Multiple Dipole Modeling and Localization from Spatio-Temporal MEG Data

John C. Mosher, *Member, IEEE*, Paul S. Lewis, *Senior Member, IEEE*, and Richard M. Leahy, *Member, IEEE*

Abstract—An array of biomagnetometers may be used to measure the spatio-temporal neuromagnetic field or magnetoencephalogram (MEG) produced by neural activity in the brain. A popular model for the neural activity produced in response to a given sensory stimulus is a set of current dipoles, where each dipole represents the primary current associated with the combined activation of a large numer of neurons located in a small volume of the brain. An important problem in the interpretation of MEG data from evoked response experiments is the localization of these neural current dipoles. We present here a linear algebraic framework for three common spatio-temporal dipole models: i) unconstrained dipoles, ii) dipoles with a fixed location, and iii) dipoles with a fixed orientation and location. In all cases, we assume that the location, orientation, and magnitude of the dipoles are unknown. With a common model, we show how the parameter estimation problem may be decomposed into the estimation of the time invariant parameters using nonlinear least-squares minimization, followed by linear estimation of the associated time varying parameters. A subspace formulation is presented and used to derive a suboptimal least-squares subspace scanning method. The resulting algorithm is a special case of the well-known MUltiple SIgnal Classification (MUSIC) method, in which the solution (multiple dipole locations) is found by scanning potential locations using a simple one dipole model. Principal components analysis (PCA) dipole fitting has also been used to individually fit single dipoles in a multiple dipole problem. Analysis is presented here to show why PCA dipole fitting will fail in general, whereas the subspace method presented here will generally succeed. Numerically efficient means of calculating the cost functions are presented, and problems of model order selection and missing moments are discussed. Results from a simulation and a somatosensory experiment are presented.

I. Introduction

A N array of Superconducting QUantum Interference Device (SQUID) biomagnetometers may be used to measure the spatio-temporal magnetoencephalogram

(MEG) produced by the brain. Given these external magnetic field measurements, one would like to compute a "neuromagnetic image," specifying the three-dimensional current density that produced the magnetic field. Accomplishing this requires inversion of the Biot-Savart law. Unfortunately, this general inverse problem is ill-posed because different neural current distributions may produce the same external field measurements. As a result, physical models of the underlying current distributions are employed (see [1] for an overview).

Of particular interest is the localization of the neural currents evoked in response to a given sensory stimulus, such as auditory or visual. An external magnetic field may be produced by this *primary* neural current, for which the simplest and most widely used composite model is the "dipole in a sphere." Here, the primary current is modeled as a *current dipole* or set of dipoles, and the head is modeled as a conductive sphere. A current dipole can accurately model neural activity localized to one site, representing the coherent activation of a large number of individual neurons [2]. The *return* or global *volume* currents are distributed over the sphere of the head, and the external magnetic field generated by the volume currents has no component normal to the head surface [3]-[5]. Thus, in this simple model, the component of the magnetic field oriented radially from the head is produced by the primary dipole current alone.

Given a suitable source and head model, the inverse problem can be reduced to the nonlinear optimization problem of computing the location and moment parameters of the set of dipoles whose field best matches the MEG measurements in a least-squares sense. Singh *et al.* [6] discussed neuromagnetic imaging (NMI), which extends the dipole model by assuming a large set of current dipoles, each with a fixed location at the center of a voxel (volume element) of a three-dimensional volume within the brain. The MEG data and image are linearly related; however, because of the large number of unknowns in the three-dimensional image, there is generally not a unique solution. Ilmoniemi *et al.* [4] described the general forward problem and presented minimum norm inverse solutions. Jeffs *et al.* [7] investigated several cost functions to select an appropriate solution from the set of feasible solutions. Dallas [8] investigated the imaging problem using a direct Fourier-based inversion approach, Alvarez [9] recently presented Fourier-based solutions for the two-dimensional case, and Wikswo *et al.* [10] have had success in imaging two-dimensional objects with their

MicroSQUID apparatus. In this paper we present methods for solving for a small parsimonious set of dipoles as a means of avoiding the ill-posedness associated with the full three-dimensional image model.

As in all modeling situations, a tradeoff exists between model complexity and generality and the ability to estimate reliably the model parameters from the given measurement data. Initial MEG dipole models used single time "snapshots" of the measured spatial magnetic field, where a spatial dipole model was fitted at an instance in time, usually at a local or global response peak. These MEG models are direct counterparts of EEG dipole models, known as *instantaneous state* dipole models [11]. To increase the complexity of the source models that can be effectively employed, researchers have begun to incorporate temporal modeling assumptions. The addition of a temporal model increases the range of the measurements that can be used in model fitting. A *spatio-temporal* dipole model and the necessary associated assumptions are presented in detail in [2], [12].

The spatio-temporal models differ in the manner in which they describe the time dependence of the data. Scherg and von Cramon [2], [13] use dipoles fixed in an unknown location *and* orientation, and therefore the time dependence is represented by a scalar time series specifying the magnitude and polarity of the current flow. Maier *et al.* [14] implicitly assume the same model, but use principal components analysis (PCA) to derive the locations. Achim *et al.* [15] compare the instantaneous state dipole model with PCA dipole fitting and spatio-temporal modeling; in their three-dipole fixed orientation spatio-temporal model, they assume that two of the dipoles are known in orientation and location (information obtained from the instantaneous state dipole model), and thus search for only one unknown dipole location and orientation. In [16], Scherg and von Cramon have fixed location, but unconstrained orientation, and they introduce the idea of *dipole source potentials*, or *regional dipoles* [12], where three *elemental dipoles* with orientations in orthogonal directions may occupy the same physical location.

These differences in formulation can be grouped into three spatio-temporal dipole models: i) unconstrained ("moving and rotating") dipoles, ii) dipoles with a fixed location ("rotating" or "regional"), and iii) dipoles with a fixed location and a fixed orientation ("fixed"). We assume that the parameters of location, orientation, and magnitude are all unknown. Our intent here is not to argue the merits of one model over another, but rather to show how each model may be solved efficiently within a common linear algebraic framework. In each case, we show that the model can be mathematically reduced to the same general expression, allowing the same approach to finding the inverse solution. This expression will unify our approach to solving the inverse problem, no matter which model we choose. Preliminary versions of this work appeared in [17], [18].

In Section II, we present each of the three data models, and in Section III, we discuss the calculation of the error function in fitting these models to spatio-temporal data. In Section III-C, we present a computer simulation to illustrate the performance of these spatio-temporal modeling techniques. Section IV presents a new suboptimal but faster method of solving the least-squares problem using a *subspace scanning* approach. In Section V, we discuss how this subspace scanning is statistically equivalent to the MUltiple SIgnal Classification method (MUSIC) [19]. We also present analysis to show that although PCA dipole fitting is similar to these subspace methods, PCA will generally fail in the multiple dipole case, whereas these methods will generally succeed. We present a simulation of this scanning method, using the same example data from Section III-C. We then present results from an actual somatosensory experiment.

II. FORWARD MODELS

As discussed in the introduction, physical models are used to represent both the neural current sources and the enclosing head shape and conductivity. Source models range from simple current dipoles to complex current surfaces. Head shape and conductivity models range from spherically symmetric conductors to finite element models based on individual anatomy. The combination of source and head model is known as the *forward model*. Given any arbitrary static current distribution, the magnetic field can be obtained from the *Biot–Savart law*. For source dipoles in a spherically symmetric conductor, Ilmoniemi *et al.* [4] point out that the source model can be reduced to consideration of just the primary source elements, regardless of the orientation of the sensors outside of the sphere. If the measurements are restricted to the radial orientation, the model simplifies further [4], [5].

Although the degree of complexity can vary greatly, in all cases the forward model specifies the complete current distribution within the head via a set of *source parameters* (e.g., locations, magnitudes). In other words the forward model provides a mapping from the source parameters to the resulting magnetic fields. In general, we can partition the parameters of any forward model into linear and nonlinear parameters. In this section, we show how this partition of spatio-temporal models provides a convenient algebraic form, a form which highlights the difficult nonlinear parameters. Although our development utilizes a specific forward model—current dipoles in a spherically symmetric conductor—the basic approach can be applied to any forward model.

A. Biot–Savart Law

In this section, we present the Biot–Savart law in a convenient discrete matrix notation for discrete source elements, which we then use to develop the spatio-temporal dipole models commonly used in MEG research. The general model in every case requires determining the unknown set of parameters $\{L, M, S\}$. The linear time varying parameters, S, can always be found using a direct pseudoinverse solution, but, in general, the time invariant

location parameters, $\underline{L}$, must be found using an iterative nonlinear minimization algorithm. The block diagonal matrix, M, depends on whether we use the rotating or fixed dipole model; in the rotating model, M is simply an identity matrix, and in the fixed model, M contains the unit orientation parameters, $\hat{M}$. The goal here is to show that each model can be expressed in a common framework and solved in a similarly efficient manner.

We begin by examining the model for a single dipole, then expand this model to account for multiple dipoles. Establishing an origin, denoting the dipole position as $\vec{L}$, and observing the ith measurement at sensor location $\vec{R}(i)$, we can write the Biot–Savart law for a current dipole as

$$\vec{B}(i) = k \frac{\vec{Q} \times (\vec{R}(i) - \vec{L})}{|\vec{R}(i) - \vec{L}|^3} \tag{1}$$

where $k = \mu_0/(4\pi)$ is a constant, $\vec{Q}$ is the dipole moment, $\vec{L}$ is the dipole location, $\vec{R}(i)$ is the ith measurement sensor location, and $\vec{B}(i)$ is the magnetic field at $\vec{R}(i)$. A SQUID biomagnetometer is used to acquire the magnetic field at position $\vec{R}(i)$, but it measures only one component of the three-dimensional field. Thus, only a scalar measurement is made:

$$B(i) = \vec{B}(i) \cdot \hat{s}(i) \tag{2}$$

where $\hat{s}(i)$ denotes the unit orientation of the ith sensor. The operation "$\cdot$" denotes the dot product of two vectors.

