ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 21 (2004) 1612–1621

Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data

Olaf Hauk*

Cognition and Brain Sciences Unit, Medical Research Council, Cambridge, UK

Received 19 August 2003; revised 6 December 2003; accepted 9 December 2003

The present study aims at finding the optimal inverse solution for the bioelectromagnetic inverse problem in the absence of reliable a priori information about the generating sources. Three approaches to tackle this problem are compared theoretically: the maximum-likelihood approach, the minimum norm approach, and the resolution optimization approach. It is shown that in all three of these frameworks, it is possible to make use of the same kind of a priori information if available, and the same solutions are obtained if the same a priori information is implemented. In particular, they all yield the minimum norm pseudoinverse (MNP) in the complete absence of such information. This indicates that the properties of the MNP, and in particular, its limitations like the inability to localize sources in depth, are not specific to this method but are fundamental limitations of the recording modalities. The minimum norm solution provides the amount of information that is actually present in the data themselves, and is therefore optimally suited to investigate the general resolution and accuracy limits of EEG and MEG measurement configurations. Furthermore, this strongly suggests that the classical minimum norm solution is a valuable method whenever no reliable a priori information about source generators is available, that is, when complex cognitive tasks are employed or when very noisy data (e.g., single-trial data) are analyzed. For that purpose, an efficient and practical implementation of this method will be suggested and illustrated with simulations using a realistic head geometry.

Keywords: Linear inverse theory; Minimum norm estimation; EEG; MEG; Source modeling; Dipole

Introduction

In recent years, considerable effort has been spent on improving the spatial resolution of source estimation procedures in the analysis of electroencephalographic (EEG) and magnetoencephalographic (MEG) data. One of the ambitious goals in this endeavor is to make the interpretation of the results comparable to metabolic imaging methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), which in turn lack the temporal resolution to track fast perceptual and cognitive processes in the human brain on the millisecond scale. Distributed inverse solutions, mostly of the linear type, are widely used to estimate the neuronal current distributions underlying the measured EEG and MEG signals, especially in studies on higher cognitive brain function (Dhond et al., 2003; Haan et al., 2000; Halgren et al., 2002; Hauk et al., 2001; Pulvermüller and Shtyrov, in press; Rinne et al., 2000). In some cases, descriptions of methods to estimate the neuronal sources from EEG and MEG signal suggest that tomography-like procedures exist, if only the right mathematical assumptions are implemented into the corresponding algorithms (Ioannides et al., 1995; Pascual-Marqui et al., 2002; Singh et al., 2002). However, it is usually implicitly or explicitly acknowledged that according to the Helmholtz principle, the bioelectromagnetic inverse problem has no unique solution (von Helmholtz, 1853). Whatever result is obtained with one method, there are still infinitely many other possible solutions equally compatible with the recorded signal.

This alone implies that the question "Which method yields the correct solution to the bioelectromagnetic inverse problem?" without further specifications is futile. EEG or MEG signals alone do not carry sufficient information to determine the precise spatial distribution of the underlying neuronal sources. This problem can be compared to the reconstruction of a three-dimensional object from its shadow: Only the shape of the object along a two-dimensional plane is given by the data, making it impossible to infer anything about its 3D structure by these data alone.

Source estimation from EEG and MEG data therefore requires reformulating the question. Two general strategies can be distinguished: (1) focusing on those solution parameters that can be estimated reliably from the data alone; (2) including a priori knowledge from other sources than the data under analysis, thus reducing the amount of parameters to be estimated to a tractable number. Applied to the example of object reconstruction from a shadow, an example for strategy 1 would be to ask "What are the maximum extensions of the object in the plane parallel to the projection screen?" An example for strategy 2 would be "I know that the shadow represents the profile of a face, and it's either Gérard Depardieu or Michael Jackson. So does the profile have a big nose?" Obviously, the information we get following strategy 1 is rather limited, but might be enough for a given purpose (e.g., if we want to push the object through our front door). Following strategy 2, we get very specific information, but if we overlooked the possibility that another big-nosed person might have been in front of the screen, our conclusion could be completely wrong.

* Cognition and Brain Sciences Unit, Medical Research Council, 15 Chaucer Road, CB2 2EF, Cambridge, UK. Fax: +44-1223-359062.
E-mail address: olaf.hauk@mrc-cbu.cam.ac.uk.
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2003.12.018

The same logic applies to source estimation procedures: Either questions must be asked that explicitly take into account the limitations of the technique, or restrictive modeling assumptions must be introduced. The problem, however, is that this can still be done in various ways, leading to a large number of procedures among which the experimenter has to choose (see Baillet et al., 2001; Dale and Halgren, 2001; Fuchs et al., 1999; Grave de Peralta Menendez et al., 1997a; Ilmoniemi, 1993; Pascual-Marqui, 1999; Vrba and Robinson, 2001 for comparisons or overviews of methods). It is therefore essential to investigate what features of an estimated current distribution are determined by the data, and how much depends on the specific method and corresponding modeling assumptions.

In this work, three approaches often employed in the literature to tackle the inverse problem will be compared: (1) the statistical approach, which aims at finding the "most likely" solution compatible with the data and possibly further constraints; (2) the minimum norm approach, which aims at finding a solution that is compatible with the data and fulfills further constraints on the amplitudes and/or covariances between source strengths; (3) the resolution optimization approach, which aims at estimating the source components as independently as possible from each other. It will be shown that all of these approaches are able to implement the same amount of a priori information, and that if this information is the same, their solutions are also the same. Most important is the case where no a priori knowledge is available at all, and the question "What is determined by the data alone?" is asked. Obviously, the amount of information present in the data cannot vary with the procedure employed to extract it, and therefore one may hope that a clear result for this case can be found. Indeed, all three approaches converge on the same method in this specific case, namely the classical minimum norm solution. This is of particular importance because even though the amount of a priori knowledge from metabolic imaging techniques or neuropsychology is continuously growing, in many experiments, the number, extension, or approximate locations of the generators are not reliably known. Obvious examples for this case are data obtained with complex cognitive paradigms, noisy data from individual subjects, with a low number of trials or on a single-trial level.

Based on these findings, an implementation for the classical minimum norm method is introduced that takes into account its limitations as well as its strengths. It is suggested for standardized routine analysis of large and complex data sets, such as in EEG and MEG studies on higher cognitive function, or for analyzing data on a single-trial level.

Theory

General

In the following, bold capital letters (like **G**) represent matrices, and bold small letters (like **s**) refer to column vectors. $\mathbf{G}_{i.}$ and $\mathbf{G}_{.i}$ represent the ith row and column of the matrix **G**, respectively. The superscript "^{T}" denotes the transposition of a vector or a matrix (e.g., $\mathbf{G}^T$ or $\mathbf{s}^T$).

The relationship between a given source distribution inside the head and the electric potential or magnetic field measured at discrete points on or above the scalp surface is linear (Geselowitz, 1967; Sarvas, 1987). The naturally continuous current distribution can be approximated by covering those brain areas that must be assumed to be activated by a large number of narrowly spaced dipole sources. This can be formulated in matrix notation as

$$\mathbf{d} = \mathbf{Ls} \tag{1}$$

where **d** is a (m*1)-vector of potentials at m electrode locations, **s** is a (n*1)-vector containing the amplitudes of n current sources with fixed locations and orientations, and **L** is the so-called (m*n)-"leadfield matrix" that takes into account the flow of volume currents due to volume conduction, and contains information about the geometry and conductivity distribution within the head. Each column of **L** contains the forward solution for one of the current sources contained in the model, that is, the signal at all sensors produced by this source alone with unit strength.

