymbol{\theta})]^T$ , where  $A_B(\boldsymbol{\theta})$  represents the MEG and  $A_E(\boldsymbol{\theta})$  the EEG response.

The matrix  $A(\boldsymbol{\theta})$  can be computed using the quasistatic approximation of the Maxwell's equations [13], [14]. Note that for the spherical head model, the third column of  $A_B(\boldsymbol{\theta})$  is zero since the radial components of the dipoles inside the head do not produce magnetic fields outside the head.

It is convenient to use spherical coordinates for the dipole moments since we utilize a spherical head model. In this case, the moment of a single dipole  $\mathbf{s}$  can be expressed as

$$\mathbf{s} = s_\nu \mathbf{u}_\nu + s_\phi \mathbf{u}_\phi + s_p \mathbf{u}_p \quad (2.2)$$

where

$$\begin{aligned} \mathbf{u}_\nu &= [\cos \nu \cos \varphi, \cos \nu \sin \varphi, -\sin \nu]^T \\ \mathbf{u}_\phi &= [-\sin \varphi, \cos \varphi, 0]^T \\ \mathbf{u}_p &= \mathbf{p}/p \end{aligned}$$

are orthonormal vectors,  $\mathbf{p}$  is the position vector of the dipole relative to the center of the sphere,  $p$  its magnitude (distance from the center),  $\nu$  the elevation angle of the dipole position, and  $\varphi$  the azimuth. The position of a single dipole is fully given by the three parameters,  $\nu$ ,  $\varphi$ , and  $p$ . The above model can be extended to multiple trials (useful to reduce the effect of noise) and to multiple temporal samples for tracking moving sources. We define the spatio-temporal data with  $K$  trials, and  $N$  time samples

$$\begin{aligned} \mathbf{y}_k(t) &= A(\boldsymbol{\theta}(t))\mathbf{s}(t) + \mathbf{n}_k(t), \\ k &= 1, \dots, K, \quad t = 1, \dots, N, \end{aligned} \quad (2.3)$$

This model can be used for either a single moving source or two stationary sources with appropriate choices of  $\boldsymbol{\theta}(t)$ 's and  $\mathbf{s}(t)$ 's.

For two stationary dipoles we have

$$\begin{aligned} \boldsymbol{\theta} &= [\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T]^T \\ &= [\nu_1, \varphi_1, p_1, \nu_2, \varphi_2, p_2]^T \\ A(\boldsymbol{\theta}) &= [A(\boldsymbol{\theta}_1), A(\boldsymbol{\theta}_2)], \\ \mathbf{s}(t) &= [\mathbf{s}_1^T(t), \mathbf{s}_2^T(t)]^T \end{aligned}$$

where the subscripts 1 and 2 correspond to each of the two dipoles. Observe that we have omitted the dependence of  $\boldsymbol{\theta}$  on  $t$ , since the locations of the dipoles,  $\boldsymbol{\theta}_1$  and  $\boldsymbol{\theta}_2$ , are assumed to be fixed, whereas their strengths and directions are allowed to change in time. For a single moving source the dipole location varies in time

$$\boldsymbol{\theta}(t) = [\nu(t), \varphi(t), p(t)]^T. \quad (2.4)$$

In this case it is possible to estimate the dipole location at each time sample independently. This would provide maximum freedom for the dipole location variations in time. However, the resulting number of unknown parameters and hence computational cost of the estimation would be quite high. To decrease the number of estimated parameters we employ basis functions to form the path of the moving dipole [3]

$$\boldsymbol{\theta}(t) = X\phi(t) \quad (2.5)$$

where  $\phi(t)$  is a  $d \times 1$  vector whose entries are known temporal basis functions and  $X$  a matrix of unknown coefficients of dimension  $3 \times d$ . In this case, we have to estimate only  $3d$  parameters which are the coefficients of the basis functions (entries of  $X$ ). Unlike  $\boldsymbol{\theta}(t)$ , the dimension  $3d$  of  $X$  is fixed and not increasing with time. Note that the number of basis functions should be chosen as a trade off between accuracy and computational complexity. Avoiding basis functions and estimating point by point is less accurate compared with using good basis functions based on prior info. Incorrect or insufficient number of basis functions would not represent the data well, and the resulting estimates would be biased. However, considering the limited spatial resolution of EEG/MEG, it is not possible to estimate very detailed paths (even if they exist); hence, it is sufficient to use a few basis functions even though the exact path cannot be represented using these, and the resulting bias is negligible.

## III. PARAMETER ESTIMATION

We briefly overview the ordinary least-squares method (OLS) we use to estimate the unknown location parameters  $\boldsymbol{\theta}(t)$  and dipole moments  $\mathbf{s}(t)$ , see also [15]. OLS estimate of  $\boldsymbol{\theta}$  is

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \sum_{t=1}^N -\bar{\mathbf{y}}(t)^T P_s(\boldsymbol{\theta}) \bar{\mathbf{y}}(t) \quad (3.6)$$

where  $\bar{\mathbf{y}}(t) = 1/K \sum_{k=1}^K \mathbf{y}_k(t)$ , and  $P_s(\boldsymbol{\theta}) = A(\boldsymbol{\theta})[A(\boldsymbol{\theta})^T A(\boldsymbol{\theta})]^{-1} A(\boldsymbol{\theta})^T$ . Here, the subscript "s" stands for stationary sources. We then estimate  $\mathbf{s}(t)$  using

$$\hat{\mathbf{s}}(t) = [A(\hat{\boldsymbol{\theta}})^T A(\hat{\boldsymbol{\theta}})]^{-1} A(\hat{\boldsymbol{\theta}})^T \bar{\mathbf{y}}(t) \quad (3.7)$$

assuming sufficient number of sensors so that  $A(\hat{\boldsymbol{\theta}})^T A(\hat{\boldsymbol{\theta}})$  is nonsingular.