Combining equations (1) and (2) yields

$$B(i) = k \frac{(\vec{R}(i) - \vec{L}) \times \hat{s}(i) \cdot \vec{Q}}{|\vec{R}(i) - \vec{L}|^3} = \vec{g}(i) \cdot \vec{Q}. \tag{3}$$

The vector $\vec{g}(i)$ can be viewed as a gain vector, relating the moment intensity of the dipole to the measurement at position $\vec{R}(i)$. If we let each gain vector be represented as a 1×3 row vector and the moment as a 3×1 column vector, then we can arrange the measurements from m locations in a matrix form:

$$\underline{B} = \begin{bmatrix} B(1) \\ \vdots \\ B(m) \end{bmatrix} = k \begin{bmatrix} \dfrac{(\vec{R}(1) - \vec{L}) \times \hat{s}(1)}{|\vec{R}(1) - \vec{L}|^3} \\ \vdots \\ \dfrac{(\vec{R}(m) - \vec{L}) \times \hat{s}(m)}{|\vec{R}(m) - \vec{L}|^3} \end{bmatrix}$$

$$\cdot [\vec{Q}] = \begin{bmatrix} \vec{g}(1) \\ \vdots \\ \vec{g}(m) \end{bmatrix} [\vec{Q}] = G(\vec{L})\vec{Q} \tag{4}$$

The matrix $G(\vec{L})$ can be considered to be the gain or relationship between a unit moment source at $\vec{L}$ and the column vector of measurement locations $\{R(i)\}$. From this form we clearly see the linear relationship between

the moment $\vec{Q}$ and the measurements $\underline{B}$. As we will show in Section III, this form also focuses our attention on the more difficult nonlinear parameters in the matrix G (we usually drop the dependence on $\vec{L}$ for notational simplicity).

Although derived for a single dipole, each column in G could also be viewed as the model for three elemental dipoles or *dipole source components*, with all three sharing the same location, but in oblique directions [16]. For simplicity, we will continue referring to these collocated elemental dipoles as one dipole, with moment $\vec{Q}$. This model easily extends to the multiple dipole (not collocated) case by superposition. For p dipoles,

$$\underline{B} = [G_1, \cdots, G_p] \begin{bmatrix} \vec{Q}_1 \\ \vdots \\ \vec{Q}_p \end{bmatrix} \tag{5}$$

or simply $\underline{B} = G(\underline{L})\,\underline{T}$, where $G(\underline{L})$ can be partitioned into the smaller matrices $G_i(\vec{L}_i)$, as defined in (4). Similarly, $\underline{T}$ may be partitioned as the concatenation of the moment vectors for each of the p dipoles. Notationally, we use over arrows, $\vec{\cdot}$, to indicate a row or column vector that is geometric in three-dimensions, such as a location, $\vec{L}_i$. We use underlines and boldface to indicate more general column vectors and matrices, which in turn may contain many such geometric vectors. For m sensors and p dipoles, vector $\underline{B}$ is $m \times 1$, matrix G is $m \times 3p$, and vector $\underline{T}$ is $3p \times 1$.

B. Dipole in a Sphere

Our model in (5) describes the external magnetic field as the sum of the individual fields from p dipoles. For simplicity in deriving the model, the biomagnetometer is assumed to make a perfect point field measurement. We also assume that this field is due to the local primary current only, as we are ignoring the global volume or return currents. In more sophisticated head and source models, the return currents, the finite coil area, and the gradiometer configuration of a practical SQUID biomagnetometer could also be included, resulting in a very similar formulation to that presented here [4], [5], [7].

Many researchers have correctly noted that only the two tangential dipole moment components need be computed for the "dipole in a sphere" model. A radially oriented dipole inside a spherically symmetric conductor produces no external magnetic field, since the field from the surface return currents cancels the field from the primary dipole current [3], [4], [5], [20], [21]. Additionally, the magnetic field normal to the surface of the sphere (i.e., radial from the center of the sphere) is due solely to the primary *tangential* dipole currents; volume or return currents in the surface of the sphere contribute nothing to the radial magnetic field.

A common MEG geometry is therefore an array of sensors arranged radially about the surface of the head, which is sensitive only to the tangentially oriented primary di-

pole currents. In this paper we will therefore assume that the radial component is immeasurable and that only the two tangential components are measurable. Thus, each submatrix G_i will be $m \times 2$, corresponding to the gain in the two tangential directions. We can refer to these two tangential components as ϕ and θ, e.g., $G_i = [\underline{G}_\phi \; \underline{G}_\theta]$. We emphasize that although we are using the dipole in a sphere model as an example, all results are sufficiently general such that extensions to other models containing all three moment components are straightforward. Indeed, the methods presented here can be applied to the EEG models as well, but we will restrict ourselves to the MEG model for simplicity.

C. Unconstrained Location and Orientation

The *instantaneous state dipole* [11] is the simplest of the dipole models, for which we consider just a single time slice of data, typically at the peak of the observed MEG response. The straightforward extension to the full temporal information is simply to treat each time slice with a separate static model. At each time slice, the locations and moments are calculated for each dipole, independently of all other time slices. Since no constraints are placed on the parameters of the dipole, this model allows both moving and rotating dipoles. The model is simply (5) with a time parameter j inserted, $\underline{B}(j) = G(j)\underline{T}(j)$, which we would solve for each j, $j = 1, \cdots, n$.

D. Fixed Location, Unconstrained Orientation

When the instantaneous dipole model is solved for several sequential time points, the location of the dipole can appear to move as a function of time. Many researchers believe it is more realistic to assume that different parts of the cortex with different cortical function are activated electrically when they perform their specific tasks [22], [23]. The "movement" seen in the instantaneous state dipole would therefore be more accurately modeled as two or more stationary dipoles which are activating electrically at different times. Rather than allowing the dipole locations to vary with time, as in the above model, this second model restricts the location of the dipoles to be constant throughout the measurement interval, but allows the moment intensities and orientations to vary. We begin with the previous model, fix the gain matrix to be a constant with respect to time, and represent the model in a compact matrix form:

$$[\underline{B}(1), \cdots, \underline{B}(n)] = G[\underline{T}(1), \cdots, \underline{T}(n)] \qquad (6)$$

or $B = GT$. Each column of the T matrix may be partitioned to represent the moments of p dipoles at time j,

$$\underline{T}(j) = \begin{bmatrix} \vec{Q}_1(j) \\ \vdots \\ \vec{Q}_p(j) \end{bmatrix} = \begin{bmatrix} \hat{M}_1(j)S_1(j) \\ \vdots \\ \hat{M}_p(j)S_p(j) \end{bmatrix} \qquad (7)$$

where each partition $\vec{Q}_i(j)$ can be represented by its unit moment orientation $\hat{M}_i(j)$ and intensity $S_i(j)$. Hence,

each row of T can represent the time series for one component of one tangential dipole [16].

E. Fixed Location and Orientation

Since no constraints are placed on the time series of the three components for each dipole, the orientation of the dipole can vary or "rotate" over time. Some researchers [2], [13] argue that physiologically a dipole orientation should not rotate, because the dipole model represents a fixed neuroanatomical structure. If we fix the unit orientation of each moment to be the same for all time slices and allow only the magnitude and polarity, $S(j)$, of the moment to vary, then we can express the matrix T from above as

$$T = \begin{bmatrix} [\hat{M}_1][S_1(1) \cdots S_1(n)] \\ \vdots \\ [\hat{M}_p][S_p(1) \cdots S_p(n)] \end{bmatrix} \qquad (8)$$

$$= \begin{bmatrix} \hat{M}_1 & & 0 \\ & \ddots & \\ 0 & & \hat{M}_p \end{bmatrix} \begin{bmatrix} S_1(1) \cdots S_1(n) \\ \vdots \ddots \vdots \\ S_p(1) \cdots S_p(n) \end{bmatrix} \qquad (9)$$

$$= MS. \qquad (10)$$

Thus, our constrained model is now separated into three components:

$$\begin{bmatrix} B(1, 1) & \cdots & B(1, n) \\ \vdots & \ddots & \vdots \\ B(m, 1) & \cdots & B(m, n) \end{bmatrix}$$

$$= [G_1, \cdots, G_p] \begin{bmatrix} \hat{M}_1 & & 0 \\ & \ddots & \\ 0 & & \hat{M}_p \end{bmatrix} \begin{bmatrix} S_1(1) \cdots S_1(n) \\ \vdots \ddots \vdots \\ S_p(1) \cdots S_p(n) \end{bmatrix}$$

$$(11)$$

or $B = GMS$. As before, G represents the $m \times 2p$ gain matrix between p unit dipoles and the array of m sensors. The $2p \times p$ block diagonal matrix M represents the fixed unit orientation moments. The moment intensity matrix S is $p \times n$.

Grouping as $B = (GM)S = AS$, we retain the same general separation into two matrices as for the other models. The difference here is that we have both the time invariant location $\underline{L}$ and moment orientation parameters M in the first matrix A, rather than just the location as before. Each column of G represents the gain of one *component* of one dipole, but each column of $A = [\underline{A}_1, \cdots, \underline{A}_p] = GM$ now represents one *complete* dipole.

F. General Model

The most general model contains both rotating and fixed dipoles. We might argue that two fixed dipoles may be so closely located that they appear in our data as one rotating dipole. We may also have rotating dipoles that rotate so

little as to appear fixed in orientation. The general model accounts for both types and will be useful in deriving further results in this paper.

We simply alter our definition of a rotating dipole to be one which *must* rotate, such that its two component time series cannot be partitioned as a fixed moment orientation and a scalar time series (i.e., the time series are linearly independent). For p_r rotating dipoles and p_f fixed dipoles, we partition the pairs of rows of T as the rank two submatrix $T_r = [T_\phi T_\theta]^T$ for rotating dipoles and the rank one submatrix $T_f = [\hat{M}S^T]$ for fixed:

$$
T = \begin{bmatrix} [\underline{T}_{\phi 1}\underline{T}_{\theta 1}]^T \\ \vdots \\ [\underline{T}_{\phi p_r}\underline{T}_{\theta p_r}]^T \\ [\hat{M}_{(p_r+1)}\underline{S}^T_{(p_r+1)}] \\ \vdots \\ [\hat{M}_p\underline{S}^T_p] \end{bmatrix}
$$

$$
= \begin{bmatrix} I_{2p_r} & & & 0 \\ & \hat{M}_{(p_r+1)} & & \\ & & \ddots & \\ 0 & & & \hat{M}_p \end{bmatrix} \begin{bmatrix} [\underline{T}_{\phi 1}\underline{T}_{\theta 1}]^T \\ \vdots \\ [\underline{T}_{\phi p_r}\underline{T}_{\theta p_r}]^T \\ \underline{S}^T_{(p_r+1)} \\ \vdots \\ \underline{S}^T_p \end{bmatrix} = MS
$$

(12)

where I_{2p_r} is a $2p_r \times 2p_r$ identity matrix. We can now express our model as $B = GT = G(MS) = HS$, where $H = GM$ is our "hybrid" gain matrix, a combination of the previously defined G_i and $\underline{A}_i$ submatrices,

$$
GM = H = [G_1, \cdots, G_{p_r}, \underline{A}_1, \cdots, \underline{A}_{p_f}]. \quad (13)
$$

The rank of this matrix H is $r = 2p_r + p_f$, which is effectively the number of *dipole source components* [16] in our model. The matrix S is the corresponding time series for *each* dipole component; if two dipole components are collocated, then by our definition they represent one rotating dipole.