The data **d** are given by the recordings, and the leadfield matrix **L** is determined by the head geometry. The bioelectromagnetic inverse problem then consists of solving Eq. (1) for the unknown source distribution **s**. If the source distribution **s** contains more independent parameters than there is independent information in the recording **d** (e.g., if $n > m$), then the source amplitudes represented by **s** cannot be estimated independently of each other (Bertero et al., 1985). In other words, Eq. (1) is underdetermined and the corresponding inverse problem is ill-posed.

Another way to state this under-determinacy is to say that there are sources $\mathbf{s}_0 \neq 0$ for which $\mathbf{Ls}_0 = 0$, that is, these sources are orthogonal to the rows of the leadfield matrix. If the number of sources is much bigger than the number of sensors, so is the dimension of the null-space compared with the space spanned by the rows of the leadfield.

One interesting consequence of this is that whenever one has found a solution **s** to Eq. (1), one can add an arbitrary solution $\mathbf{s}_0$ such that $\mathbf{s} + \mathbf{s}_0$ is another solution of Eq. (1). In fact, every source distribution **s** can be separated into parts $\mathbf{s}_1$ and $\mathbf{s}_2$ such that

$$\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2, \quad \mathbf{p} = \mathbf{Ls}$$

$$\mathbf{Ls}_1 = \mathbf{p}, \quad \mathbf{Ls}_2 = 0 \tag{2}$$

that is, into one part that is orthogonal to the columns of the leadfield matrix, and another part that is not (Bertero et al., 1985; Menke, 1989).

Resolution of linear estimation techniques

If a linear estimator **G** is multiplied to the data, the relationship between the real but unknown source **s** and its estimate $\hat{\mathbf{s}}$ is (using Eq. (1))

$$\hat{\mathbf{s}} = \mathbf{Gd} = \mathbf{GLs} = \mathbf{Rs} \tag{3}$$

The newly defined matrix **R** describes the "distortion" or "blurring" of the real source by the measurement and estimation technique, and was therefore termed "resolution matrix". It has dimension (n*n) but highest possible rank m, so it is, unfortunately, not invertible in the usual case $m < n$, which is another way to state the under-determinacy of linear inverse problems (Grave de Peralta Menendez et al., 1997a; Menke, 1989).

Alternatively, each row $\mathbf{G}_{i.}$ of the linear estimator **G** is the estimator for one current source $\hat{s}_i$ in $\hat{\mathbf{s}}$. This way, the resolution vector $\mathbf{r}_i^T = \mathbf{G}_{i.}\mathbf{L}$ (i.e., the ith row of the resolution matrix **R**) can be defined for each source estimate $\hat{s}_i$. The vector $\mathbf{r}_i$ describes the influence each source contained in **s** would have on the estimate $\hat{s}_i$.

Approaches to tackle the inverse problem

Null-space approach

In the case where nothing is known about the source distribution to be estimated, one might want to determine that part of the source that is solely determined by the data. Formally, one would like to find a solution $\hat{\mathbf{s}}$ for which

$$\mathbf{L}\hat{\mathbf{s}} = \mathbf{d} \tag{4}$$

that is, which produces the recorded signal, but which also cannot be separated into parts $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$ such that

$$\hat{\mathbf{s}} = \hat{\mathbf{s}}_1 + \hat{\mathbf{s}}_2 \text{ and } (\mathbf{L}\hat{\mathbf{s}}_1 = \mathbf{0} \text{ or } \mathbf{L}\hat{\mathbf{s}}_2 = \mathbf{0}) \tag{5}$$

In other words, $\hat{\mathbf{s}}$ shall not contain any nonvanishing part that taken by itself would not produce any measurable signal at all sensors. This implies that $\hat{\mathbf{s}}$ should not contain parts that are orthogonal to the rows of the leadfield matrix $\mathbf{L}$, and thus $\hat{\mathbf{s}}$ must be a linear combination of the rows of $\mathbf{L}$, that is

$$\mathbf{s} = \mathbf{L}^T\mathbf{w} \tag{6}$$

where $\mathbf{w}$ stands for a weighting vector to be determined (Bertero et al., 1985).

Since $\mathbf{d} = \mathbf{Ls}$, it follows that $\mathbf{d} = \mathbf{LL}^T\mathbf{w}$, and if the rows of $\mathbf{L}$ are linearly independent (which is the case if the observations from different sensors are independent), this can be inverted to yield

$$\mathbf{w} = (\mathbf{LL}^T)^{-1}\mathbf{d}, \text{ and thus } \mathbf{s} = \mathbf{L}^T\mathbf{w} = \mathbf{L}^T(\mathbf{LL}^T)^{-1}\mathbf{d} \tag{7}$$

The resulting matrix $\mathbf{L}^T(\mathbf{LL}^T)^{-1}$ is the "minimum norm pseudoinverse" (MNP) or "Moore–Penrose Inverse" for the linear system in Eq. (1) (Golub and van Loan, 1996). In this section, it was derived solely under two requirements: (1) the solution should explain the data $\mathbf{d}$; (2) it should not contain any part of the null space, which taken by itself would not produce any measurable signal at any of the recording sites. The second requirement is plausible in the case where no a priori information about the sources is available, and the solution shall only contain features predicted by the recorded data alone. In the following, it will be shown that several seemingly very different approaches yield the MNP if no further a priori information is provided.

Minimum norm approach

A unique solution to the inverse problem stated in Eq. (1) can be found by combining constraints on the solution (e.g., that it has minimal power) and constraints on the data it predicts (e.g., that it comes close to the measured data). A general formulation in a linear framework is to require

$$(\hat{\mathbf{s}} - \hat{\mathbf{s}}_0)^T\mathbf{C}_s(\hat{\mathbf{s}} - \hat{\mathbf{s}}_0) = \min \tag{8}$$

for the solution, where $\hat{\mathbf{s}}$ is the estimated solution, $\hat{\mathbf{s}}_0$ is an a priori approximation of the solution, and $\mathbf{C}_s$ is a weighting matrix, representing the metric associated with the source space (e.g., a priori knowledge about the approximate locations or covariances of sources), together with

$$(\mathbf{L}\hat{\mathbf{s}} - \mathbf{d})^T(\mathbf{L}\hat{\mathbf{s}} - \mathbf{d}) = \min \tag{9}$$

where $\mathbf{L}\hat{\mathbf{s}}$ are the predicted data, and $\mathbf{d}$ are the measured data.

If the matrix $\mathbf{C}_s$ is positive definite (and therefore invertible), the solution to this problem is (Grave de Peralta Menendez and Gonzalez Andino, 1998):

$$\hat{\mathbf{s}} = \hat{\mathbf{s}}_0 + \mathbf{C}_s^{-1}\mathbf{L}^T(\mathbf{LC}_s^{-1}\mathbf{L}^T)^{-1}(\mathbf{d} - \mathbf{L}\hat{\mathbf{s}}_0) \tag{10}$$

If no a priori model $\hat{\mathbf{s}}_0$ is included, this equation reduces to

$$\hat{\mathbf{s}} = \mathbf{C}_s^{-1}\mathbf{L}^T(\mathbf{LC}_s^{-1}\mathbf{L}^T)^{-1}\mathbf{d} \tag{11}$$