For a single moving source, the OLS estimate of  $X$  is

$$\hat{X} = \arg \min_X \sum_{t=1}^N -\bar{\mathbf{y}}(t)^T P_m(X) \bar{\mathbf{y}}(t) \quad (3.8)$$

where

$P_m(X) = A(X\phi(t))[A(X\phi(t))^T A(X\phi(t))]^{-1} A(X\phi(t))^T$  and subscript “m” stands for moving source. We then estimate  $\mathbf{s}(t)$  using

$$\hat{\mathbf{s}}(t) = [A(\hat{X}\phi(t))^T A(\hat{X}\phi(t))]^{-1} A(\hat{X}\phi(t))^T \bar{\mathbf{y}}(t). \quad (3.9)$$

It is possible to generalize the estimation to the generalized least-squares method (GLS) [15] for correlated noise with known covariance  $\Sigma = s^2 W$ , where  $s$  reduces to the estimate of the noise variance for the uncorrelated case. A whitening filter  $W^{-1/2}$  (obtained from the Cholesky decomposition [17]) can be used to pre-whiten the data. In this case,  $\mathbf{y}(t)$  will be replaced by  $W^{-1/2} \mathbf{y}(t)$ , and  $A(\boldsymbol{\theta}) \mathbf{s}(t)$  will be replaced by  $W^{-1/2} A(\boldsymbol{\theta}) \mathbf{s}(t)$ , resulting in the unexplained data  $W^{-1/2} \mathbf{y}(t) - W^{-1/2} A(\boldsymbol{\theta}) \mathbf{s}(t)$  to be used instead of  $\mathbf{y}(t) - A(\boldsymbol{\theta}) \mathbf{s}(t)$ . It is also possible to extend to unknown covariance as in [16], [18], [19].

It can be seen that the cost functions (3.6) and (3.8) are equivalent to the negative likelihood function for a normally distributed and uncorrelated noise. Hence, OLS and maximum likelihood (ML) estimation are equivalent in this case. There exist alternative methods to calculate these estimates such as MUSIC [15], [20] however we will not pursue them here since our primary concern is the model selection.

#### IV. MODEL SELECTION

We describe the model selection methods we use to distinguish between the moving dipole and two stationary dipole sources. For each method we first explain the criterion used and then discuss its advantages and disadvantages. We calculate the value of the criterion under both models and select the one with larger or smaller value depending on the method.

##### A. Residual Variance

This method simply measures the error between the real data and its estimate. We define residual variance as follows [21]:

$$r^2(V) = \frac{\sum_{t=1}^N [\bar{\mathbf{y}}(t) - A(\hat{\boldsymbol{\theta}}(t)) \hat{\mathbf{s}}(t)]^T V^{-1} [\bar{\mathbf{y}}(t) - A(\hat{\boldsymbol{\theta}}(t)) \hat{\mathbf{s}}(t)]}{\sum_{t=1}^N \bar{\mathbf{y}}(t)^T V^{-1} \bar{\mathbf{y}}(t)} \quad (4.10)$$

where  $V$  is a positive definite matrix, which we choose to be the identity matrix in this paper (another  $V$  could be used to assign different penalties to the errors at different sensors.) The correct model and a perfect fit of the parameters together would give a residual variance equal to zero. We calculate the residual variance for the two alternative models and choose the one that makes the residual variance closer to zero.

The residual variance is doubly noncentral F-distributed, since the denominator is noncentral  $\chi^2$  distributed. The number of degrees of freedom (DOFs) depends on the value of  $V$ . The residual variance may seem intuitively appealing, but it has a number of drawbacks. First, it has a low value for large noise variance even if the model is correct, although averaging the

data may mitigate this effect. Secondly, it cannot use any prior information we may have regarding the noise. Also, for large data power it gets closer to zero, indicating a better fit. It has a low value when the performance of the parameter estimation is poor even if the model is correct. These factors often give misleading results when the residual variance method is used for the model selection. However, we have included it for comparison purposes and its common use in the literature.

##### B. Smirnov Test

In this method [22], the selection is based on finding the model that makes the distribution of the “unexplained data” closer to the normal distribution. The unexplained data can be written as a spatio-temporal matrix

$$E_k = Y_k - [A(\hat{\boldsymbol{\theta}}(1)) \hat{\mathbf{s}}(1) \dots A(\hat{\boldsymbol{\theta}}(N)) \hat{\mathbf{s}}(N)] \quad (4.11)$$

where

$$Y_k = [\mathbf{y}_k(1) \dots \mathbf{y}_k(N)]. \quad (4.12)$$

For the correct model  $E_k$  consists of the noise and the estimation error, whereas for an incorrect model we also have a modeling error. We calculate  $E_k$  under both models, and choose the one that is closer to a normal distribution. The Smirnov test is used at this point by providing a criterion to test the normality of the unexplained data, i.e., the elements of  $E_k$ . The criterion suggested by Smirnov is

$$\epsilon_S = \sum_{k=1}^K \sum_{t=1}^N \sum_{i=1}^m [S(e_{ik}(t)) - F(e_{ik}(t))]^2 f(e_{ik}(t)) \quad (4.13)$$

where  $e_{ik}(t)$  is a single sample of the unexplained data, i.e., an element of  $E_k$ ,  $S(e_{ik}(t))$  is the empirical cumulative distribution function (CDF) of the noise constructed using these samples,  $F(e_{ik}(t))$  the true CDF, and  $f(e_{ik}(t))$  the true probability density function, i.e., the normal probability density function. We calculate  $S(e_{ik}(t))$  by counting the number of noise samples belonging to a certain interval, and calculate  $F(e_{ik}(t))$  considering the normal distribution in this interval. This process, enabling us to handle continuous distributions, is called binning, hence the Smirnov test is said to be a member of the binned tests.

The measurements at all the sensors and time instants, i.e., all the elements of  $E_1, E_2, \dots, E_K$  can be used separately as the samples to construct  $S(e_{ik}(t))$  and calculate  $\epsilon_S$ , since we assume uncorrelated sensor noise in space and time.

This criterion simply measures weighted squared difference between the assumed true CDF and the sample CDF of the noise obtained using the measurements. It is zero for the correct model and large sample size, and large for an incorrect model. We calculate  $\epsilon_S$  under the two alternative models and choose the one with smaller  $\epsilon_S$ .

The Smirnov test drastically suffers from small number of samples, which occurs when we have small numbers of trials, or time samples and sensors.