The rotating and fixed models are just specializations of this model, and the instantaneous dipole model is just this model for one time slice. For no fixed dipoles ($p_f = 0$), M is simply a $2p_r \times 2p_r$ identity matrix, and we have the rotating dipole model of (6), with $3p$ unknown location parameters in the gain matrix $H(\underline{L}, M) = G(\underline{L})$. Similarly, for no rotating dipoles ($p_r = 0$), we have the fixed dipole model of (11), with $4p$ unknown location and constrained unit moment parameters in the gain matrix $H(\underline{L}, M) = A(\underline{L}, M)$. As we will show in Section V, successful localization requires that we determine the rank r, i.e., the number of dipole components, but not necessarily the number of rotating and fixed dipoles.

III. THE LEAST-SQUARES SOLUTION

In this section, we show how a least squares error is generated by implicitly or explicitly projecting the measurements onto the orthogonal complement of the subspace formed from the appropriate forward model. We use the simple dipole in a sphere model as an illustration, where the nonlinear parameters are the locations of the dipoles and possibly the fixed orientation. The cost function is shown to be a function of these nonlinear parameters only, thus reducing the number of parameters to be searched. The nonlinear parameters are iteratively adjusted to minimize this error through standard nonlinear minimization techniques. The optimal nonlinear parameter set is then used to perform a linear least squares fit for the linear parameters. We present approaches for efficiently calculating these cost functions, then conclude with a three dipole simulation example.

A. Separation of Linear Parameters

Consider the general model of p_r rotating dipoles and p_f fixed dipoles for our data, $B = H(\underline{L}, M)S$, where $H(\underline{L}, M)$ is $m \times r$, S is $r \times n$, and $r = 2p_r + p_f$. $H(\underline{L}, M)$ has $3p$ unknown location parameters and p_f unknown constrained moment parameters. We collect n time samples from each of m SQUID biomagnetometers and form a spatio-temporal matrix of data $F = [\underline{F}(1) \cdots \underline{F}(n)]$. We model this data as $F - B = N$, where N represents the error between the measurements and our model. We define our measure of fit as the square of the Frobenius norm,

$$
J_{LS} = \sum_{i=1}^{m} \sum_{j=1}^{n} \{F(i, j) - B(i, j)\}^2 = \|F - B\|_F^2
$$

$$
= \|F - H(\underline{L}, M)S\|_F^2. \quad (14)
$$

The goal is to find the set $\{\underline{L}, M, S\}$ that minimizes this error.

The simple approach is to use this cost function directly in an iterative minimization algorithm, which would search for all five parameters for each one of p dipoles at each point of n time instances, for an overall total of $5pn$ parameters [24]. Thus a three dipole model would require searching a 15 parameter space at every time instance. This same cost function, however, can be viewed as a projection minimization that can greatly reduce the computational cost, yet incorporates the information of all the time slices.

Given $\underline{L}$ and M (hence, H), a solution for the matrix S that will minimize J_{LS} is

$$
S = H^\dagger F \quad (15)
$$

where $H^\dagger$ is the well-known pseudoinverse solution (for full column rank H),

$$
H^\dagger = (H^T H)^{-1} H^T \quad (16)
$$

or the more general Moore–Penrose pseudoinverse (minimum norm) solution,

$$
H^\dagger = V\Sigma^+ U^T, \quad (17)
$$

where $H = U\Sigma V^T$ is a singular value decomposition (SVD) and Σ^+ is the diagonal matrix Σ with its nonzero elements inverted [25]. This minimization for S will hold for *all* sets $\{\underline{L}, M\}$, including the optimal set $\{\underline{L}, M\}^*$ that minimizes the cost function J_{LS}.

We can replace S with this pseudoinverse solution before solving for $\{\underline{L}, M\}$. The cost function can be equivalently expressed as

$$J_{LS} = \|F - HS\|_F^2 = \|F - H(H^\dagger F)\|_F^2$$
$$= \|(I - HH^\dagger)F\|_F^2 = \|(I - P_H)F\|_F^2 = \|P_H^\perp F\|_F^2.$$

$$(18)$$

The matrices P_H and $P_H^\perp$ are *projection* matrices: P_H projects data onto the column space of the matrix H, and $P_H^\perp$ is the orthogonal complement projection, that is, the projection of the data onto the left null space of H. Thus, the squared error can be explicitly computed from the projection of the data matrix, F, onto the left null space.

We have used the separation of the unknown parameters into linear and nonlinear components and have factored out the linear moments. While this method has often been used by other researchers [12], [14], [15], [23], the mathematical details have not always been stated or explained explicitly; [26], [27] give a full mathematical justification for this approach. The benefit is that J_{LS} is now a function of only the parameters in H. An iterative minimization routine need only explicitly consider this reduced subset of parameters, which can considerably reduce the convergence time.

This cost function can be minimized directly by a non gradient-based method, such as the Nelder–Meade simplex, used by many researchers because of its simplicity and apparent robustness to local minima [15]. Alternatively, gradient-based methods are typically faster, but require either analytical or numerical partial derivatives of the projection matrix. Simple analytic expressions of the partials of the projection matrix are derived in [26], [27], and these expressions, in turn, require only partials of the gain matrix H. In practice, however, any slight change in the head or source model requires a recalculation of the partials, and the simplex method or other nongradient methods are therefore usually preferred for their simplicity [14], [15].

The computational complexity of the least-squares estimation problem is highly dependent on the number of nonlinear parameters that must be estimated. In the instantaneous state dipole model, the location of each dipole must be computed independently for each time slice i. In comparison, the number of nonlinear parameters in the rotating dipole model drops dramatically. For p dipoles and n time slices, in both cases we have $2pn$ linear parameters, but the gain matrix G is a function of only $3p$ location parameters for the rotating dipole model, rather than $3pn$ for the instantaneous dipole model. Thus, the rotating dipole model requires an iterative search of only the $3p$ nonlinear location parameters, followed by a simple $2pn$ linear fit for the moment parameters. For the fixed

dipole model, the dimensionality of the search space in the iterative minimization algorithm is increased relative to the rotating dipole model from $3p$ parameters (locations only) to $4p$ parameters (locations and constrained unit orientations). Only the pn parameters of the time-series magnitude and polarity of the moments in the matrix S can be calculated using a simple linear fit.

An approximate approach to the fixed dipole model is to group the model as $B = G(MS)$, which is effectively identical to the rotating dipole model $B = GT$. Once T is found, we form a second equation, $MS = T$, from (10), and solve for M and S. The advantage of this approach is that $MS = T$ can be solved efficiently using the SVD. Each set of two rows of T represents the time series for the two tangential components of one dipole. If the dipole is truly fixed in orientation, then this $2 \times n$ matrix partition is of rank one. An SVD of this partition of the matrix will give the best rank one fit, and a simple analysis of the singular values will confirm the quality of this fit. If the rank one fit from the SVD is perfect, then the solution is optimal; however, in general this approximate approach is not guaranteed to give the same result as that which would be obtained by solving $B = AS$. The appeal lies in keeping only $3p$ unconstrained parameters instead of $4p$ constrained parameters in the iterative search space. An example of this method is shown in Section III-C. One possible extension of this approach may be to embed this two-step process into each iterative error calculation, rather than solving $MS = T$ only once at the end of the iterations.

B. Error Function Computation

The greatest computational burden in fitting the multiple dipole model occurs in the iterative nonlinear minimization routine, which must repeatedly form the nonlinear gain matrix H, then solve the inverse matrix problem for cost function J_{LS} for different sets of parameters. By using the SVD and the QR decomposition [25], we can significantly reduce the number of multiplications required. This analysis also leads naturally into Section IV, where we present an alternative method of solving the least-squares problem using these same decompositions.

If the number of time samples, n, is greater than the number of sensors, m, then the use of an SVD of F gives an efficient form for calculating the above error function. Decomposing as $F = U\Sigma V^T$, the least-squares cost function can be rewritten as

$$J_{LS} = \|P_H^\perp F\|_F^2 = \|P_H^\perp U\Sigma V^T\|_F^2$$
$$= \|P_H^\perp U\Sigma\|_F^2 = \|P_H^\perp W\|_F^2$$

$$(19)$$

where we are able to drop the term V^T because orthogonal matrices preserve the F-norm. Since F is $m \times n$, then the diagonal matrix Σ of singular values has at most m nonzero terms. Therefore $W = U\Sigma$ is only $m \times m$, as opposed to the larger $m \times n$ data matrix F.

Greater savings occur if the number of nonzero singular terms, r, is less than m, i.e., F is not of full row rank,

because then W would have correspondingly fewer columns, r. With noise considerations, the singular values are almost always strictly greater than zero, so we can amend r to be the number of singular terms "significantly" greater than zero. The corresponding r components in the decomposition are the *principal components* that "adequately" describe the data, where the user must decide what is adequate and what is significant. Using these significant components reduces the computational cost but results in a suboptimal least-squares solution. Section IV more fully exploits this possible decomposition. A two-stage approach would be to use this reduced r set for the coarse fit, then return to the full m set of components for the complete fit.

If an SVD is used to calculate the pseudoinverse of H, then a further reduction in the number of multiplications can occur. Denoting the decomposition as $H = [U_r U_{m-r}]\Sigma V^T$, where H is $m \times r$, then let U_r contain the columns corresponding to the r non-zero singular values, and let U_{m-r} correspond to the $m - r$ zero singular values, where r is the rank of H. Then $P_H^{\perp} = U_{m-r}U_{m-r}^T$, and the cost function can be calculated as

$$J_{LS} = \|P_H^{\perp} W\|_F^2 = \|U_{m-r}^T W\|_F^2 \qquad (20)$$

$$= \|W\|_F^2 - \|P_H W\|_F^2 = \|W\|_F^2 - \|U_r^T W\|_F^2 \qquad (21)$$

where the second form follows from the well-known equality $\|W\|_F^2 = \|P_H W\|_F^2 + \|P_H^{\perp} W\|_F^2$.