The matrix $\mathbf{C}_s$ might be used to incorporate a priori information about brain areas in which active sources are expected, that is, if appropriate fMRI results are available (Dale et al., 2000). However, if sources can be expected at any location in the source space, each location in the source space must be given equal weight. In this case, $\mathbf{C}_s$ is the identity matrix, and the MNP is obtained:

$$\hat{\mathbf{s}} = \mathbf{L}^T(\mathbf{LL}^T)^{-1}\mathbf{d} \tag{12}$$

Regularization

So far, only the case that the data shall be explained completely by the solution was considered. However, in realistic situations, noise is present in the data. One approach is to substitute the constraint on the predicted data by

$$(\mathbf{L}\hat{\mathbf{s}} - \mathbf{d})^T\mathbf{C}_d(\mathbf{L}\hat{\mathbf{s}} - \mathbf{d}) = \varepsilon > 0 \tag{13}$$

where $\mathbf{C}_d$ is a positive definite weighting matrix, representing the "reliability" of the sensors (e.g., by their standard deviations or covariances), and ε reflects the part of the data that shall remain unexplained, that is, which is due to noise. The solution then changes to

$$\hat{\mathbf{s}} = \hat{\mathbf{s}}_0 + \mathbf{C}_s^{-1}\mathbf{L}^T(\mathbf{LC}_s^{-1}\mathbf{L}^T + \lambda\mathbf{C}_d^{-1})^{-1}(\mathbf{d} - \hat{\mathbf{s}}_0) \tag{14}$$

From this expression, weighted minimum norm solutions like that of Dale and Sereno (1993), Pascual-Marqui et al. (2002), and Wagner et al. (1996) can be derived. λ is the "regularization parameter", which needs to be determined such that ε reaches an optimal value. Without an a priori model $\hat{\mathbf{s}}_0$, and giving equal weight to all channels and source locations (i.e., taking the identity matrix for $\mathbf{C}_d$ and $\mathbf{C}_s$), the resulting solution is

$$\hat{\mathbf{s}} = \mathbf{L}^T(\mathbf{LL}^T + \lambda\mathbf{I}_d)^{-1}\mathbf{d} \tag{15}$$

which is the minimum norm solution with Tikhonov regularization (Bertero et al., 1988; Tikhonov, 1963).

Maximum likelihood approach

If the acquired data and the assumed solution follow a known probability distribution, it would be reasonable to determine the "most likely" solution to Eq. (1), that is, to maximize the likelihood $P(\mathbf{s},\mathbf{d})$ of a solution $\mathbf{s}$ given a specific measurement $\mathbf{d}$ (Clarke, 1989; Tarantola, 1994). This likelihood function can be separated into two separate parts: (1) the likelihood $P(\hat{\mathbf{s}})$ that a solution $\hat{\mathbf{s}}$ occurs independently of the data (i.e., a priori knowledge about the source distribution); (2) the likelihood $P(\mathbf{d},\hat{\mathbf{s}})$ that data $\mathbf{d}$ are a consequence of the source $\hat{\mathbf{s}}$ (i.e., the forward model).

The total distribution P($\hat{\mathbf{s}}$,**d**) is then proportional to the product of the latter two (Tarantola, 1994):

$$\mathbf{P}(\hat{\mathbf{s}}, \mathbf{d}) \sim \mathbf{P}_{\hat{s}d}(\mathbf{d}, \hat{\mathbf{s}})\mathbf{P}_s(\hat{\mathbf{s}})$$

One can assume that these distributions are Gaussian, that is

$$\mathbf{P}\ (\mathbf{s}) \sim \exp\{-(\hat{\mathbf{s}} - [\mathbf{s}])^T \mathbf{C}_s (\hat{\mathbf{s}} - [\mathbf{s}])\} \tag{16}$$

$$\mathbf{P}\ (\mathbf{d}, \hat{\mathbf{s}}) \sim \exp\{-(\mathbf{d} - \mathbf{L}\hat{\mathbf{s}})^T \mathbf{C}_d (\mathbf{d} - \mathbf{L}\hat{\mathbf{s}})\} \tag{17}$$

where exp{ } denotes the exponential function and [ ] the expectation value operator. It can be shown that a combination of Gaussian distributions is Gaussian as well, and the solution is (Tarantola, 1994):

$$\hat{\mathbf{s}} = [\mathbf{s}] + \mathbf{C}_s^{-1}\mathbf{L}^T(\mathbf{L}\mathbf{C}_s^{-1}\mathbf{L}^T + \lambda\mathbf{C}_d^{-1})^-(\mathbf{d} - \mathbf{L}[\mathbf{s}]) \tag{18}$$

where $^-$ indicates the generalized inverse, which should be used in this case since the expression to be inverted does not necessarily have full rank.

The regularization parameter λ was introduced to allow for continuous weighting of the influence of $\mathbf{C}_d$. This expression is identical to the weighted minimum norm solution in Eq. (14).

The matrix $\mathbf{C}_d$ can be used to incorporate differential reliability of the sensors (e.g., a metric or variance of the data). If no noise is taken into account, it can be neglected and the regularization parameter λ set to zero. If some areas in the brain are more likely to be activated than others, this information can be included in matrix $\mathbf{C}_s$, analogous to the weighted minimum norm approach described above. If no such information and no a priori model [**s**] exist, $\mathbf{C}_s$ should be chosen to be the identity matrix, and thus

$$\hat{\mathbf{s}} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{d} \tag{19}$$

again the MNP is obtained.

Regularization

If the data contain noise, the regularization parameter λ should be larger than zero. If all sensors have the same noise level, the matrix $\mathbf{C}_d$ is proportional to the identity matrix, i.e. $\mathbf{C}_d = \lambda\mathbf{I}$. This yields

$$\hat{\mathbf{s}} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T + \lambda\mathbf{I})^{-1}\mathbf{d} \tag{20}$$

which is the expression for Tikhonov regularization already derived in the previous section (Eq. (15)).

Backus–Gilbert approach

Each row $\mathbf{G}_{i.}$ of the linear estimator **G** is the estimator for one current source $\hat{s}_i$ in $\hat{\mathbf{s}}$ (see Resolution of linear estimation techniques section above). Ideally, $\mathbf{G}_{i.}$ should highly correlate with the potential distribution of the source of interest, $\hat{s}_i$, but would be orthogonal to all the others, that is, to all $\hat{s}_j$ with $j \neq i$. More formally, the values of the resolution vector $\mathbf{r}_i^T = \mathbf{G}_{i.}\mathbf{L}$ should be maximal at its ith element, and vanish anywhere else. This approach to underdetermined linear inverse problems was first suggested by Backus and Gilbert (1968), and a similar strategy was developed by Capon (1969). The ideal choice for $\mathbf{r}_i$ would be a (n*1)-vector containing zeroes everywhere except for the position of the source of interest (the ith element), at which it should be 1. This corresponds to the ith column of the (n*n)-identity matrix $\mathbf{I}_{.i}$ (Jackson, 1972).

However, the construction of an optimal $\mathbf{r}_i$ is restricted by the condition $\mathbf{r}_i^T = \mathbf{G}_{i.}\mathbf{L}$, that is, it is a linear combination of the rows of the leadfield matrix. One can only expect to optimize the behavior of the estimator $\mathbf{G}_{i.}$ to some degree. This can be done by defining the ideal resolution vector $\mathbf{t}_i$ and formulating a measure for its distance to the feasible resolution vector $\mathbf{r}_i$:

$$\min(\mathbf{W}(\mathbf{r}_i - \mathbf{t}_i))^2 \tag{21}$$

$$\Rightarrow \min((\mathbf{G}_i.\mathbf{L} - \mathbf{t}_i^T)\mathbf{W}^T\mathbf{W}(\mathbf{G}_{i.}\mathbf{L} - \mathbf{t}_i^T)^T) \tag{22}$$

$$\Rightarrow \min(\mathbf{G}_i.\mathbf{S}\mathbf{G}_{i.}^T - 2\mathbf{G}_i.\mathbf{u} + const) \tag{23}$$

with

$$\mathbf{S} = \mathbf{L}\mathbf{W}^T\mathbf{W}\mathbf{L}^T$$

and

$$\mathbf{u} = \mathbf{L}\mathbf{W}^T\mathbf{W}\mathbf{t}_i^T$$

W is a weighting matrix that might give differential weights to the components of the difference between $\mathbf{r}_i$ and $\mathbf{t}_i$. For example, the original idea of Backus and Gilbert (1968) was to increasingly penalize contributions from sources distant to the location of interest, such that activity produced at these distant sites projects less on the linear estimator than nearer ones.