##### C. Chi-Square Method

In this method [23], we compare the unexplained data with the pure noise. If the energy of the unexplained data is much larger than the noise variance, we can argue that the model is not

correct. We use the data from each trial separately to construct the test statistic since the trials are independent

$$\chi^2 = \frac{\sum_{t=1}^N [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)] [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)]^T}{\sigma^2} \quad (4.14)$$

where  $\mathbf{e}_i(t)$  is the unexplained data vector of dimension  $m \times 1$  obtained from the  $i$ th trial. This test is also related to measuring the normality of the unexplained data. When the model is correct the unexplained data coincides with the pure noise only and hence the statistic given by (4.14) is  $\chi^2$  distributed. When the model is not correct, the unexplained data is not normally distributed, making the statistic distributed differently from  $\chi^2$ .

Assuming the noise is white and normal with known variance, this statistic is  $\chi^2$  distributed with  $KNm - q$  DOFs, where  $q$  is the number of parameters to be estimated, since we are adding  $KNm$  independent normally distributed random variables squared with  $q$  parameters estimated. The independence of the terms can be seen by recalling that, (i) the trials are independent, (ii) the unknown location parameters are assumed to be deterministic, (iii) the time samples are independent since  $\mathbf{e}_k(t_i)$  and  $\mathbf{e}_k(t_j)$  are normally distributed and  $E\{\mathbf{e}_k(t_i)\mathbf{e}_k^T(t_j)\} = E\{\mathbf{e}_k(t_i)\}E\{\mathbf{e}_k^T(t_j)\}$  for all  $i, j$ . The term  $E\{\mathbf{e}_k(t_i)\mathbf{e}_k^T(t_j)\}$  is not equal to  $\delta_{ij}$ , since  $\mathbf{e}_k(t_i)$  is the unexplained data (not pure noise), and it will have components other than noise for a wrong model. We compute this statistic for both models and select the one that is less significant, since the model with less significance is more successful in explaining the total error by the noise error.

The  $\chi^2$  test has the advantage of incorporating noise information, i.e., the variance, however it must be known *a priori*.

#### D. Lack of Fit

This method [24] is a modification of the  $\chi^2$  test for the case of unknown noise variance. We replace  $\sigma^2$  by its unbiased estimator ( $\hat{\sigma}^2 = \sum_{k=1}^K \sum_{t=1}^N (\mathbf{y}_k(t) - \bar{\mathbf{y}}(t))^T (\mathbf{y}_k(t) - \bar{\mathbf{y}}(t)) / ((mKN)(K - 1))$ ). The lack of fit (LOF) statistic is then

$$\text{LOF} = \frac{\sum_{t=1}^N [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)] [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)]^T}{(KNM - q)\hat{\sigma}^2} \quad (4.15)$$

We noted previously that  $\sum_{t=1}^N [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)] [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)]^T / \sigma^2$  is  $\chi^2$  distributed with  $KNm - q$  DOFs. Similarly,  $\hat{\sigma}^2$  is  $\chi^2$  distributed since it is the sum of normal distributed random variables squared. To be exact,  $mNK(K - 1)\hat{\sigma}^2/\sigma^2$  is  $\chi^2$  distributed with  $mNK(K - 1)$  DOFs. So the statistic in (4.15) is the ratio between two  $\chi^2$  random variables divided by their DOFs. Hence, it is  $F$  distributed with  $KNm - q$  and  $mNK(K - 1)$  DOFs.

For the correct model and true parameter values, the unexplained data becomes normally distributed and therefore the statistic in (4.15) is  $F$  distributed. If the model is not correct, this statistic is not  $F$  distributed. Similarly to the  $\chi^2$  test, we compute the statistics for both models and select the model that is less significant.

The LOF test incorporates the effect of noise and furthermore it has the advantage that it does not require prior knowledge of the noise statistics.

#### E. Akaike's Information Criterion

Akaike's information criterion (AIC) [25] measures the goodness of a model using the log-likelihood function. AIC also has a penalty term for the number of parameters, therefore it favors simpler models and is defined as

$$\text{AIC} = -2\log(p_\theta(\mathbf{y})) + 2q \quad (4.16)$$

where  $p_\theta(\mathbf{y})$  is the data likelihood function, and  $q$  is the number of parameters to be estimated. Substituting the normal pdf for the likelihood function

$$\text{AIC} = \log((2\pi)^{mNK} \sigma^2) + \frac{\sum_{t=1}^N [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)] [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)]^T}{\sigma^2} + 2q. \quad (4.17)$$

We calculate the AIC values under the two alternative models, and select the smaller one. We can generalize AIC to the unknown noise variance case by substituting  $\sigma^2$  by its unbiased estimate  $\hat{\sigma}^2$ .

#### V. LIMITING ANALYSIS

We show that all the above model selection methods select the correct model with probability one when the number of trials  $K$  goes to infinity and one of the candidate models is correct. First we introduce some notation to simplify the presentation.

- $A_0(t)$  is the true  $A(\boldsymbol{\theta}(t))$  which is unknown but assumed to be equal to one of the possibilities considered.
- $\mathbf{s}_0(t)$  is the true dipole moment vector  $\mathbf{s}(t)$ .
- $\hat{A}(t)$  is the estimated  $A(\boldsymbol{\theta}(t))$  after a model is selected.
- $\hat{\mathbf{s}}(t)$  is the estimated dipole moment vector after a model is selected.

Observe that  $\mathbf{e}_k(t)$  plays a crucial role in all the model selection methods, since it is the unexplained data. With the above notation

$$\mathbf{e}_k(t) = \mathbf{y}_k(t) - \hat{A}(t)\hat{\mathbf{s}}(t) = A_0(t)\mathbf{s}_0(t) - \hat{A}(t)\hat{\mathbf{s}}(t) + \mathbf{n}_k(t) \quad k = 1, \dots, K \quad (5.18)$$

since  $\mathbf{y}_k(t) = A_0(t)\mathbf{s}_0(t) + \mathbf{n}_k(t)$ . Note that the OLS estimates are consistent, i.e., converge to the true parameter values when  $K \rightarrow \infty$ . This can be seen by writing the OLS cost function as

$$\sum_{t=1}^N \sum_{k=1}^K \left\{ \mathbf{y}_k^T(t) \mathbf{y}_k(t) - (\hat{A}(t)\hat{\mathbf{s}}(t))^T \mathbf{y}_k(t) - \mathbf{y}_k^T(t) (\hat{A}(t)\hat{\mathbf{s}}(t)) + (\hat{A}(t)\hat{\mathbf{s}}(t))^T (\hat{A}(t)\hat{\mathbf{s}}(t)) \right\} \quad (5.19)$$