The selection between the two forms depends on the rank r and the decomposition method used, since either projection matrix may also be efficiently computed by using the QR decomposition of the gain matrix H. The choice of SVD or QR decomposition is application dependent. In general, the calculation of an SVD is more expensive, because it calculates the two eigenvector spaces U and V by iteratively converging to a solution. However, "economy" SVD versions [28], [29] can be run in which only the principal component eigenvectors are calculated, i.e., U_r above, resulting in considerable savings if the rank r of H is small relative to m. By comparison, QR decomposition is noniterative, and it outperforms a full SVD calculation. If the rank of H is large, then QR decomposition generally outperforms even the economy SVD. Either method of decomposing H will outperform the undecomposed projection matrix, $P_H = HH^{\dagger}$. These decomposition approaches also have better numerical properties than inverting $H^T H$ or solving via Gaussian elimination [25]. We will use these decompositions as a basis for forming a subspace scanning method in Section IV.

C. Simulation

We conclude this section by presenting the results of simulations in which the spatio-temporal models described are applied to a set of simulated MEG data. Data were simulated for the case of three dipoles, all three with fixed locations, two with fixed moment orientations and the third with a rotating orientation. The data were com-
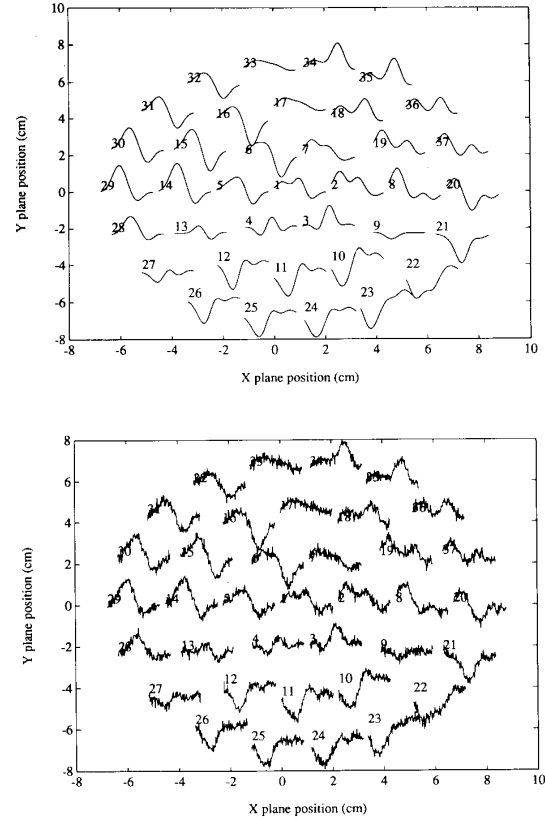

Fig. 1. Simulated MEG data for 37 radially oriented sensors, 100 time samples from each. Sensors are positioned on an imaginary sphere of radius 12 cm, centered on the origin, with Sensor 1 located 12 cm above the (x, y) plane at (x, y, z) position (0, 0, 12). Three dipoles were simulated about 3 cm radially below the sensors numbered 9, 13, and 17. Zero-mean Gaussian noise with a standard deviation equal to one-tenth the peak was added, for an SNR of approximately 10 dB. Each numbered trace represents the time response for the corresponding sensor, with the plots arranged in their approximate spatial position in the (x, y) plane.

puted for an array of 37 closely spaced sensors radially oriented and positioned on the surface of an imaginary sphere of radius 12 cm.

Because the head model was assumed to be a spherically symmetric conductor and the sensors were arranged radially outside of the sphere, then only the fields due to the primary tangential dipole currents were computed, as discussed in Section II. However, rather than compute the parameters in a spherical or rotated coordinate system, we employ a Cartesian coordinate system and solve for the three constrained moment parameters per dipole, using the pseudoinverse form of (17). A total of 100 time samples were generated and corrupted by additive white Gaussian noise with an SNR of 10 dB. SNR is computed as the ratio of the average magnetic field measurement power to the variance of the noise. The resulting simulated MEG data are shown in Fig. 1.

The parameters for this data were estimated using two of the models discussed: a) fixed location but unconstrained orientation; and b) fixed location and orientation.

TABLE I

ESTIMATED LOCATIONS AND MOMENTS. THE LOCATIONS WERE FOUND USING A NELDER–MEADE SIMPLEX
ALGORITHM FOR THE ROTATING DIPOLE MODEL. THE NUMBER OF DIPOLES, 3, WAS ASSUMED KNOWN, AND
THE SIMPLEX ALGORITHM SEARCHED FOR THE BEST 9 LOCATION PARAMETERS THAT FIT THE DATA IN A
LEAST-SQUARES SENSE. ONCE THE LOCATIONS WERE OPTIMIZED, THE TIME SERIES WERE FOUND VIA A
SIMPLE LINEAR LEAST-SQUARES FIT. THE TIME SERIES FOR THE SECOND AND THIRD DIPOLES WERE
DECOMPOSED INTO A RANK 1 MODEL, FROM WHICH THE FIXED MOMENT ORIENTATIONS WERE FOUND

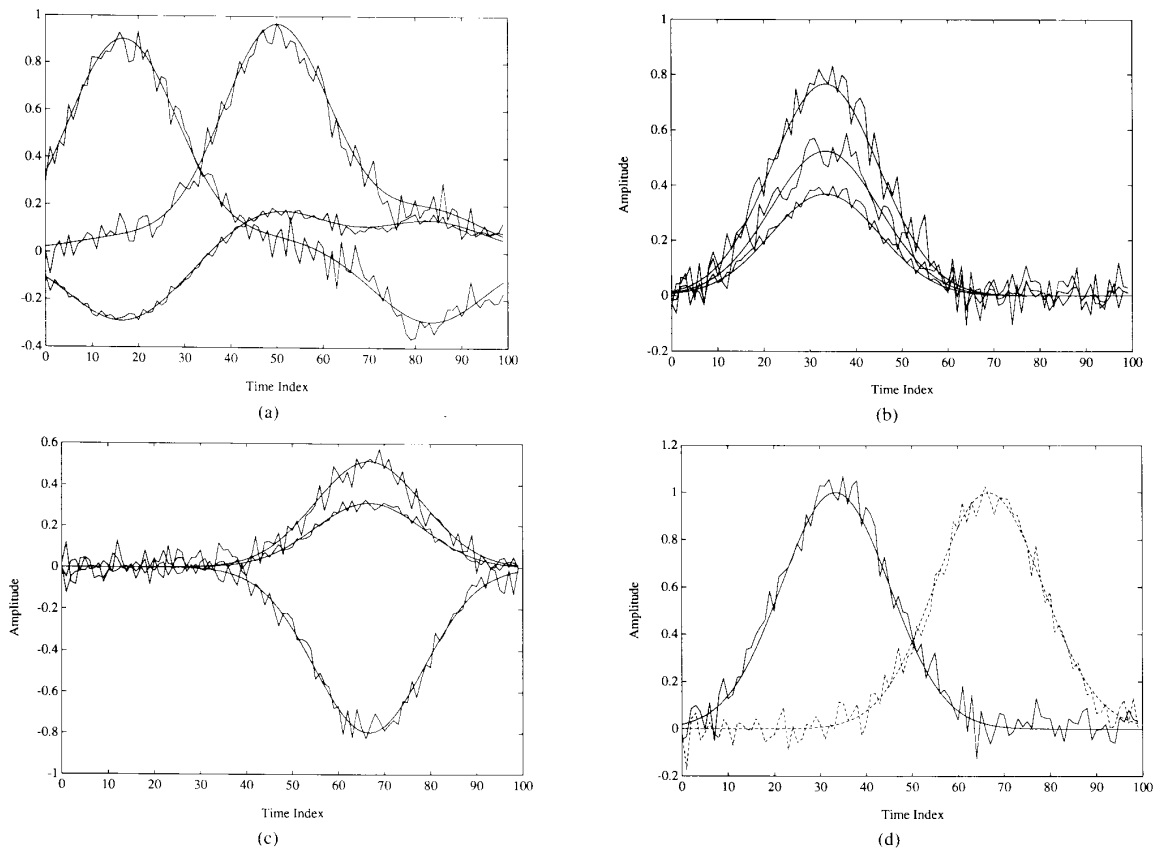|	Dipole 1			Dipole 2			Dipole 3		
	True and Estimated Location (cm)								
	L_x	L_y	L_z	L_x	L_y	L_z	L_x	L_y	L_z
True	2.800	-1.700	8.300	-2.900	-1.600	8.300	0.000	3.300	8.400
Est.	2.817	-1.691	8.335	-2.910	-1.594	8.225	-0.056	3.320	8.358
	True and Estimated Moments (from SVD)								
	M_x	M_y	M_z	M_x	M_y	M_z	M_x	M_y	M_z
True	(N/A, rotating)			0.770	0.525	0.369	0.516	-0.797	0.313
Est.				0.770	0.518	0.373	0.507	-0.800	0.320

(a)

(b)

(c)

(d)

Fig. 2. Estimated moment time series for three dipoles. Three dipoles of tangential orientation (no radial component) were
given overlapping time series and projected into cartesian coordinates, one time series per coordinate per dipole. The dipole
model for the unconstrained moment orientation was then used in a simplex algorithm to find the locations. The nine time series
were found with a simple least-squares fit and plotted for each of the three components of the three dipoles in (a)–(c). An SVD
was then performed on the second and third dipole time series [(b) and (c)] to approximate the dipole model for constrained
orientations, and the resulting dipole moment magnitude and polarity is plotted in (d).

For the "rotating" model, the locations of the three di-
poles were estimated using a Nelder–Meade simplex
search to minimize J_{LS} in (18) over the dipole location
parameters; (15) was then used to find the moment time
series. The true and estimated locations are listed in Table
I and the estimated time series are shown in Fig. 2 over-
layed with the original simulated time series.

For the fixed orientation and location model, rather than

iteratively search for dipole locations *and* orientations, we used the two-step approximate method discussed in Section III-A. In this method, we use the results of the rotating model above, which searched the nine parameter location space only. The identified time series were then fit to a rank one model (per dipole) via an SVD, resulting in the time series displayed in Fig. 2. The true and identified moments for the two fixed dipoles are displayed in Table I. Because the third dipole actually had a rotating moment, then the SVD of its time series properly revealed a poor rank one fit.

IV. A NEW APPROACH TO SOURCE LOCALIZATION

In this section we introduce a new MEG inverse algorithm. Although it is generally suboptimal in a least-squares sense, this algorithm has the strong advantage of "scanning" quickly with a *one* dipole search, rather than the p-dipole search necessary in a complete fit. In Section V-A, we discuss how this algorithm is analogous to the statistically derived MUSIC direction-finding algorithm for polarized sources [19]. Section V-C examines the related method of PCA dipole fitting, but shows where this method differs and why it generally fails. We first develop the algorithm for the rotating dipole, then extend it to include the fixed dipole model. We then consider the general hybrid case of both fixed and rotating dipoles.