The solution to this minimization problem is unique, and yields the estimator (Grave de Peralta Menendez et al., 1997a)

$$\mathbf{G}_{i.} = \mathbf{S}^{-1}\mathbf{u} \tag{24}$$

If the target vector $\mathbf{t}_i$ is chosen as the ideal resolution vector, that is, as one row of the identity matrix $\mathbf{I}_{i.}$ and if equal weight is given to the minimization of the difference between the obtainable resolution vector and the ideal one ($\mathbf{r}_i - \mathbf{t}_i$) (i.e., $\mathbf{W} = \mathbf{I}$), then $\mathbf{S} = \mathbf{L}\mathbf{L}^T$ and $\mathbf{u} = \mathbf{L}_{.i}$ (which is the ith column of the leadfield matrix), yielding

$$\mathbf{G}_{i.} = (\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}_{.i}. \tag{25}$$

Computing such an estimator for all solution points, one again obtains the MNP

$$\mathbf{G} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1} \tag{26}$$

The choice of the matrix **W** is critical for the resolution properties of the linear estimator. For example, special care has to be taken in the case of vector fields (Grave de Peralta Menendez and Gonzalez Andino, 1999). Furthermore, due to the flexibility in the choice of **W**, it is possible to obtain weighted minimum norm solutions of the type described above (Grave de Peralta Menendez et al., 1997b).

Regularization

If we assume that signal and noise are additive, then the linear estimate obtained with an inverse matrix **G** can be analogously decomposed:

$$\hat{\mathbf{s}} = \mathbf{G}\mathbf{d} = \mathbf{G}\mathbf{d}_s + \mathbf{G}\mathbf{d}_n = \hat{\mathbf{s}}_s + \hat{\mathbf{s}}_n \tag{27}$$

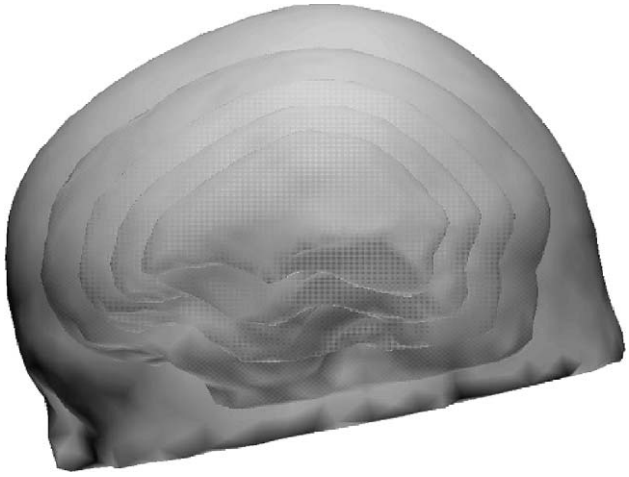

Fig. 1. Parts of the simulation set-up used in this paper. Shown are (from outside to inside) the skin and the liquor surface used in the BEM model, and shell 1, shell 2, and shell 3 used as the source space.

or elementwise:

$$\hat{s}_i = \mathbf{G}_{i.}\mathbf{d} = \mathbf{G}_{i.}\mathbf{d}_s + \mathbf{G}_{i.}\mathbf{d}_n = \hat{s}_{is} + \hat{s}_{in} \tag{28}$$

$\hat{\mathbf{s}}_n$ is the part of the estimate due to noise and should therefore be minimized. This can be accomplished in a least-squares framework for each element of $\hat{\mathbf{s}}$ individually:

$$\min([\mathbf{G}_{i.}\mathbf{n}]^2) \Rightarrow \min([\mathbf{n}^T(\mathbf{G}_{i.}\mathbf{G}_{i.}^T)\mathbf{n}]) \Rightarrow \min(\mathbf{G}_{i.}[\mathbf{n}^T\mathbf{n}]\mathbf{G}_{i.}^T) \Rightarrow \min(\mathbf{G}_{i.}\mathbf{C}_n\mathbf{G}_{i.}^T) \tag{29}$$

[ ]denotes the expectation value operator and $\mathbf{C}_n$ the noise covariance matrix.

This expression can be included in Eq. (3) with a regularization parameter λ that indicates how much weight is given to the suppression of noise:

$$\min((\mathbf{W}(\mathbf{G}_i.\mathbf{L} - \mathbf{t}_i^T))^2 + \lambda(\mathbf{G}_i.\mathbf{C}_n\mathbf{G}_i^T)) \tag{30}$$

This yields the unique solution:

$$\mathbf{G}_{i.} = (\mathbf{S} + \lambda\mathbf{C}_n)^{-1}\mathbf{u} \tag{31}$$

For the simplifications mentioned above (i.e., $\mathbf{W} = \mathbf{I}$, $\mathbf{t}_i = \mathbf{I}_{.i}$), this becomes

$$\mathbf{G}_{i.} = (\mathbf{L}\mathbf{L}^T + \lambda\mathbf{I})^{-1}\mathbf{L}_{.i}, \text{ or } \mathbf{G} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T + \lambda\mathbf{I})^{-1} \tag{32}$$

which is the formula for Tikhonov regularization as in Eq. (15).

One advantage of this approach is that the trade-off between resolution (given by the resolution vector $\mathbf{r}_i$) and the influence of noise (given by the variance of the source estimate at a specific location) can be controlled for each estimator $\mathbf{G}_{i.}$ (and thus for each estimation point in the source space) separately. Estimators of this kind are sometimes called "beamformers" (Barnes and Hillebrand, 2003; Sekihara et al., 2002; Vrba and Robinson, 2001) or "minimum variance spatial filters" (Van Veen et al., 1997), since the aim is to create spatial filters $\mathbf{G}_{i.}$ that focus only on activity generated in a circumscribed brain area, and are not sensitive to activity outside these areas and noise.