Substituting  $\mathbf{y}_k(t) = A_0(t)\mathbf{s}_0(t) + \mathbf{n}_k(t)$ , we have

$$\sum_{t=1}^N \sum_{k=1}^K \left\{ (A_0(t)\mathbf{s}_0(t) - \hat{A}(t)\hat{\mathbf{s}}(t))^T (A_0(t)\mathbf{s}_0(t) - \hat{A}(t)\hat{\mathbf{s}}(t)) + (A_0(t)\mathbf{s}_0(t) - \hat{A}(t)\hat{\mathbf{s}}(t))^T \mathbf{n}_k + \mathbf{n}_k^T (A_0(t)\mathbf{s}_0(t) - \hat{A}(t)\hat{\mathbf{s}}(t)) + \mathbf{n}_k^T(t) \mathbf{n}_k(t) \right\} \quad (5.20)$$

and letting  $K \rightarrow \infty$  we obtain

$$NK\sigma^2 + \lim_{K \rightarrow \infty} \sum_{t=1}^N \sum_{k=1}^K \left\{ (A_0(t)\mathbf{s}_0(t) - \hat{A}(t)\hat{\mathbf{s}}(t))^T \times (A_0(t)\mathbf{s}_0(t) - \hat{A}(t)\hat{\mathbf{s}}(t)) \right\}. \quad (5.21)$$

The OLS cost function is minimized for the true parameter values and correct model. This can be seen by realizing that the summation term is always nonnegative and zero only for the true parameter values, and also considering that one of the candidate models is correct which makes it possible to obtain an  $\hat{A}(t)$  which is exactly equal to  $A_0(t)$ . In this case, the OLS cost is simply  $NK\sigma^2$ . As for  $\mathbf{s}(t)$ , its estimate is then equal to the true value of  $\mathbf{s}(t)$ :

$$\begin{aligned}\hat{\mathbf{s}}(t) &= [A_0(t)^T A_0(t)]^{-1} A_0(t)^T \bar{\mathbf{y}}(t) \\ &= [A_0(t)^T A_0(t)]^{-1} A_0(t)^T A_0(t) \mathbf{s}_0(t) \\ &= \mathbf{s}(t)\end{aligned}$$

since  $\bar{\mathbf{y}}(t) = A_0(t) \mathbf{s}_0(t)$  for  $K \rightarrow \infty$ .

In the following, we analyze each of the model selection methods and show that all of them choose the correct model with probability one for  $K \rightarrow \infty$ , when one of the candidate models is actually correct.

#### A. Residual Variance

We have

$$\bar{\mathbf{y}}_k(t) = \frac{1}{K} \sum_{k=1}^K \mathbf{y}_k(t) = \frac{K}{K} A_0(t) \mathbf{s}_0(t) + \frac{1}{K} \sum_{k=1}^K \mathbf{n}_k(t). \quad (5.22)$$

For large values of  $K$ , that last term vanishes since the noise is zero mean, therefore

$$\lim_{K \rightarrow \infty} \bar{\mathbf{y}}_k = A_0(t) \mathbf{s}_0(t). \quad (5.23)$$

With the above limit of  $\bar{\mathbf{y}}_k(t)$  the residual variance becomes zero if and only if the model is correct assuming that  $|\mathbf{s}_0^T(t) A_0^T(t) A_0(t) \mathbf{s}_0(t)|$  is bounded. The correct model makes the term  $\bar{\mathbf{y}}(t) - \hat{A}(\hat{\boldsymbol{\theta}}(t)) \hat{\mathbf{s}}(t)$  zero, and an incorrect model which makes  $\hat{A}(\hat{\boldsymbol{\theta}}(t)) \hat{\mathbf{s}}(t)$  different from  $A_0(t) \mathbf{s}_0(t)$  for arbitrary dipole locations and moments results in a residual variance larger than zero. Therefore, the residual variance becomes zero if and only if the model is correct, guaranteeing the selection of the correct model.

#### B. Smirnov Test

For the correct model,  $E_k$  in (4.11) becomes zero-mean normal since the OLS method gives the true parameter values for  $K \rightarrow \infty$ . Also, we have infinitely many samples of the normal distribution when  $K \rightarrow \infty$ . Hence  $S(e)$  becomes equal to  $F(e)$ , and  $\epsilon_S$  becomes zero for the correct model, when  $K \rightarrow \infty$ . For an incorrect model, when  $\hat{A}(t) \hat{\mathbf{s}}(t)$  fails to be equal to  $A_0(t) \mathbf{s}_0(t)$  for arbitrary dipole locations and moments,  $E_k$  will not be zero mean, which will result in a positive  $\epsilon_S$ . Therefore, the Smirnov test chooses the correct model with probability one when  $K \rightarrow \infty$ .

#### C. Chi-Square and Lack of Fit

Consider  $\sum_{t=1}^N [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)] [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)]^T$ , a common factor in the test statistics of the  $\chi^2$  test and LOF. We can see that this is equal to the OLS cost function, hence the arguments made at the beginning of this section are valid. When the DOFs are the same, the correct model will always be less significant and hence both methods will select the correct model. For both the  $F$  and  $\chi^2$  distributions, there is a factor  $K$

in the DOFs as a multiplying term. Hence for any model with arbitrary number of parameters to be estimated we will have the same DOFs when  $K \rightarrow \infty$ . Therefore, both methods select the correct model with probability one when  $K \rightarrow \infty$ .

#### D. Akaike's Information Criterion

When  $K \rightarrow \infty$ , the first and second term of (4.17) dominate and the penalty term vanishes. The first term is independent of the model, hence we are left with  $\sum_{t=1}^N [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)] [\mathbf{e}_1^T(t), \dots, \mathbf{e}_K^T(t)]^T$ , and the previous arguments regarding this term are valid, i.e., this term will be minimum if and only if the model is correct. Therefore AIC chooses the correct model with probability one when  $K \rightarrow \infty$ .