A. Order Selection

Perhaps one of the greatest problems in MEG analysis is determining the number of dipoles: if too few are selected for any of the models, then the calculated dipoles are biased by the missing dipoles; if too many dipoles are specified, then spurious dipoles are introduced, which may be indiscernible from the true dipoles. Since the computational cost and numerical sensitivity of most iterative minimization increases dramatically with the number of parameters, then too many dipoles also adds needless computational burden.

Other authors have shown that examination of the spatial surface topography can be misleading if time series are overlapping and/or dipoles are placed such that one maximum potential cancels another [15], [20]. As an alternative, analysis of the dimensionality or rank of the data matrix F is often made in an attempt to determine the true number of dipoles. In general, for p diodes, the rank of the model data matrix B will be limited by

$$\text{Rank } (B) = \min (\text{Rank } (H), \text{Rank } (S)) \leq 2p. \quad (22)$$

The upper limit comes from either H or S, since each dipole moment component in S has a corresponding column in H, with at most two moment components per dipole. For the lower limit, the problems usually arise in S, where the matrix is of full row rank only if all time series are linearly independent. If all dipoles are rotating, then S is at most rank $2p$; however, if all dipoles are fixed in orientation, then S is at most rank p. If any combination of the spatially distinct dipoles have linearly dependent

time series (perfectly correlated), then the rank of S drops accordingly.

For the general model with p_r rotating and p_f fixed dipoles, the rank of H is $r = 2p_r + p_f$. As we will show, we do not explicitly need to know p_r and p_f, if the time series are sufficiently independent and the SNR sufficiently large.

B. Rotating Dipole Model

We consider first the case where all dipoles rotate, that is $H(\underline{L}, M) = G(\underline{L})M = G(\underline{L})$. From (18), we express the least-squares cost function as

$$J_{LS} = \|F - GT\|_F^2 = \|P_G^\perp F\|_F^2 \quad (23)$$

where G is $m \times 2p$, T is $2p \times n$, and M is the identity matrix, since all dipoles are assumed rotating. We can interpret the least-squares problem as trying to find the gain matrix G whose orthogonal subspace projector $P_G^\perp$ minimizes J_{LS}. Since G is of rank $2p$, then the orthogonal complement projector $P_G^\perp$ is of rank $m - 2p$.

In the method developed here, we first find the best orthogonal projector $P^\perp$, regardless of the gain matrix, and then find the gain matrix G that best fits this projector. The first step in this method is to minimize over all possible orthogonal projectors of rank $m - 2p$, for which we form the first cost function,

$$J_1 = \|P^\perp F\|_F^2. \quad (24)$$

Minimizing J_1 over all $P^\perp$ is equivalent to finding the best rank $2p$ projections of F. From [30, corollary 2.3], the best rank $2p$ approximation of F is formed from the first $2p$ components of the SVD. Hence, we decompose F as $F = U \Sigma V^T = [U_{2p} U_{m-2p}] \Sigma V^T$, where U_{2p} contains the $2p$ left singular vectors associated with the $2p$ largest singular values, and U_{m-2p} contains the remaining $m - 2p$ left singular vectors. The best rank $2p$ approximation of F is given then by $F_{2p} = (U_{2p} U_{2p}^T) F$. The best orthogonal projector is therefore

$$P^\perp = U_{m-2p} U_{m-2p}^T. \quad (25)$$

Once we have formed this best orthogonal projector $P^\perp$, then the second step is to find the gain matrix G most orthogonal to this projector. Orthogonality between $P^\perp$ and $G = [G_1 \cdots G_p]$ implies $P^\perp$ is orthogonal to each G_i, where G_i is the $m \times 2$ gain matrix for a single dipole. Since the ith term depends only on the location of the ith dipole, our scanning function for rotating dipoles, $J_r(i)$, is derived from this orthogonality,

$$J_r(i) = \frac{\|P^\perp G_i\|_F^2}{\|G_i\|_F^2} = \|U_{m-2p}^T \bar{G}_i\|_F^2, \quad (26)$$

where $\bar{G}_i = G_i / \|G_i\|_F$ is the normalized gain matrix. Normalization is necessary so that a small value for $J_r(i)$ is an indication of closeness to orthogonality and is not simply due to a relatively small gain.

If $P^\perp$ is a reasonable approximation to the optimal $P_G^\perp$, and a G exists such that $\|P^\perp G\|_F^2 \approx 0$, then each of the G_i submatrices will be orthogonal to $P^\perp$, and

$J_r(i) \approx 0$ when evaluated for each correct location (the conditions under which these approximations hold true will be discussed in Section V). Our scanning method therefore is to search over all possible *one*-dipole locations and at each location evaluate $J_r(i)$, looking for minima. The explicit steps will be presented in Section V. The general approach is to evaluate $J_r(i)$ over a fine three-dimensional grid, plot its *inverse*, and look for p sharp "spikes."

Provided the gain matrices are linearly independent over the scanning space, the scanning function will not display more than p spikes, since the existence of more, for instance $p + 1$ spikes, would suggest there are $m + 2$ orthogonal column vectors in an m-dimensional space, which is not possible. W may however find fewer than p spikes. Each evaluation of $J_r(i)$ actually represents the projection of the column vectors for *two* tangential components in G_i. If the dipole is *fixed* in orientation, then only a linear combination of the two vectors may be orthogonal, and $J_r(i)$ would not be a minimum at the dipole location. The next subsection extends this scanning method to accommodate these fixed orientation dipoles.

C. Fixed Orientation Dipoles

We now assume all dipoles are fixed, such that our model for the fixed orientation dipole, as shown in (11), is $B = [G(\underline{L})M]S = A(\underline{L}, M)S$, where S is the scalar time series and M is the matrix of fixed unit orientation moments. The gain matrix G is still $m \times 2p$, but M is $2p \times p$ and S is $p \times n$, where p is the number of dipoles. Thus, $A(\underline{L}, M)$ is an $m \times p$ matrix and the full least-squares cost function is $J_{LS} = \min_A \|P_A^\perp F\|_F^2$, where $P_A^\perp$ is an orthogonal projector onto an $m - p$ subspace, rather than an $m - 2p$ subspace as for $P_G^\perp$. We otherwise proceed identically as above, arriving at an analogous cost function for fixed dipole,

$$J_f(i) = \frac{\|U_{m-p}^T \underline{A}_i\|_2^2}{\|\underline{A}_i\|_2^2} = \|U_{m-p}^T \tilde{A}_i\|_2^2 \qquad (27)$$

where $\tilde{A}_i = \underline{A}_i / \|\underline{A}_i\|_2$ is the normalized gain matrix for fixed dipoles, U_{m-p} now contains the $m - p$ left singular vectors associated with the $m - p$ smallest singular values, and we note that for vectors the L_2 norm is equivalent to the Froebinius norm.

Note that $J_f(i)$ is now a function of four parameters. If $J_f(i)$ were evaluated in the same manner as the rotating dipole scanning function, then not only would we need to scan over all possible locations $\vec{L}$, but at each location we would also have to scan over all possible constrained unit orientations $\hat{M}$, thus extending our scan from three dimensions to four dimensions. However, we now show that we can instead continue to search over only the three-dimensional locations $\vec{L}$ and at each location implicitly select the *best* orientation $\hat{M}$ which minimizes our cost function. We achieve this by using the separability of $\underline{A}_i$ as $\underline{A}_i = G_i \hat{M}_i$ and applying the constraint $\hat{M}_i^T \hat{M}_i = 1$, since $\hat{M}_i$ is, by our definition, a unit moment orientation vector.

For a given three-dimensional location $\vec{L}$, $G_i(\vec{L})$ is

completely specified, so we seek to minimize the cost function $J_f(i)$ with respect to $\hat{M}_i$,

$$J_f(i) = \frac{\|U_{m-p}^T \underline{A}_i\|_2^2}{\|\underline{A}_i\|_2^2} \qquad (28)$$

$$= \frac{\|U_{m-p}^T G_i \hat{M}_i\|_2^2}{\|G_i \hat{M}_i\|_2^2} \qquad (29)$$

$$= \frac{\hat{M}_i^T G_i^T P^\perp G_i \hat{M}_i}{\hat{M}_i^T G_i^T G_i \hat{M}_i} \qquad (30)$$

subject to $\hat{M}_i^T \hat{M}_i = 1$.

From (30) we recognize that minimizing this problem is equivalent to finding the minimum *generalized eigenvalue* (e.g., [31, p. 277], [30, p. 470]) of the expression

$$G_i^T P^\perp G_i \hat{M}_i = \lambda G_i^T G_i \hat{M}_i. \qquad (31)$$

If we denote the SVD of G_i as $G_i = U_{G_i} \Sigma_{G_i} V_{G_i}^T$, where U_{G_i} contains only the principal eigenvectors associated with the nonzero singular values, then the generalized eigenvalue problem can be expressed as

$$[V_{G_i} \Sigma_{G_i} U_{G_i}^T] P^\perp [U_{G_i} \Sigma_{G_i} V_{G_i}^T] \hat{M}_i$$
$$= \lambda [V_{G_i} \Sigma_{G_i} U_{G_i}^T][U_{G_i} \Sigma_{G_i} V_{G_i}^T] \hat{M}_i, \qquad (32)$$

$$U_{G_i}^T P^\perp U_{G_i} [\Sigma_{G_i} V_{G_i}^T \hat{M}_i] = \lambda [\Sigma_{G_i} V_{G_i}^T \hat{M}_i] \qquad (33)$$

such that we can now solve the equivalent simpler minimum eigenvalue problem,

$$J_f(i) = \lambda_{\min}\{U_{G_i}^T P^\perp U_{G_i}\} \qquad (34)$$

where $\lambda_{\min}\{\cdot\}$ denotes the minimum eigenvalue of the bracketed term.

Thus we need not explicitly scan for or calculate the best moment orientation that minimizes $J_f(i)$, but rather calculate just the eigenvalue associated with this moment. Consequently, the fixed dipole scanning procedure is the same as for the rotating model, except that at each location in our 3-D scanning grid we calculate the smallest eigenvalue of the bracketed term. For the fixed p-dipole model we expect to find p locations where $J_f(i) \approx 0$. The optimal moment orientation at each minimum of $J_f(i)$ is found as the eigenvector associated with $\lambda_{\min}$.