Implementation of the minimum norm method

Motivation

The classical minimum norm solution is obtained by all of the abovementioned approaches in the case of nonexisting a priori knowledge about the source to be estimated. It would therefore be applicable in many situations where minimum modeling

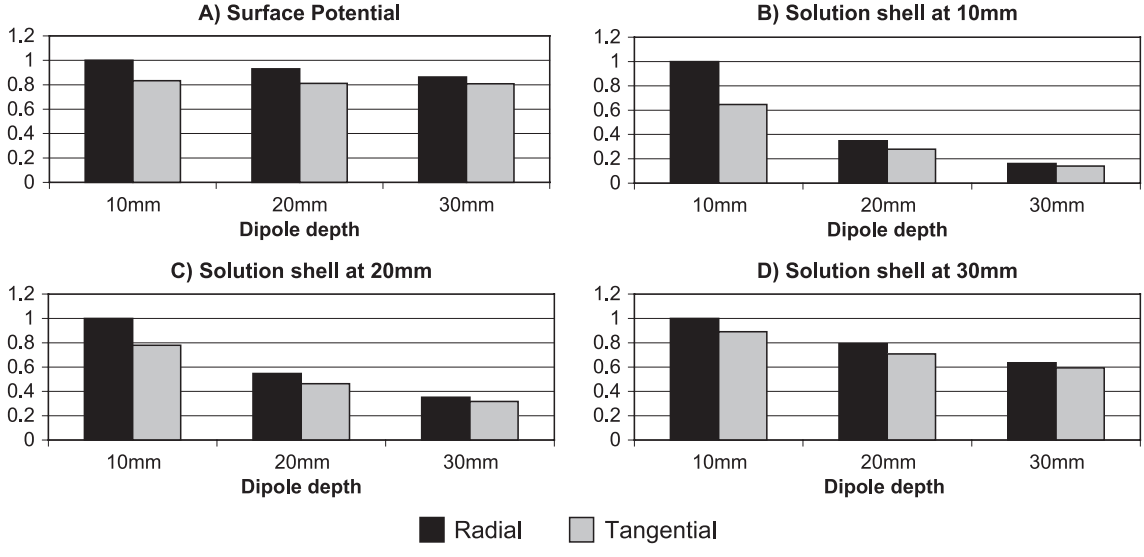

Fig. 2. Amplitude of the potential distribution (A) and the minimum norm estimates (B–D) in dependence of source depth. Radial and tangential sources were placed approximately beneath the vertex, 10, 20, and 30 mm below the cortical surface, respectively. Amplitudes were determined as the RMS across all electrodes or dipoles, respectively, and normalized to the maximum value in each diagram. For each source, a solution was computed on all three shells combined. The amplitudes on separate shells for these solutions, but for the same dipole sources, are shown in different diagrams (B–D).

assumptions are required, such as localization of complex cognition-related brain activity, analysis of continuous EEG and MEG, or analysis on a single-trial level. In the following, an implementation of the minimum norm method will be described that optimally exploits the strengths of this method, and allows for efficient data processing, statistical analysis of results, and visualization.

It has been described previously that the minimum norm method is not able to retrieve the depth of dipolar sources, even in the absence of noise (Fuchs et al., 1999; Grave de Peralta-Menendez and Gonzalez-Andino, 1998). The best we can therefore hope for is a two-dimensional projection of the true current distribution. This might at least tell us in which section of the brain a center of activity is located. This led several research groups to compute their source estimates on a two-dimensional surface right from the beginning (Edlinger et al., 1997; Hämäläinen and Ilmoniemi, 1994; Uutela et al., 1999; van Burik and Pfurtscheller, 1999). A consequence of this approach is that sources that are possibly active in the brain are not included in the source model. Although the depth of the source cannot be deduced from the estimate, one should still expect that the estimate becomes more reliable if no such modeling error is introduced. Considering the two-dimensionality of information in minimum estimates on the one hand, and the desirability of a three-dimensional source space for most accurate modeling on the other, it is suggested here to use a ‘‘shell model’’ for source reconstruction. Such a model could consist of several surfaces at different distances to the head center. Furthermore, including sources at different depths into the model allows to explicitly investigate the effect of source depth on the estimates.

Simulation configuration

The following simulations were carried out using the 152-subject T1 average brain of the Montreal Neurological Institute (MNI). Image processing was accomplished using the software package CURRY (Neuroscan Labs, Sterling, USA). The surfaces for skin, skull, liquor, and brain compartments were extracted interactively using a region growing algorithm and three-dimensional closing operations (Wagner et al., 1995). One hundred forty-eight electrodes were regularly distributed on a sphere and projected onto the skin surface of the MNI brain.

The brain surface was eroded by 10 mm (shell 1), 20 mm (shell 2), and 30 mm (shell 3), respectively, and the corresponding surfaces were used as shells in the source space for the simulations. This set-up is shown in Fig. 1. The shells contained a similar number of source locations each (996, 991, and 915, respectively), that is, the density of points was different for these shells. Therefore, the forward solutions for dipoles on shell 2 and shell 3 were each weighted by the factors 6.5/8 and 4.9/8, respectively. Side lengths of triangles connecting dipole locations were 8 mm on shell 1, 6.5 mm on shell 2, and 4.9 mm on shell 3. At each location, three orthogonal dipoles were placed (equivalent to one rotating dipole), two of them approximately tangential to the scalp surface and one of them radial.

The leadfield for this source configuration (i.e., the forward solutions for all dipoles included in the model with unit strength) was computed in a boundary element model (BEM) constructed from the MNI standard brain. CURRY standard parameters using the isolated problem approach and constant triangle potentials on three compartments (scalp, skull, liquor) were used. The whole BEM set-up comprised 4245 nodes. Side lengths of triangles within the surface were 10 mm for the skin surface, 9 mm for the skull surface, and 8 mm for the liquor surface.

Results

Fig. 2 presents the amplitudes obtained in the surface potential and on different minimum norm solution shells, in dependence of the depth of the dipole generating the surface potential. One radial

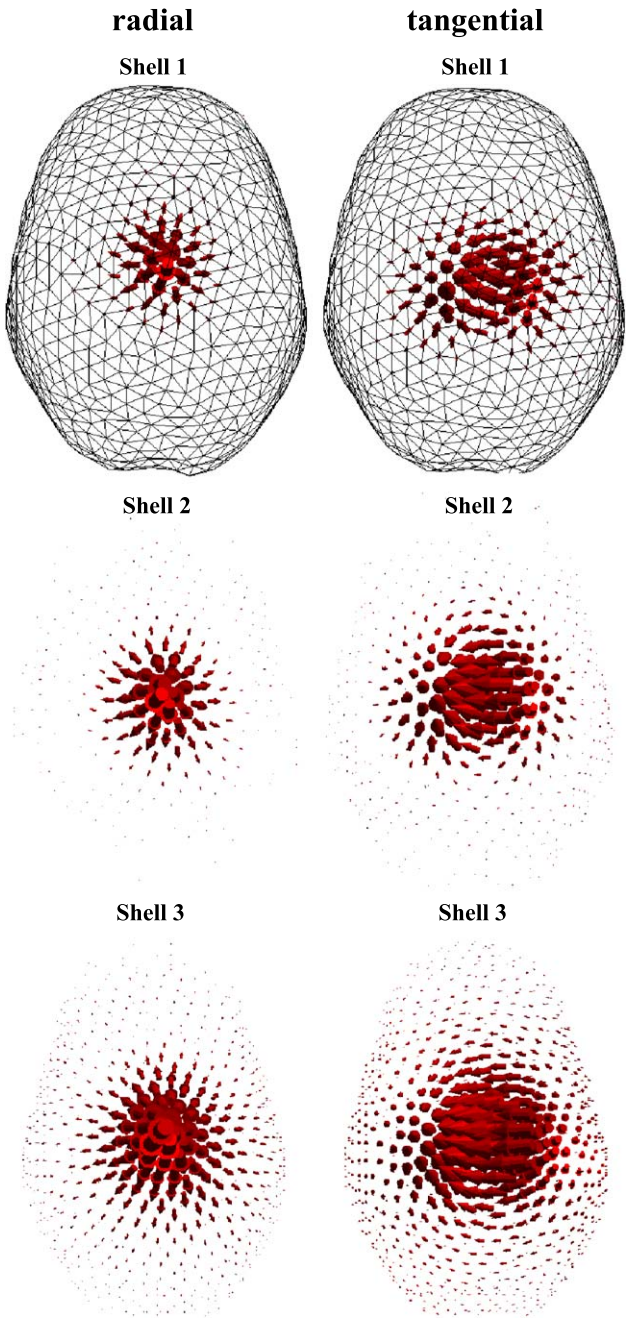

Fig. 3. Current source distribution including dipole orientations estimated for superficial radial and tangential sources (10 mm beneath vertex). The solutions were obtained for all three shells combined. The distributions for different shells (10, 20, and 30 mm below cortical surface) are displayed in separate images, each of which is scaled to its maximum value for better visualization. The corresponding RMS amplitudes are displayed in Fig. 2B.