## VI. NUMERICAL EXAMPLES

We test the methods using computer simulations and analyze their performance by varying a number of factors. In our simulations, the parameter estimation and model selection are performed based on 40 trials for the evoked responses and we have repeated these experiments 50 times to be able to compare the performances of the methods using the percentage of correct decisions. We have used  $m_B = 37$  MEG and  $m_E = 37$  EEG sensors which are positioned in rings of 1, 6, 12, 18 sensors each separated by 12 degrees. These positions are chosen to approximate a realistic system following [26], [27]. The MEG sensors are radially oriented at 11 cm's and EEG sensors at 9 cm's from the center of the head with a radius of 9 cm's. For EEG we have used a three-layer (scalp, brain, and skull) head model, with conductivity values 0.33, 0.33, and 0.0042 ( $\Omega\text{m}$ )<sup>-1</sup>; and radii 8.8, 8.1, and 8.5 cm's respectively [28]. We have calculated the SNR using the average power of the signals at all sensors

$$\text{SNR} = \frac{1}{K} \sum_{k=1}^K \frac{\text{Tr}[Y_k Y_k^T] / m}{\sigma^2} \quad (6.24)$$

where "Tr" represents the trace operation.

We have used dipoles tangential to the sphere since the radial components are silent for MEG. We have created two dipole locations in depths of 1 to 4 cm's with steps of 0.5 cm's. The two dipole locations are assumed to be in the same depth at angles  $\varphi = 20, 180$  deg.,  $\varphi = 30, 170$  deg.,  $\varphi = 40, 160$  deg.,  $\varphi = 50, 150$  deg.,  $\varphi = 60, 140$  deg.,  $\varphi = 70, 130$  deg.,  $\varphi = 80, 120$  deg.,  $\varphi = 90, 110$  deg. and a fixed  $\nu = 30$  degrees. For the strength of a single moving dipole, we have assumed a half-sine variation at a frequency of 20 Hz. For the strengths of the two stationary dipoles we have used a half-sine for one of them and a half-cosine time variation for the other one, so that the strength of one dipole is increasing while the other one is decreasing.

For the model with a single moving dipole, the source is assumed to travel linearly between the two locations with a speed of 0.4 cm/ms. We have used 100 time samples for both models. These values are chosen to mimic a real case, and the sensor signals under the two models look close to each other. For the model described above we have estimated the locations of the stationary dipoles, the path of the moving dipole, and the dipole moments using the OLS method as described in Section III. For the moving dipole model, three basis functions are used:

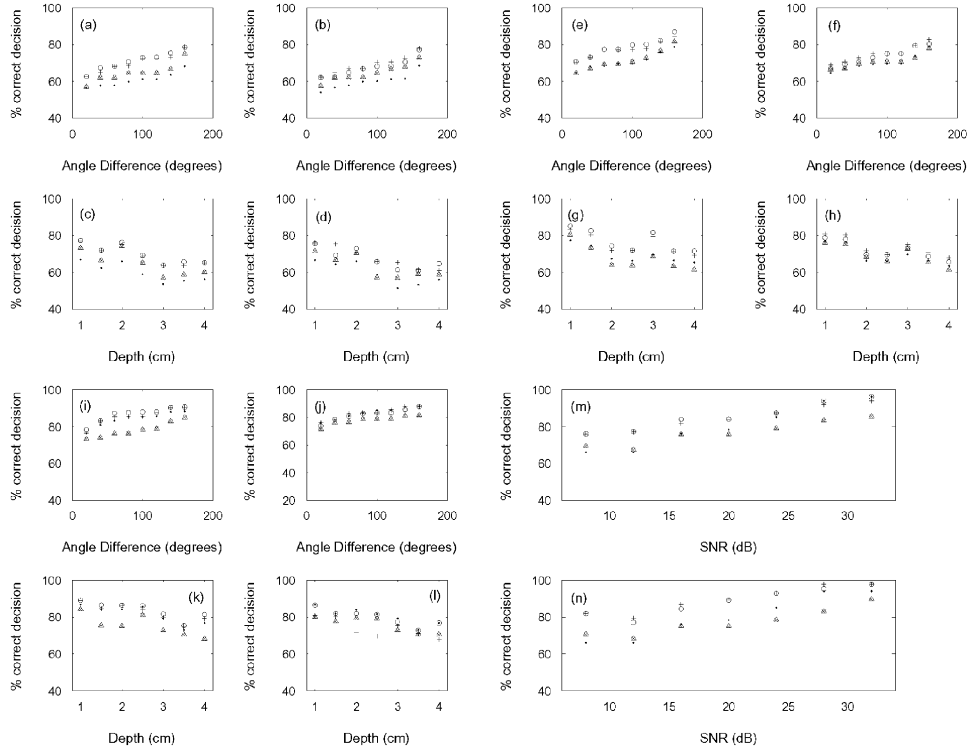


Fig. 1. (a)–(l) Comparisons of the percentages of correct decision for all of the methods. The symbol “·” is used for residual variance, “o” for the Smirnov test, “x” for  $\chi^2$ , “Δ” for LOF, and “+” for AIC. (a) Shows the results when the correct model is with two stationary dipoles and (b) with a single moving dipole, both as a function of source angle separation; (c) and (d) give similar plots but as a function of depth this time, all for SNR = 8. (e)–(h) Same as (a)–(d) except for SNR = 16. (i)–(l) Same as (a)–(d) except for SNR = 24. (m), (n) Comparisons of the percentages of correct decision for all the methods as a function of SNR. (m) Shows the results when the correct model is two stationary dipoles and (n) the correct model is a moving dipole.

$\phi_1(t) = 1, \phi_2(t) = t, \phi_3(t) = t^2$ . The same basis functions are used for all three components,  $\varphi, \nu$ , and  $p$  so that  $\phi(t) = [\phi_1(t), \phi_2(t), \phi_3(t)]^T$ , and  $X$  is a  $3 \times 3$  matrix of unknown coefficients.