Note that calculating this cost function requires finding the smallest eigenvalue of only a 2×2 matrix. Also note that depending on the dimension p of the orthogonal projector, this small 2×2 matrix may be more efficiently formed as $U_{G_i}^T(I - U_p U_p^T) U_{G_i} = I - (U_p^T U_{G_i})^T (U_p^T U_{G_i})$, where U_p again represents the p left singular vectors associated with the p largest singular values of F. Thus, this cost function may be calculated with relatively little computational burden.

D. Fixed and Rotating Dipoles

In the mixed case, we have p_r rotating dipoles and p_f fixed dipoles for a total of $p = p_r + p_f$ dipoles. Our gen-

eral model is therefore

$$B = HS = [[\underline{G}_{\phi 1}\underline{G}_{\theta 1}], \cdots ,$$

$$[\underline{G}_{\phi p_r}\underline{G}_{\theta p_r}], \underline{A}_1, \cdots , \underline{A}_{pf}] \begin{bmatrix} [\underline{T}_{\phi 1}\underline{T}_{\theta 1}]^T \\ \vdots \\ [\underline{T}_{\phi p_r}T_{\theta p_r}]^T \\ \underline{S}_1^T \\ \vdots \\ \underline{S}_{pf}^T \end{bmatrix}. \quad (35)$$

We have two related scanning functions for rotating and fixed dipoles, $J_r(i)$ and $J_f(i)$. Examining $J_r(i)$, we see that

$$J_r(i) = \|P^\perp \tilde{G}_i\|_F^2 = \frac{\|P^\perp U_{G_i}\Sigma_{G_i}\|_F^2}{\|\Sigma_{G_i}\|_F^2} \quad (36)$$

$$= \frac{s_1^2\|P^\perp U_{G_i}\underline{e}_1\|_2^2 + s_2^2\|P^\perp U_{G_i}\underline{e}_2\|_2^2}{s_1^2 + s_2^2} \quad (37)$$

where s_1 and s_2 are the two nonzero singular values associated with G_i. The vectors $\underline{e}_1 = [1, 0]^T$ and $\underline{e}_2 = [0, 1]^T$ form an orthogonal unit moment orientation pair that spans the row space of G_i. By comparison, $J_f(i)$ can be viewed as

$$J_f(i) = \lambda_{\min}\{U_{G_i}^T P^\perp U_{G_i}\} \quad (38)$$

$$= \|P^\perp U_{G_i}\hat{M}_i^*\|_2^2 \quad (39)$$

where $\hat{M}_i^*$ is the optimal $\hat{M}_i$ that minimizes $J_f(i)$ over all unit moment orientations at that location. Comparing $\hat{M}_i^*$ with $\underline{e}_1$ and $\underline{e}_2$ above, we can describe the rotating dipole scanning function as a weighted average measure of how *all* dimensions of G_i project onto the noise subspace, while the fixed dipole function is a measure of how *one* optimal dimension of G_i projects. *Thus when $J_r(i) \approx 0$, then $J_f(i) \approx 0$.*

The rank of H is $r = 2p_r + p_f$ and is the crucial piece of information for this model. We otherwise proceed as for the *fixed* dipole model. At each location, we calculate the hybrid scanning function

$$J_h(i) = \lambda_{\min}\{U_{G_i}^T U_{m-r}U_{m-r}^T U_{G_i}\}$$

where U_{m-r} now contains the $m - r$ left singular vectors associated with the $m - r$ smallest singular values, as compared to $m - 2p$ and $m - p$ vectors for the two previous models. If at a location i we have a *fixed* dipole, then $G_i\hat{M}_i$ represents a linear combination of the vectors in G_i that is orthogonal to the noise subspace projector $(U_{m-r}U_{m-r}^T)$, and we obtain $J_h(i) \approx 0$. Correspondingly, if at a location i we have a *rotating* dipole, then G is already orthogonal to the noise subspace projector, regardless of the fixed orientation $\hat{M}_i$ we attempt to assign, and we still obtain $J_h(i) \approx 0$. *Thus we can use the fixed dipole scanning method for both rotating and fixed dipoles.* We summarize the complete steps in the following section,

where we show that this same algorithm can be derived under the proper statistical conditions.

V. MUSIC

A. Signal and Noise Subspaces

In the previous section we examined the rotating, fixed, and hybrid models separately. In this section we proceed directly to the general hybrid model $B = H(\underline{L}, M)S$ and assume that the hybrid gain matrix may represent any mix of fixed or rotating dipoles. We state the assumptions necessary for proving some of the MUSIC assertions:

AH (Gain Matrix Assumption) The $m \times r$ hybrid gain matrix H, $m > r$, is of full column rank r for p dipoles. In other words, the gain columns of the dipole components cannot be combined to simulate the gain columns of a third dipole component. For p_r rotating and p_f fixed, $p = p_r + p_f$ and $r = 2p_r + p_f$.

AS (Asynchronous Assumption) The moment time series for different dipole components are asynchronous or linearly independent, i.e., the time series of one component is not simply a scalar multiple of the time series from another component, nor can any combination of time series form another time series. Thus the time series matrix S is also of full rank r.

AW (Noise Whiteness) The additive noise is considered temporally and spatially zero-mean white noise with variance σ^2, such that the expectation of the outer product of the $m \times n$ noise matrix is $\mathcal{E}\{N(i)N(i)^T\}$ = $\sigma^2 I$, where n is the number of time slices. This requirement may be eased by prewhitening of the data, if the noise statistics are known.

The key assumption is that spatially distinct dipoles have linearly independent time series over the measured time segment; however, no constraint is imposed on whether or not the dipole moment is rotating, i.e., the method will work for either fixed or rotating dipoles.

Our model for noiseless data with m sensors, n time slices, and r elemental dipole components is $B = HS$, where H is $m \times r$, $m > r$, and S is $r \times n$, $r \leq n$. Consider the model for the noisy data under the assumption AW of zero mean white noise, $F = HS + N$. The spatial autocorrelation of the data is then

$$R_F \equiv \mathcal{E}\{F(i)F(i)^T\} = \mathcal{E}\{[HS(i) + N(i)]$$

$$\cdot [HS(i) + N(i)]^T\} = HR_S H^T + \sigma^2 I \quad (40)$$

where $\mathcal{E}\{\cdot\}$ is the expectation operator, and $R_S = \mathcal{E}\{S(i)S(i)^T\}$; by assumption AS, this correlation matrix R_S is of full rank. The square symmetric matrix R_F may be written in terms of its eigendecomposition as

$$R_F = \Phi\Lambda\Phi^T = [\Phi_s \quad \Phi_n] \begin{bmatrix} \Lambda_s & \\ & \Lambda_n \end{bmatrix} [\Phi_s \quad \Phi_n]^T \quad (41)$$

where we define Λ_s as the diagonal matrix containing the r largest eigenvalues and Φ_s as the matrix containing the corresponding eigenvectors. By our assumptions, it is well

known that the eigenvalue equal to the variance of the noise, $\lambda = \sigma^2$, repeats with multiplicity $m - r$ [19]. Accordingly, $\Lambda_n = \sigma^2 I$, and Φ_n is the matrix containing the corresponding $m - r$ eigenvectors.

Comparing Equations (40) and (41) and using assumptions AH and AS, it is straightforward to show that the space spanned by Φ_s is identical to that spanned by HR_SH^T; therefore, Φ_s is said to span the *signal* subspace. The space spanned by Φ_n is the orthogonal complement of the signal subspace and is referred to as the *noise* subspace. Based on these observations, it can be shown that the quantity

$$\lambda_{\min}\{U_{G_i}^T \Phi_n \Phi_n^T U_{G_i}\}, \tag{42}$$

is zero for any matrix $G_i = U_{G_i} \Sigma_{G_i} V_{G_i}^T$ corresponding to a true dipole location [19]. Thus we can determine the dipole locations exactly using Equation (42) provided R_F, and hence Φ_n, is known exactly.

In practice, MUSIC approximates R_F by $\hat{R}_F = (1/n)FF^T$. Estimates of the signal and noise subspaces, $\hat{\Phi}_s$ and $\hat{\Phi}_n$, are formed using an eigendecomposition of $\hat{R}_F$. Using this approximation for Φ_n, (42) is now equivalent to (34), which was derived using a deterministic suboptimal least-squares approach. Thus the scanning method presented in Section IV is also equivalent to the MUSIC algorithm for "polarized sources," as defined in [19], when the assumptions presented here hold true.

B. Summary of the MUSIC Algorithm

Summarizing the MUSIC algorithm:

1) Given the $m \times n$ data matrix F for m sensors and n time samples, perform the eigendecomposition of the estimate $\hat{R}_F = (1/n)FF^T = \hat{\Phi}\hat{\Lambda}\hat{\Phi}^T$. Order the eigenvalues, such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$. Equivalently, perform the SVD of F, where the eigenvalues are the square of the singular values.

2) Select the separation point $1 \leq r < m$ between the signal and noise subspace eigenvalues. By assumptions AH and AS, $r = 2p_r + p_f$, where p_r is the number of rotating dipoles and p_f is the number of fixed dipoles; therefore, r is the number of elemental dipoles. While theoretically $\lambda_{\min} = \sigma^2$ repeats with multiplicity $(m - r)$, in practice there is some spread among the smaller eigenvalues, depending on the number of time slices n used to estimate R_F. If the signals are of sufficient strength and sufficiently uncorrelated during the time interval, then a distinct drop in the magnitude of eigenvalues will occur between λ_r and λ_{r+1}. (A more detailed treatment of the order determination problem is given recently by Chen *et al.* [32].) Form the estimated matrices $\hat{\Phi}_s$ and $\hat{\Phi}_n$ from the corresponding signal and noise eigenvectors.

3) Over a fine grid of three-dimensional locations $\{(x, y, z)_i\}$, calculate the corresponding $m \times 2$ gain matrix G_i for each location, obtain the principal left eigenvectors U_{G_i} of G_i, using an SVD such

that $G_i = U_{G_i} \Sigma_{G_i} V_{G_i}^T$, and evaluate $J_h(i) = \lambda_{\min}\{U_{G_i}^T \hat{\Phi}_n \hat{\Phi}_n^T U_{G_i}\}$, i.e., the minimum eigenvalue of the bracketed expression. Form two-dimensional slices through the three-dimensional space, e.g., (x, y) planes for constant z, and plot the function $Z\{(x, y, z)_i\} = 1/J_h(i)$ as contours, images, or oblique mesh plots.