and one tangential dipole were considered at each location. The amplitude was determined as the root-mean-square across all electrodes or dipoles, respectively. The general pattern of the results is not surprising: For radial as well as tangential dipoles, the amplitude decreases with increasing source depth, and for each source location, radial dipoles produce larger activity than tangential ones. More interestingly, the decline of amplitude with source depth for each minimum norm solution shell is steeper than for the surface potential. In other words, deeper dipoles are suppressed more strongly on any of the solution shells than in the plain EEG data. The more superficial the shell, the steeper the decline. In conclusion, the sensitivity of minimum norm estimates to deep dipole sources is worse than for the surface potential. This depth sensitivity depends on the depth of the solution shell, and improves with increasing depth.

Fig. 3 presents minimum norm solutions obtained for one radial and one tangential dipole placed near the vertex on shell 1. Each image is normalized to its maximum value, such that also smaller details of the distributions remain visible. The current estimates are shown with their orientations, that is, the length and orientation of the arrows at each source location indicate the strength and direction of the dipole moment, respectively. The amplitudes corresponding to these images can be found in Fig. 2B. Because these dipoles were located on the most superficial shell, and in an

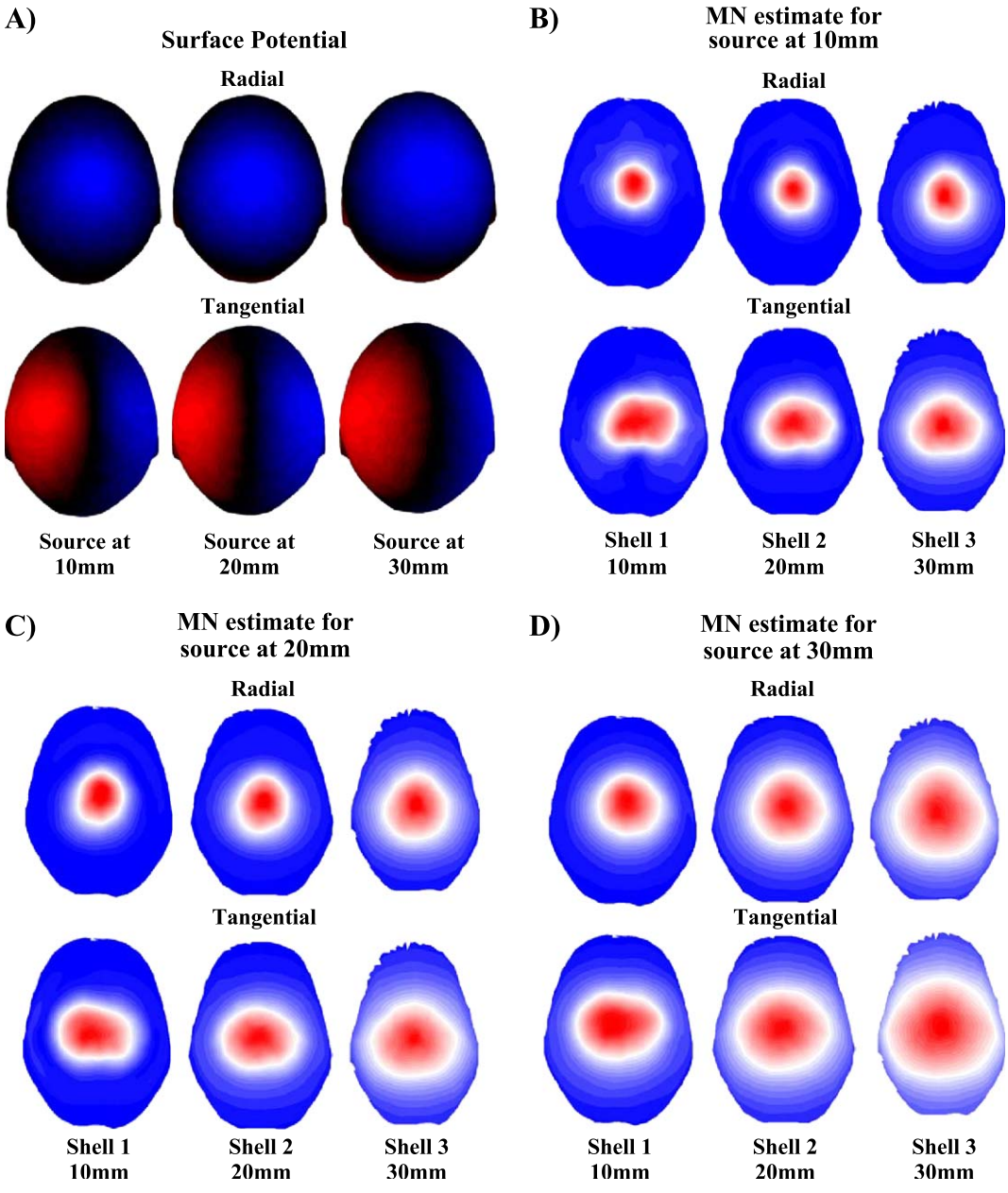

Fig. 4. Potential (A) and intensity distributions for current source estimates (B–D) obtained for radial and tangential sources at different depths (10, 20, 30 mm below the cortical surface beneath the vertex). The source estimates for sources at different depths are displayed in separate boxes. Each image is scaled to its maximum value for better visualization of its spatial extent.

area that is well-covered by the electrode array, the corresponding source estimates illustrate the best limit of the resolution power of this method.

To estimate the spatial extent of the source estimates in dependence of the solution shell and the depth of the dipoles, intensity maps are presented in Figs. 4B–D. Radial and tangential dipole sources were placed beneath the vertex, at depths of 10, 20, and 30 mm below the cortical surface. The surface potential distributions corresponding to these dipole sources are displayed in Fig. 4A. Like for Fig. 3, all intensity maps were normalized to their maximum value, so that smaller details can still be identified. The color scale was set such that white contour areas reflect half the intensity between zero and maximum intensity within each map. Red color therefore indicates the upper and blue color indicates the lower intensity range.

Superficial sources produce less extended solutions than deeper ones, and tangential sources produce more widespread peaks than radial ones. It is noteworthy, however, that the centers of these peaks always appear near the true dipole source location, and that none of the images in Figs. 3 and 4 contain any significant local maxima distant from the target location (“ghost sources”). Furthermore, the distributions on deeper solution shells are generally more widespread than those on superficial shells, though the centers of the peaks are approximately at the same locations. Shell 1 produces the sharpest images, and is therefore superior with respect to spatial resolution. This is in contrast with the results obtained from Fig. 2, where shell 1 was found to be least sensitive to deep sources. This suggest that a trade-off exists between spatial resolution (or the degree of blurring) and depth sensitivity of a source estimate.

Discussion

Theory

The theoretical part of this paper compared three main strategies to tackle the bioelectromagnetic inverse problem: the statistical, the minimum norm, and the resolution optimization approach. The main results are:

(1) All three approaches can incorporate the same constraints on the source distribution.
(2) All three approaches allow the handling of noisy data, and can take into account different noise levels at different sensors or the covariances between them.
(3) In the case where there is no a priori information implemented and the noise level is assumed to be the same in all sensors, all three approaches yield Tikhonov regularization.
(4) In the case where there is no a priori information and no regularization implemented, all three approaches yield the minimum norm pseudoinverse (MNP).