Fig. 1(a) and (b) shows the results when the model with stationary dipoles and a single moving dipole is correct respectively, for SNR = 8. The percentages of correct decisions are plotted as a function of the separation angle of the two source locations, after averaging over results for different depths. Fig. 1(c) and (d) shows similar plots, but this time as a function of the depth of the locations averaging over data for different angle separations. Similar plots are given for SNR = 16 in Fig. 1(e)–(h), and for SNR = 24 in Fig. 1(i)–(l). In Fig. 1(m)–(n), we plot the percentages of correct decisions as a function of SNR after averaging over data for different source angle separations and depths. Fig. 1(m) is for the case when the correct model is two stationary dipoles, and Fig. 1(n) is for the true single moving dipole case.

Additionally, assuming 80% correct decision is satisfactory, we present the cutoff SNR values that achieves this performance for all methods in Table I.

Observe that all of the methods perform well at high SNR values. We could expect the models accounting for the noise variance to perform well also in low SNR, but this is not the case in practice since low SNR adversely affects the estimation of the parameters. Hence there will be unexplained data due to the estimation error in addition to the error stemming from pure noise or selecting an incorrect model. It is also observed that the methods perform better for well separated sources. This is

TABLE I  
CUT-OFF SNR VALUES FOR 80% CORRECT DECISIONS ON THE AVERAGE

	Res. Var.	Smirnov Test	$\chi^2$	LOF	AIC
Stat.	20dB	16dB	20dB	20dB	14dB
Mov.	20dB	17dB	20dB	20dB	17dB

also expected and related to the spatial performance of the estimation methods. Deeper sources result in poor performance of the methods, hence misfitting of the data plays a more important role.

In general, the residual variance performs poorly compared with other methods as explained in Section IV. The Smirnov and AIC tests perform better than the other three methods. Also the  $\chi^2$  and LOF tests give very similar results, since the only difference is the estimation of the noise variance. For closely separated locations, the Smirnov test seems to perform slightly better than others.

Observe that the methods are more successful when the correct model is stationary. This can be explained by the path estimation which makes the parameter estimation less accurate. The AIC method tends to choose the stationary model, and hence performs well when the correct model is stationary. This is due to the penalty term for the number of parameters to be estimated in the AIC, giving preference to the model with less parameters.

## VII. APPLICATION TO EPILEPSY DATA

We apply the methods of residual variance, LOF, and AIC to real MEG data of epilepsy. It is not possible to use the Smirnov

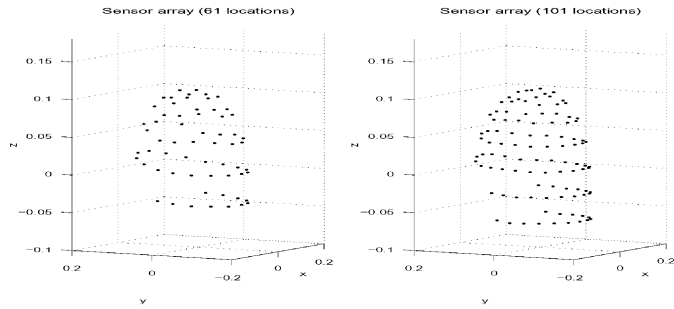


Fig. 2. Sensor array configurations for real MEG data, (a) with 61 recording locations and (b) the system with 101 recording locations.

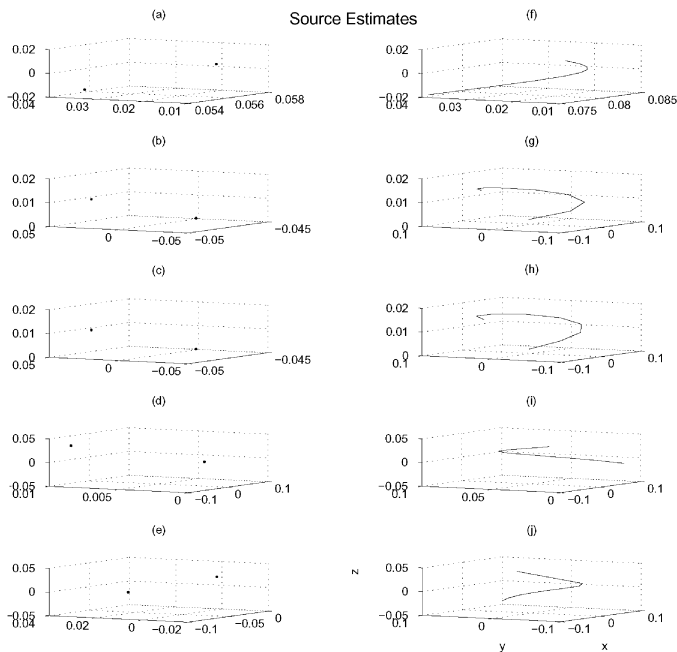


Fig. 3. Estimated source positions for all 5 real MEG data sets. Left: results assuming the stationary sources model and right: assuming moving source model. The symbol “.” shows the estimated locations of the stationary sources, and the continuous line shows the estimated path of the single moving source.

and  $\chi^2$  tests since they require precise knowledge of the noise variance. In real life, we can only estimate the noise variance which brings us to the LOF test instead of  $\chi^2$  test.

We use five data sets taken from three subjects. All measurements are made using single trials, and filtered at 6–50 Hz for noise reduction. We have 104 to 196 time samples for the data, covering a time duration of approximately 0.35 to 0.65 s. A pair of data sets are obtained using a pair of planar gradiometers at 61 locations resulting in 122 sensors, and the remaining three using 102 pairs of planar gradiometers resulting in 204 sensors as shown in Fig. 2. For the moving source, we have used three basis functions:  $\phi_1(t) = 1$ ,  $\phi_2(t) = t$ , and  $\phi_3(t) = t^2$ . Fig. 3 shows the estimated source positions under the two alternative models. Fig. 3 (a)–(e) are the estimated positions assuming the stationary sources model, and Fig. 3(f)–(j) assuming the moving source model. The results of the model selection methods is given in Table II. For each data set, we give the value of the test statistics for the residual variance, Smirnov, LOF, and AIC tests. The selected model is listed in the last column of this table. We

TABLE II

RESULTS OF MODEL SELECTION METHODS FOR REAL MEG DATA. THE FIRST AND SECOND COLUMN FOR EACH METHOD SHOWS THE CRITERION VALUE UNDER THE STATIONARY SOURCES MODEL AND THE MOVING SOURCE MODEL RESPECTIVELY, AND THE THIRD COLUMN SHOWS THE SELECTED MODEL (STATIONARY OR MOVING)

Set	Res. Var.			LOF			AIC		
1	0.39	0.51	S	1.003	1.009	S	87125	87533	S
2	0.41	0.24	M	1.007	1.001	M	65132	64876	M
3	0.40	0.24	M	1.007	1.001	M	65127	64867	M
4	0.45	0.22	M	1.007	1.001	M	54982	54112	M
5	0.56	0.33	M	1.009	1.004	M	59871	58994	M
	S	M	Sel						

observe that all the model selection methods agree for all data sets, making the decisions quite confident.