4) At each sharp "spike" (minimum of J_h), check to see if the entire subspace of G_i is orthogonal to the noise subspace (both eigenvalues of $\{U_{G_i}^T \hat{\Phi}_n \hat{\Phi}_n^T U_{G_i}\}$ are approximately zero), indicating that the dipole is rotating. Alternatively, if memory storage is not an issue, then at each point when evaluating $J_h(i)$ we can also calculate the rotating cost function $J_r(i) = \|\hat{\Phi}_n^T U_{G_i} \Sigma_{G_i}\|_F^2 / \|\Sigma_{G_i}\|_F^2$. Locations where $J_r(i) \approx J_h(i) \approx 0$ indicate rotating dipoles. If the dipole is indeed fixed (only $J_h(i) \approx 0$), we estimate its orientation by calculating the eigenvector associated with $\lambda_{\min}$. We repeat this analysis until we find p_r rotating dipoles and p_f fixed dipoles such that $r = 2p_r + p_f$. We can refine the estimate of the locations by either using a finer grid in these areas, or by using these estimates as the initialization point for a p-dipole least-squares search.

5) Form the hybrid gain matrix $H(\underline{L}, \underline{M})$ and solve for the time series, $S = H^\dagger F$.

Note that we do not explicitly need to know the number of rotating versus fixed dipoles, p_r versus p_f; instead, we need only $r = 2p_r + p_f$. The values of p_r and p_f are then found by the MUSIC algorithm. We also note that Step 4 indicates the possibility of detecting rotating dipoles by examining not only $J_r(i)$, but by also examining the other eigenvalues of $\{U_{G_i} \hat{\Phi}_n \hat{\Phi}_n^T U_{G_i}\}$ found in the calculation of $J_f(i)$; however, we have not studied differences in bias and variance among these various measures of fit.

C. Comparison to Other Dipole Fitting Methods

1) Least-Squares Scanning: It should be emphasized that the MUSIC scanning procedure is quite different from fitting the full data with a single dipole. The appeal of the one-dipole model is the relative simplicity with which we can form G_i. An alternative to the MUSIC approach would be to simply fit a one-dipole model at each point in a scanning grid, i.e., compute the function

$$J_{LSS}(i) = \|P_{G_i}^\perp F\|_F^2. \tag{43}$$

"Least-squares scanning" is then the evaluation of this function as the single dipole is scanned through the head region. The dipoles are assumed to lie at the locations corresponding to the local minima of J_{LSS}. Since we are fitting a one-dipole model, this method will naturally work when there is a single source. However, this method generally fails; for example, when there are multiple sources that are closely spaced or that generate fields of greatly differing intensities.

Least-squares scanning is equivalent to the beamsteering approach of conventional direction-finding and suffers

from the same problems of poor resolution and inter-source interference. The MUSIC approach itself was first proposed in the direction-finding arena to overcome these problems [19].

2) PCA Dipole Fitting: The MUSIC algorithm fits a one dipole model to a subspace derived from the data. PCA dipole fitting is a related method often examined in MEG analysis, usually with poor results [14], [15], [33], [34]. Although PCA dipole fitting also scans a one-dipole space derived from a similar decomposition, the subspace framework presented here allows us to examine this method and show where it fails.

PCA dipole fitting begins with the identical step of decomposing the data matrix F into its orthogonal components, $F = U\Sigma V^T$, then selecting the signal subspace. As described by Maier *et al.* [14] and analyzed by Achim *et al.* [15], the columns of U are the spatial distribution of the principal components, and the columns of V are the corresponding time functions. The matrix $W = U\Sigma'$ describes the "factor loadings," where Σ' contains only the r principal singular terms; this definition of W is consistent with our definition in Section III.

The PCA method fits a series of r *single* dipole models to the principal components as follows. The ith dipole location and moment are chosen as a least squares fit to a linear combination of the factor loadings,

$$\min_{L_i, Q_i} \|G_i(\vec{L}_i)\vec{Q}_i - W\underline{C}_i\|_2^2, \qquad (44)$$

where $G_i(\vec{L}_i)$ is the gain matrix for a dipole at location $\vec{L}_i$ and $\vec{Q}_i$ is the dipole moment. In [15], Achim *et al.* chose the individual terms of $\underline{C}_i$ to be $C_{ii} = 1$ and $C_{ij} = 0$ for $i \neq j$. In this case a single dipole is fit to each of the columns of W. Alternative choices of the rotation factors $\{C_{ij}\}$, such as Varimax, are discussed and analyzed in [33], [34] and elsewhere.

The error in (44) can be written as

$$J_{\text{PCA}}(i) = \|W\underline{C}_i - G_i\vec{Q}\|_2^2 = \|W\underline{C}_i - G_i(G_i^\dagger W\underline{C}_i)\|_2^2$$

$$= \|P_{G_i}^\perp W\underline{C}_i\|_2^2 \qquad (45)$$

where $P_{G_i}^\perp$ represents the orthogonal projection for a single dipole. The limitations of PCA dipole fitting are now apparent. PCA dipole fitting will succeed only if the coefficients in $\underline{C}_i$ are correctly selected, such that $W\underline{C}_i$ lies in the two-dimensional subspace spanned by G_i. However, this requires that we know the dipole location before we begin. As an example of the inaccuracy of PCA, Achim *et al.* [15] show a case in which PCA severely mislocates one of three dipoles in a *noiseless* simulation.

D. Simulation and Example

In this section, we present results using the MUSIC algorithm for both simulated data and experimental somatosensory data. We begin by using the same simulation model and data as in Section III-C, where at each of 37 sensor locations we simulate 100 time samples. Since by design we have two fixed dipoles and one rotating, then

our data model has a rank of four for three dipoles. We perform an SVD of the simulated noisy data matrix F and plot the singular values, of which the first ten are shown in Fig. 3.

The abrupt drop between values 4 and 5 give a clear indication in this simulation that the number of elemental dipoles is 4. We form the noise subspace estimate $\hat{\Phi}_n$ from the eigenvectors associated with singular values 5 to 37. We form the gain matrix at each voxel in the region $x = -5$ to 5 cm, $y = -5$ to 5 cm, and $z = 6$ to 9.5 cm, at 0.5 cm intervals. The minimum eigenvalue at each position using (34) is found, then the inverses of these values are formed into two-dimensional images for fixed z.

The results are shown in Fig. 4. The cost function shows three distinct peaks in Fig. 4(a) between $z = 8$ and $z = 9$ cm. The true locations, given in Table I, are just below the plane $z = 8.5$ cm and agree well with the positions shown in the figure. Fig. 4(b) shows the inverse of the second eigenvalue found in the evaluation of (34). The figure shows a distinct peak for the single rotating dipole in this simulation. We could refine our estimates by either scanning more finely around the $z = 8.5$ plane or by using these scanning estimates as an initialization point for a full three-dipole least-squares fit.

For the somatosensory experiment, the data were generated by vibrotactile stimulation (using a piezoelectric speaker element) of the right thumb, the right ring finger, and then both digits simultaneously. The intent was that the evoked field pattern for "both digits" might reflect a summation of the fields evoked by the stimulation of thumb and ring finger alone. The data were collected during eleven placements of a seven-sensor, second-order gradiometer system, at each placement averaging 300 trials. The data were digitized for 300 ms (100 prestimulus and 200 poststimulus) at 1 kHz and were filtered on line between 1 and 100 Hz. Fig. 5 displays the average absolute time response across all sensors for the three experiments, "ring," "thumb," and "both." Since piezoelectric stimulation creates a large stimulus artifact, the data were partitioned into a pre-stimulus interval up to 0 ms and a poststimulus interval after 24 ms.

Fig. 6 plots the first ten singular values from the decomposition of the spatio-temporal data matrices. The upper three curves are from the SVD of the poststimulus data matrices, and the lower three correspond to the prestimulus. The shape of the poststimulus curves and the merging of the prestimulus curves leads us to select the first six eigenvectors as our signal subspace. (For subspace orders five and four, we also obtained results similar to those presented below.) Fig. 7 displays the results of a 5 mm grid MUSIC scan for the "both digits" stimulation, using the first six principal components of the poststimulus interval. Each subimage represents an axial slice of the head in five millimeter increments along the z-axis, with the left ear at the top of the image and the nose at the right. The head coordinate system used was the x-axis through the nose, the y-axis through the left ear, and the z-axis up through the top of the head.
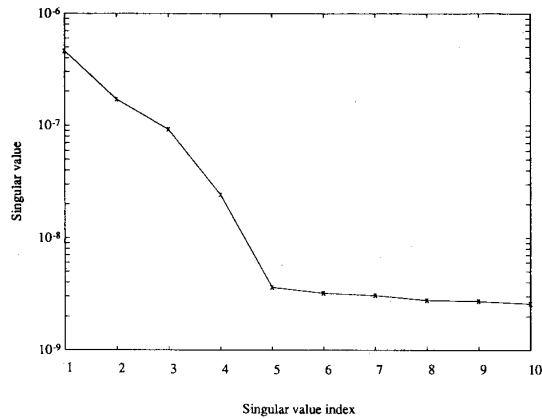
Fig. 3. Singular values for simulated noisy data matrix F. For clarity, only the first 10 of 37 singular values are plotted. The abrupt drop between singular values 4 and 5 gives a clear indication in this simulation that the number of elemental dipoles is 4. The noise singular values corresponding to indexes 5 and upward are seen to be approximately equal.
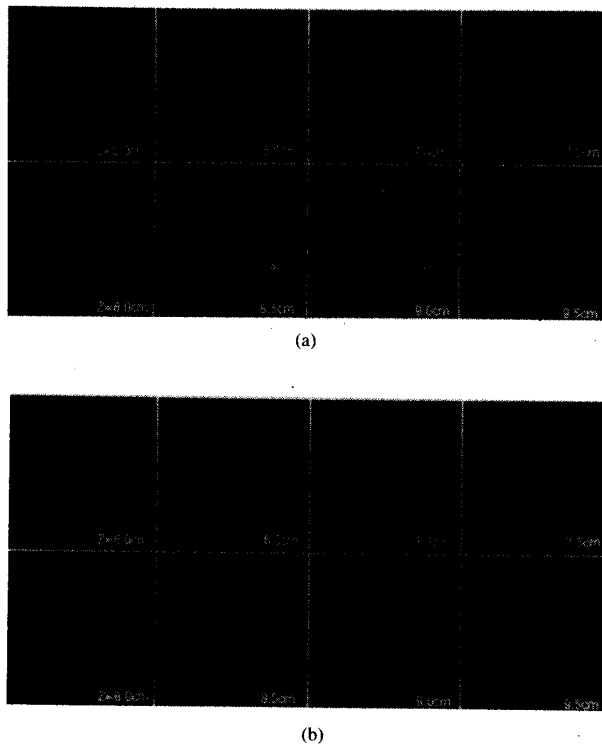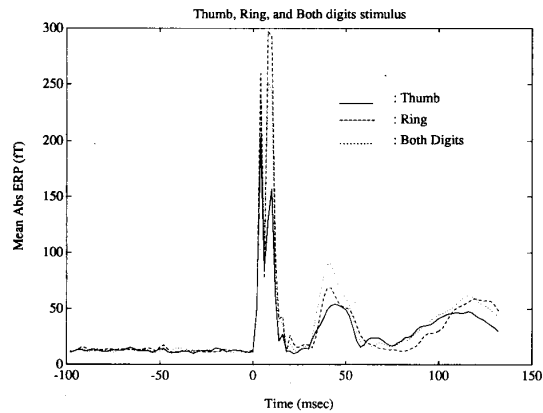