Point (1) implies that even though experimenters who chose different approaches can interpret their results differently (e.g., talk about the “most likely” solution in one case and about the “least energy consuming” in another), the actual source distributions might be exactly the same. Importantly, although the derivation of a solution within a certain framework might be complex, the crucial question to be asked is whether there is any reliable a priori information available for the problem at hand that can be implemented by the corresponding method. What is therefore needed to obtain a more detailed image of the source configuration underlying a measured potential distribution is not a more complicated mathematical approach. It would be more essential to provide reliable physiological information, for example, about the number or the approximate location of possibly active sources, which could then be incorporated into any of the outlined approaches. Unless such information is available, the classical minimum norm solution presents us with an, in some well-defined terms, optimal approach, that is, it only includes sources predicted from the data alone.

A simple example on the danger of accepting “invisible” sources can be constructed for the case of MEG. In a spherical head model, radial dipoles do not produce any measurable field at any sensor around the head, and therefore present a part of the null space (Sarvas, 1987). They could be added to any other solution obtained by any method, and would not increase the discrepancy between the predicted and the measured magnetic field (e.g., the residual variance). If such radial sources in some brain region are part of a source distribution, for example, as a result of a constraint that favors sources in that particular area, they might be interpreted as evidence for the activation of that brain area. However, the presence of these sources would not be supported by the data. It is therefore one thing to claim that a feature of the solution does not contradict the data and a certain hypothesis, but yet another thing to prove that the data themselves support this feature.

There are two ways of dealing with this problem. Either one can avoid it completely by using the MNP, or one can check with any given solution if the features of interest lie in the null space. If the latter option is chosen, the following aspects should be carefully considered:

(1) an explicit and clear justification for the a priori information must be provided,
(2) it should be clearly described how (e.g., by which of the abovementioned approaches) it was implemented,
(3) it should be demonstrated how much information is actually obtained from the data, for example, by comparison with the minimum norm solution.

Practical implementation

This paper suggests an efficient implementation of the MNP. In simulations, it produced single activation peaks around the correct two-dimensional spherical locations. Previous studies reported that the localization accuracy of the MNP in 3D space, usually measured as the distance between the true location of a simulated dipole and the largest peak in the distributed source estimate, is inferior to weighted minimum norm or nonlinear solutions (Fuchs et al., 1999; Grave de Peralta-Menendez and Gonzalez-Andino, 1998). However, the theoretical considerations of this paper show that the properties of the MNP reflect fundamental properties of the measurement configuration, and are not just a limitation of this particular method itself. For example, its inability to determine the depth of a source indicates that this information is simply not contained in the data. The potential or magnetic field generated by a single dipole at some depth inside the brain can always as well be explained by a distributed solution on a surface with an arbitrary radius. All these solutions would be fully compatible with the recorded data, and one could pick out any of them to claim that the source is located at that specific depth. However, the simulations of

this study and earlier reports suggest that the MNP can provide a blurred two-dimensional projection of the true sources (Hauk et al., 2002). If spherical coordinates are employed and the radius is neglected, the MNP can produce activation peaks directly above the true sources. This can already hint at the number and approximate location of sources in a particular data set, and can provide valuable information for more constrained source analysis (e.g., dipole modeling). In any case, this information is more reliable than the raw ERPs, and can easily be subjected to statistical analysis to compare for differences between experimental conditions (Hauk et al., 2001, 2002).

It should be emphasized that the conclusion of this paper is not that reliable a priori information does not exist at all and that the classical minimum norm solution should always be preferred to any other method. Obviously, in the case where reliable a priori information is available, it can significantly increase the resolution power of the analysis, and also localization in depth might become possible (Babiloni et al., 2003; Dale et al., 2000; Phillips et al., 2002). In the most fortunate case, one can assume a fixed number of focal sources, which can be modeled as equivalent dipoles, as for some perceptual processes (Di Russo et al., 2003; Lütkenhöner et al., 2003). The assumption of single focal sources can also be implemented in linear distributed methods (Grave de Peralta Menendez et al., 2001). One can expect that the more constrained the model and the less parameters to estimate, the more accurate the result will be, given the constraints are correct. In this paper, it was argued that if one does not want to run the risk of implementing incorrect constraints that might distort the solution, the MNP is a convenient and informative alternative.

Because it is linear, its localization accuracy for the case of multiple sources can easily be deduced by the superposition principle. The whole toolbox of linear inverse theory is at hand to objectively quantify the performance of this method (Craig and Brown, 1986; Fuchs et al., 1999; Grave de Peralta-Menendez and Gonzalez-Andino, 1998; Menke, 1989; Parker, 1994; Tarantola, 1994). This extends to parameters not considered in the simulations of this paper, as, for example, the influence of noise. Backus and Gilbert (1970) demonstrated analytically that already a small amount of regularization (i.e., with little loss of spatial resolution) already significantly increases the stability of the solution. Therefore, the present simulation results can be compared to data with high signal-to-noise ratio.

The suggested implementation allows efficient data reduction, because the critical parameter source depth is explicitly parameterized in the multiple-shell model. The data can be conveniently visualized, for example, for only one or few solution shells that might already carry the relevant information to be presented. The orientation of dipole sources in distributed source estimates is commonly not shown. One reason might be the difficulty to visualize them in a convenient manner, as, for example, on three-dimensional source spaces. Another reason might be that some experimenters are not interested in the orientation of the sources, but only in their strength. However, even if the intensity at one brain site is the same between two conditions, their source orientations could still be in completely opposite directions, which would be a dramatic difference that might be missed by analyzing intensity distributions alone. The shell approach suggested in this paper would make the visualization of source orientations conveniently possible.

For these reasons, this method and the presented implementation are well suited for standardized source estimation. It is suggested as a powerful analysis tool when much complex data have to be processed. This includes studies on higher cognitive function in which individual source estimation preceding statistical analysis is intended, or if source estimation is intended on a single-trial level before averaging, for example, as a preprocessing tool for time series analyses.

Acknowledgment

I would like to thank Rhodri Cusack for valuable comments on an earlier version of this manuscript.

References

Babiloni, F., Babiloni, C., Carducci, F., Romani, G.L., Rossini, P.M., Angelone, L.M., Cincotti, F., 2003. Multimodal integration of high-resolution EEG and functional magnetic resonance imaging data: a simulation study. NeuroImage 19, 1–15.

Backus, G.E., Gilbert, J.F., 1968. The resolving power of gross earth data. Geophys. J. R. Astron. Soc. 16, 169–205.

Backus, G.E., Gilbert, J.F., 1970. Uniqueness in the inversion of inaccurate gross earth data. Philos. Trans. R. Soc. Lond., A 266, 123–192.

Baillet, S.M., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. IEEE Signal Process. Mag., 14–30.

Barnes, G.R., Hillebrand, A., 2003. Statistical flattening of MEG beamformer images. Hum. Brain Mapp. 18, 1–12.

Bertero, M., De Mol, C., Pike, E.R., 1985. Linear inverse problems with discrete data: I. General formulation and singular system analysis. Inverse Problems 1, 301–330.

Bertero, M., De Mol, C., Pike, E.R., 1988. Linear inverse problems with discrete data: II. Stability and regularization. Inverse Problems 4, 573–594.

Capon, J., 1969. High-resolution frequency–wavenumber spectrum analysis. Proc. I.E.E.E. 57, 1408–1419.