## VIII. CONCLUSION

In this paper, we have compared five statistical methods for distinguishing between a single moving source and two stationary sources that may result in similar EEG/MEG measurements. These two models may appear for example in intralobular activities or Jacksonian seizures (e.g., epilepsy). The methods of residual variance, Smirnov test,  $\chi^2$  test, LOF, and AIC were reviewed and applied to the problem.

The limiting behavior of the methods were investigated analytically. We showed that when the number of trials goes to infinity and one of the alternatives is correct, all of the methods select the correct model with probability one. We have also conducted computer simulations to compare the methods, varying parameters such as the separation of the two source locations, depths of the sources and SNR. The methods were observed to perform better for well separated sources, sources that are closer to the surface, and high SNR. Additionally, we have applied the methods to real MEG data for epilepsy for certain cases where both models may have been valid as discussed in Section I. We observed that, the results of all methods agreed for all cases, giving us the information about the model that should be used.

We have assumed for simplicity that the additive noise is uncorrelated in space and time. When this assumption fails (that is an incorrect noise covariance is used), only the covariance of the measurements will be affected and the primary factor for distinguishing between the models is the mean. However, we should note that a severely incorrect noise covariance may overcome the effect of difference in the means, particularly when the means are close (when it is more difficult to distinguish between the models). The models we have used are certainly far from representing a realistic scenario of the electrical activity in the brain. However, focal sources are used very commonly to represent activities confined in a small area even though the activity may, in fact, have a local spread. The moving dipole model also assumes that the activity consists of delta function like spikes which is a simplification of the electrical activity in the brain. Nevertheless, in this paper we aim to make a coarse distinction between moving and stationary models, and do not claim that the epileptic spikes (moving or stationary) can perfectly be represented using delta like functions. We should also note that

DOFs of the models (number of samples, and number of parameters that determine the path of the moving dipole model) will affect the goodness of fit. However, this effect is likely to cause significant errors when the two compared models are nested. For nested models, even a simpler model (a special case of the more complex model) may result in a better fit when compared to the more complex model when the number of time samples are different. However, the models that are compared in this paper are nonnested models and hence are less effected than the differences between the goodness of fit of a particular model. That is, more number of samples may result in a better fit for a certain model, but this is not likely to result in the selection of an incorrect model since the two models are fundamentally different unlike nested models.

As a future research, it would be interesting to develop confidence criteria for the selection methods. The criteria would be useful to choose the method to be used or combine several model selection methods. Also it is possible to generalize the methods to correlated noise. Another interesting subject would be to explore sequential tests in terms of the number of trials,  $K$ . It is worth to investigate a test that decides whether the current trials are sufficient or more trials are needed for a confident decision.

## REFERENCES

- [1] R. Hari, K. Reinikainen, E. Kaukoranta, M. Hamalainen, R. Ilmoniemi, A. Penttinen, J. Salminen, and D. Teszner, "Somatosensory evoked cerebral magnetic fields from SI and SII in man," *Electroencephalogr. Clin. Neurophysiol.*, vol. 57, pp. 254–263, 1984.
- [2] J. C. de Munck, "The estimation of time varying dipoles on the basis of evoked potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 77, pp. 156–160, 1990.
- [3] A. Nehorai and A. Dogandzic, "Estimation of propagating dipole sources by EEG/MEG sensor arrays," in *Proc. 32nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 1998, pp. 304–308.
- [4] S. Baillet and L. Garnero, "A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 5, pp. 522–534, May 1999.
- [5] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai, "On detection of the number of signals when the noise covariance is arbitrary," *J. Multi-Variate Anal.*, vol. 77, pp. 156–160, 1986.
- [6] J. P. Reilly, K. M. Wong, P. C. Yip, and Q.-T. Zhang, "On information theoretic criteria for determining the number of signals in high resolution array processing," *IEEE Trans. Signal Process.*, vol. 38, no. 11, pp. 1959–1971, Nov. 1990.
- [7] T. R. Knosche, E. M. Berends, H. R. A. Jagers, and M. J. Peters, "Determining the number of independent sources of the EEG: A simulation study on information criteria," *Brain Topogr.*, vol. 11, pp. 111–124, 1998.
- [8] J. L. Waldrop, H. M. Huizenga, R. P. P. Grasman, K. B. E. Bocker, J. C. de Munck, and P. C. M. Molenaar, "Model selection in electromagnetic source analysis with an application to VEF's," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1121–1129, Oct. 2002.
- [9] J. T. Murphy, H. C. Kwan, W. A. MacKay, and W. C. Wong, "Physiological basis for focal motor seizures and the Jacksonian march phenomena," *Can. J. Neurol. Sci.*, vol. 7, no. 2, pp. 79–85, 1980.
- [10] P. Teale, J. Sheeder, D. C. Rojas, J. Walker, and M. Reite, "Sequential source of the M100 exhibits inter-hemispheric asymmetry," *Neuroreport*, vol. 9, no. 11, pp. 2647–2652, 1998.
- [11] R. L. Rogers, A. C. Papanicolaou, S. B. Baumann, C. Saydjari, and H. M. Eisenberg, "Neuromangetic evidence of a dynamic excitation pattern generating the N100 auditory response," *Electroencephalogr. Clin. Neurophysiol.*, vol. 77, no. 3, pp. 237–240, 1990.
- [12] C. 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.
- [13] M. Hamalainen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Modern Phys.*, vol. 65, pp. 413–497, 1993.
- [14] G. Edlinger, P. Wach, and G. Pfurtscheller, "On the realization of an analytic high-resolution EEG," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 6, pp. 736–745, Jun. 1998.
- [15] A. Dogandzic and A. Nehorai, "EEG/MEG spatio-temporal dipole source estimation and array design," in *High-Resolution and Robust Signal Processing*, A. Gershman, Y. Hua, and Q. Cheng, Eds. New York: Marcel Dekker, 2003, ch. 7, pp. 393–442.
- [16] —, "Estimating 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.
- [17] J. R. Scott, *Matrix Analysis for Statistics*. New York: Wiley, 1997.
- [18] H. M. Huizenga, J. C. de Munck, L. J. Waldrop, and R. 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.
- [19] J. C. de Munck, H. M. Huizenga, L. J. Waldrop, 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.
- [20] R. M. Leahy, P. S. Lewis, and J. C. Mosher, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 6, pp. 541–557, Jun. 1992.
- [21] H. M. Huizenga and P. C. M. Molenaar, "Estimating and testing the sources of evoked potentials in the brain," *Multivariate Behavioral Res.*, vol. 29, pp. 237–267, 1994.
- [22] R. Monzingo and T. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
- [23] S. Supek and C. J. Aine, "Simulation studies of multiple dipole neuro-magnetic source localization: Model order and limits of source resolution," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 6, pp. 529–540, Jun. 1993.
- [24] J. W. Neill, "Testing for lack of fit in nonlinear regression," *Ann. Statist.*, vol. 16, pp. 733–740, 1988.
- [25] H. Akaike, "Information and an extension of the likelihood principle," presented at the Proc. 2nd Int. Symp. Information Theory, Suppl. Problems of Control and Information Theory, Budapest, Hungary, 1973, pp. 267–281.
- [26] R. T. Johnson, W. C. Black, and D. S. Buchanan, "Source detectability for multi-channel biomagnetic sensors," in *Proc. 10th Int. Conf. Biomagn.*, Santa Fe, NM, Feb. 1996.
- [27] J. C. Mosher, M. E. Spencer, R. M. Leahy, and P. S. Lewis, "Error bounds for EEG and MEG dipole source localization," *Electroencephalogr. Clin. Neurophysiol.*, vol. 86, pp. 303–321, 1993.
- [28] L. A. Geddes and L. E. Baker, "The specific resistance of biological material, a compendium of data for the biomedical engineer and physiologist," *Med. Biol. Eng.*, no. 5, pp. 271–293, 1967.



**İmam Şamil Yetik** was born in Istanbul, Turkey, in 1978. He received the B.Sc. degree in electrical and electronics engineering from Bogazici University, Istanbul, Turkey, in 1998, the M.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2000, and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Chicago, IL, in 2004.

His research interest is mainly statistical sensor array signal processing and image processing with applications to biomedicine.



**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.

After graduation he worked as a Research Engineer for Systems Control Technology, Inc., in Palo Alto, CA. From 1985 to 1989, he was Assistant Professor and from 1989 to 1995 he was Associate Professor with the Department of Electrical Engineering at Yale University, New Haven, CT. In 1995,

he joined the Department of Electrical Engineering and Computer Science at 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. In 2001, he was named University Scholar of the University of Illinois. He holds a joint professorship with the ECE and Bioengineering Departments at UIC. His research interests are in signal processing, communications, and biomedicine.

Dr. Nehorai is Vice President-Publications and Chair of the Publications Board of the IEEE Signal Processing Society. He is also a member of the Board of Governors and of the Executive Committee of this Society. He was Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from January 2000 to December 2002, and is currently a Member of the Editorial Board of *Signal Processing*, the *IEEE Signal Processing Magazine*, and *The Journal of the Franklin Institute*. He is the founder and Guest Editor of the special columns on Leadership Reflections in the *IEEE Signal Processing Magazine*. He has previously been an Associate Editor of IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, IEEE JOURNAL OF OCEANIC ENGINEERING, and *Circuits, Systems, and Signal Processing*. He served as Chairman of the Connecticut IEEE Signal Processing Chapter from 1986 to 1995, and a Founding Member, Vice-Chair, and later Chair of the IEEE Signal Processing Society's Technical Committee on Sensor Array and Multichannel (SAM) Processing from 1998 to 2002. He was the co-General Chair of the First and Second *IEEE SAM Signal Processing Workshops* held in 2000 and 2002. He was co-recipient, 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 received the Faculty Research Award from the UIC College of Engineering in 1999 and was Adviser of the UIC Outstanding Ph.D. Thesis Award in 2001. 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.

**Jeffrey David Lewine** received the Ph.D. degree in neuroscience from the University of Rochester, Rochester, MN, in 1989.

Following that he was appointed as a Director's Fellow at Los Alamos National Laboratory, where he worked on the development of MEG technology. He has previously been on the faculty of the University of New Mexico and the University of Utah. At present, he is an Associate Professor of Neurology at the University of Kansas Medical Center and the Director of the Magnetoencephalography program at the Hoglund Brain Imaging Center at KUMC. He also holds an adjunct appointment at the University of Illinois at Chicago. His research focuses on the utilization of brain imaging technologies for the evaluation of neurological and psychiatric disease.



**Carlos H. Muravchik** (S'81–M'83–SM'99) was born in Argentina, June 11, 1951. He graduated as an Electronics Engineer from the National University of La Plata, La Plata, Argentina, in 1973. He received the M.Sc. in statistics (1983) and the M.Sc. (1980) and Ph.D. (1983) degrees in electrical engineering, from Stanford University, Stanford, CA.

He is a Professor at the Department of the Electrical Engineering of the National University of La Plata and a member of its Industrial Electronics, Control and Instrumentation Laboratory (LEICI). He is also a member of the Comision de Investigaciones Cientificas de la Pcia. de Buenos Aires. He was a Visiting Professor at Yale University, New Haven, CT, in 1983 and 1994, and at the University of Illinois at Chicago in 1996, 1997, 1999, and 2003. His research interests are in the area of statistical signal and array processing with biomedical, control, and communications applications, and nonlinear control systems.

Since 1999, Dr. Muravchik is a member of the Advisory Board of the journal *Latin American Applied Research* and is currently an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.