Clarke, C.J.S., 1989. Probabilistic methods in a biomagnetic inverse problem. Inverse Problems 5, 999–1012.

Craig, I.J.D., Brown, J.C., 1986. Inverse Problems in Astronomy: A Guide to Inversion Strategies for Remotely Sensed Data. Adam Hilger, Bristol.

Dale, A.M., Halgren, E., 2001. Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. Curr. Opin. Neurobiol. 11, 202–208.

Dale, A.M., Sereno, M.I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J. Cogn. Neurosci. 5, 162–176.

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.

Dhond, R.P., Marinkovic, K., Dale, A.M., Witzel, T., Halgren, E., 2003. Spatiotemporal maps of past-tense verb inflection. NeuroImage 19, 91–100.

Di Russo, F., Martinez, A., Hillyard, S.A., 2003. Source analysis of event-related cortical activity during visuo-spatial attention. Cereb. Cortex 13, 486–499.

Edlinger, G., Pfurtscheller, G., van Burik, M., Wach, P., 1997. Post-movement EEG synchronization studied with different high resolution methods. Brain Topogr. 10, 103–113.

Fuchs, M., Wagner, M., Kohler, T., Wischmann, H.A., 1999. Linear and nonlinear current density reconstructions. J. Clin. Neurophysiol. 16, 267–295.

Geselowitz, D.B., 1967. On bioelectric potentials in an inhomogeneous volume conductor. Biophys. J. 7, 1–11.

Golub, G.H., van Loan, C.F., 1996. Matrix Computations, third ed. Johns Hopkins Univ. Press, Baltimore.

Grave de Peralta Menendez, R., Gonzalez Andino, S., 1998. Distributed source models: standard solutions and new developments. In: Uhl, C. (Ed.), Analysis of Neurophysiological Brain Functioning. Springer-Verlag, Heidelberg, pp. 176–201.

Grave de Peralta-Menendez, R., Gonzalez-Andino, S.L., 1998. A critical analysis of linear inverse solutions to the neuroelectromagnetic inverse problem. IEEE Trans. Biomed. Eng. 45, 440–448.

Grave de Peralta Menendez, R., Gonzalez Andino, S.L., 1999. Backus and Gilbert method for vector fields. Hum. Brain Mapp. 7, 161–165.

Grave de Peralta Menendez, R., Hauk, O., Gonzalez Andino, S., Vogt, H., Michel, C., 1997a. Linear inverse solutions with optimal resolution kernels applied to electromagnetic tomography. Hum. Brain Mapp. 5, 454–467.

Grave de Peralta Menendez, R., Gonzalez Andino, S., Hauk, O., Spinelli, L., Michel, C.M., 1997b. A linear inverse solution with optimal resolution properties: WROP. In: Tilg, B., Wach, P. (Eds.), Graz: Schiele&Schoen GmbH, Berlin, pp. 53–56.

Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity: I. Method descriptions and simulations. Brain Topogr. 14, 131–137.

Haan, H., Streb, J., Bien, S., Rosler, F., 2000. Individual cortical current density reconstructions of the semantic N400 effect: using a generalized minimum norm model with different constraints (L1 and L2 norm). Hum. Brain Mapp. 11, 178–192.

Halgren, E., Dhond, R.P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J.D., Dale, A.M., 2002. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. NeuroImage 17, 1101–1116.

Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32, 35–42.

Hauk, O., Rockstroh, B., Eulitz, C., 2001. Grapheme monitoring in picture naming: an electrophysiological study of language production. Brain Topogr. 14, 3–13.

Hauk, O., Keil, A., Elbert, T., Müller, M.M., 2002. Comparison of data transformation procedures to enhance topographical accuracy in time-series analysis of the human EEG. J. Neurosci. Methods 113, 111–122.

Ilmoniemi, R.J., 1993. Models of source currents in the brain. Brain Topogr. 5, 331–336.

Ioannides, A.A., Liu, M.J., Liu, L.C., Bamidis, P.D., Hellstrand, E., Stephan, K.M., 1995. Magnetic field tomography of cortical and deep processes: examples of "real-time mapping" of averaged and single trial MEG signals. Int. J. Psychophysiol. 20, 161–175.

Jackson, D.D., 1972. Interpretation of inaccurate, insufficient and inconsistent data. Geophys. J. R. Astron. Soc. 28, 97–109.

Lütkenhöner, B., Krumbholz, K., Lammertmann, C., Seither-Preisler, A., Steinsträter, O., Patterson, R.D., 2003. Localization of primary auditory cortex in humans by magnetoencephalography. NeuroImage 18, 58–66.

Menke, W., 1989. Geophysical Data Analysis: Discrete Inverse Theory. Academic Press, San Diego.

Parker, R.L., 1994. Geophysical Inverse Theory. Princeton Univ. Press, Princeton.

Pascual-Marqui, R.D., 1999. Review of methods for solving the EEG inverse problem. IJBEM 1, 75–86.

Pascual-Marqui, R.D., Esslen, M., Kochi, K., Lehmann, D., 2002. Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review. Methods Find. Exp. Clin. Pharmacol. 24 (Suppl. C), 91–95.

Phillips, C., Rugg, M.D., Friston, K.J., 2002. Anatomically informed basis functions for EEG source localization: combining functional and anatomical constraints. NeuroImage 16, 678–695.

Pulvermüller, F., Shtyrov, Y., 2004. Automatic processing of grammar in the human brain as revealed by the mismatch negativity. NeuroImage (in press).

Rinne, T., Alho, K., Ilmoniemi, R.J., Virtanen, J., Naatanen, R., 2000. Separate time behaviors of the temporal and frontal mismatch negativity sources. NeuroImage 12, 14–19.

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

Sekihara, K., Nagarajan, S.S., Poeppel, D., Marantz, A., Miyashita, Y., 2002. Application of an MEG eigenspace beamformer to reconstructing spatio-temporal activities of neural sources. Hum. Brain Mapp. 15, 199–215.

Singh, K.D., Barnes, G.R., Hillebrand, A., Forde, E.M., Williams, A.L., 2002. Task-related changes in cortical synchronization are spatially coincident with the hemodynamic response. NeuroImage 16, 103–114.

Tarantola, A., 1994. Inverse Problem Theory, second ed. Elsevier, Amsterdam.

Tikhonov, A.N., 1963. Solution of incorrectly formulated problems and the regularization method. Sov. Math., Dokl. 4, 1035–1038.

Uutela, K., Hamalainen, M., Somersalo, E., 1999. Visualization of magnetoencephalographic data using minimum current estimates. NeuroImage 10, 173–180.

van Burik, M., Pfurtscheller, G., 1999. Functional imaging of postmovement beta event-related synchronization. J. Clin. Neurophysiol. 16, 383–390.

Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880.

von Helmholtz, H., 1853. Über einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern, mit Anwendung auf die thierisch-elektrischen Versuche. Annals of Physics and Chemistry 89, 211–233, 353–377.

Vrba, J., Robinson, S.E., 2001. Signal processing in magnetoencephalography. Methods 25, 249–271.

Wagner, M., Fuchs, M., Wischmann, H.A., Ottenberg, K., Doessel, O., 1995. Cortex segmentation from 3D MR images for MEG reconstructions. In: Baumgartner, C., et al. (Eds.), Biomagnetism: Fundamental Research and Clinical Applications. Elsevier/IOS Press, Amsterdam, pp. 433–438.

Wagner, M., Fuchs, M., Wischmann, H.A., Drenckhahn, R., Koehler, T., 1996. Smooth reconstruction of cortical sources from EEG or MEG recordings. NeuroImage 3, S168.