Fig. 5. Average absolute evoked field. The data were rectified across all sensor locations and averaged to give an indication of the temporal activity.
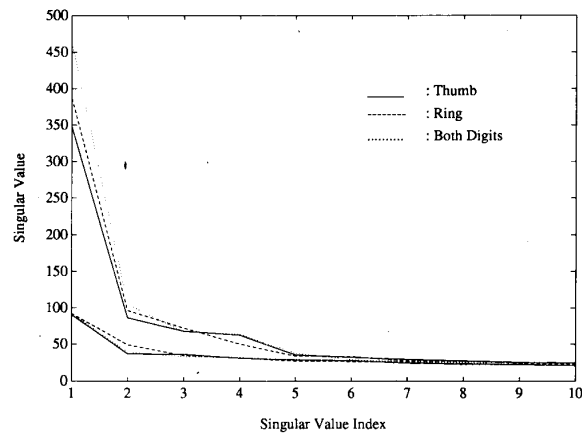

(a)

Fig. 6. Singular values of somatosensory data matrices. The top set of curves correspond to the three sets of poststimulus data (data after 24 ms), and the bottom set corresponds to the prestimulus data. We selected order six as our signal subspace; similar results were obtained for orders five and four.

(b)

Fig. 4. Simulation results. Each subimage is a two-dimensional slice in the (x, y) plane for $z = 6$ cm to 9.5 cm. Each (x, y) slice was formed at 0.5 cm intervals, from -5 to 5 cm in both the x and y directions. The simulation data are identical to the least-squares example, where the SNR is approximately 10 dB. The top set of images (a) is from the fixed dipole function, $J_f(i)$, (34). Here we have encoded the image as white to represent the minima in the cost function. The lower set of images (b), also from the fixed dipole model, shows the second (nonminimum) eigenvalue; a minimum in the second eigenvalue indicates the presence of a rotating dipole, as discussed in the text. The single minimum here correctly identifies the one rotating dipole. The true locations, given in Table I, are approximately in the $z = 8.3$ cm plane, with good agreement in the positions indicated here.

One millimeter scans were then centered on the observed peaks seen in the 5 mm scans. Figs. 8–10 display axial scans in one millimeter increments, from $x = -2$ to 2 cm, $y = 3$ to 7 cm, and $z = 6.1$ to 8 cm. We see a clear shift in the response among the three data sets in both the y and z coordinates. Fig. 11 summarizes the results from the $z = 7.3$ centimeter slice, where we have overlayed contour plots with full gray-scale images. Here we clearly see the distinct separation between the peaks for the thumb and the ring finger stimulation. The response due to the "both digits" stimulation actually peaks in the $z = 6.9$ cm plane, but we see in this $z = 7.3$ cm plane the indications of a possible summation of the response from the two somatosensory centers. A more thorough analysis would include discussions of noise correlation, order selection, and physiological interpretation, but we defer such analysis, since the emphasis here is simply to illustrate the utility of MUSIC analysis with real data.
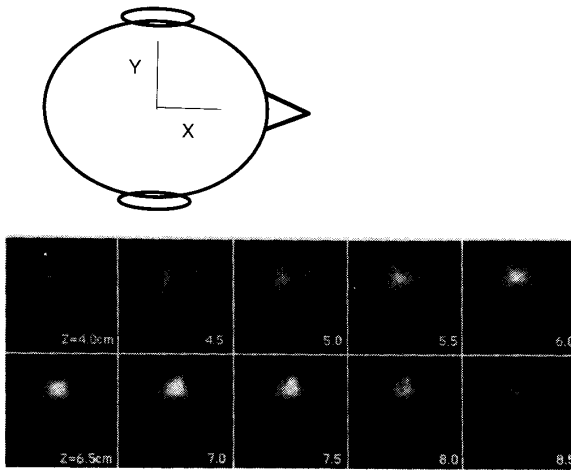
Fig. 7. Axial scans of the "both digits" stimulation case. Each subimage is a two-dimensional slice in the (x, y) plane for x = −10 to 10 cm, y = −10 to 10 cm, and z = 4 to 8.5 cm. The scanning grid increment is 0.5 cm in all directions. The head schematic gives the orientation of the axial slices. Based on these results, a 1 mm scanning grid was then formed around the indicated minima.
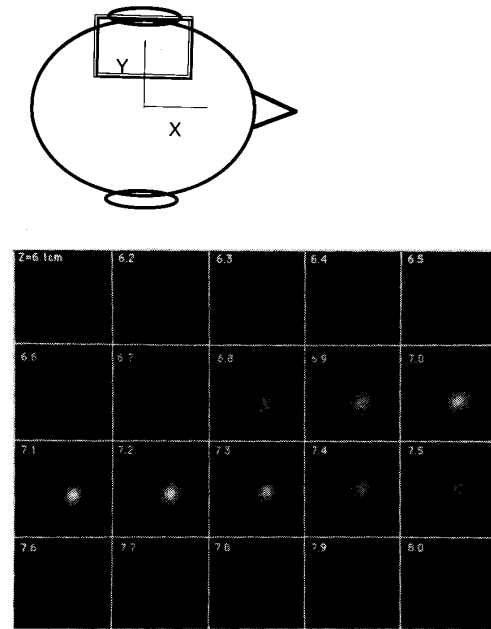

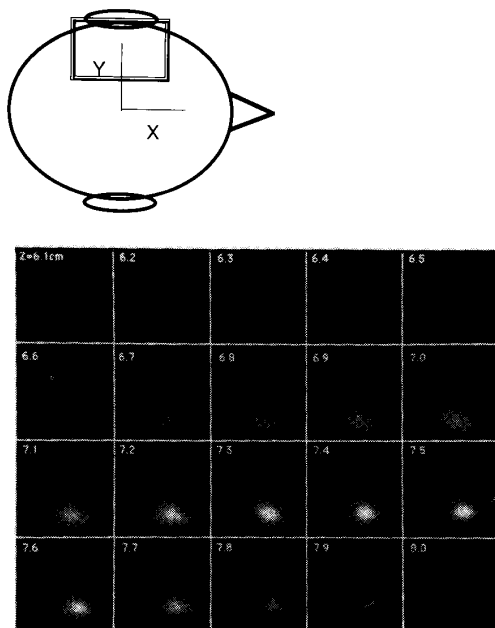Fig. 9. One millimeter axial scans of thumb data.

Fig. 8. One millimeter axial scans of ring data. Each subimage is a two-dimensional slice in the (x, y) plane for x = −2 to 2 cm, y = 3 to 7 cm, and z = 6.1 to 8 cm. The scanning increment is 0.1 cm in all directions. The head schematic gives the orientation of the axial slices, and the box indicates the region scanned.
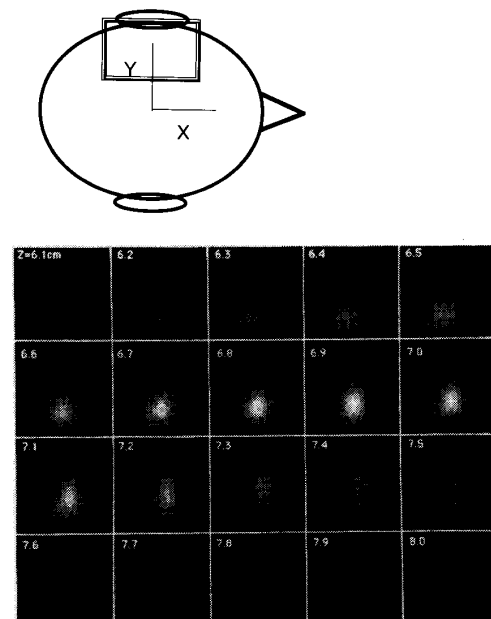

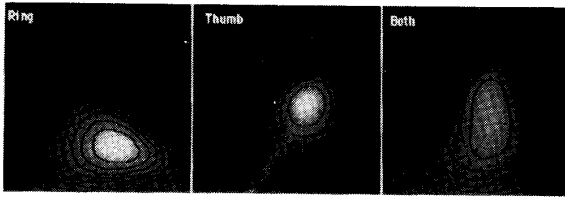Fig. 10. One millimeter axial scans of both digits data.

Fig. 11. Comparisons of the one millimeter scans at $z = 7.3$ cm. The (x, y) dimensions remain $x = -2$ to 2 cm and $y = 3$ to 7 cm. Contour plots of the axial slices are overlayed with a gray scale image of the data. On the left is the "ring" data, in the middle is the "thumb" data, and on the right is the "both digits" data.

VI. CONCLUSIONS

We have presented general descriptive models for spatio-temporal MEG data and have shown the separability of the linear moment parameters and nonlinear location parameters in the MEG problem. A forward model with current dipoles in a spherically symmetric conductor was used as an example; however, other more advanced MEG models, as well as many EEG models, can also be formulated in a similar linear algebra framework. A subspace methodology and computational approach to solving the conventional least-squares problem was then presented.

A new scanning approach, equivalent to the statistical MUSIC method, was then developed. This subspace method scans three-dimensional space with a one dipole model, making it computationally feasible to scan the complete head volume. Although PCA dipole fitting is related to this subspace method, we show how PCA dipole fitting fails, while the new scanning method generally succeeds. Least-squares and MUSIC demonstrations were presented using simulated noisy data and actual somatosensory MEG data. The MUSIC scanning method may also prove useful in constrained problems, such as scanning of the cortical regions of the brain using locations obtained from an MRI scan. In a constrained optimization over a very *irregular* reconstruction region, local minima abound, and conventional error function searches may be unable to locate the global minimum.

In general, MUSIC may fail when the noise is of sufficient strength to corrupt the estimates of the noise subspace, when the time series in S are strongly correlated, or when the sources are closely spaced. Many other authors have analyzed the performance of MUSIC, particularly the sensitivity of the results to errors in estimation of the noise subspace, and have also suggested many modifications of the algorithm [35]-[38]. These results may prove useful in improving the subspace scanning method described here. The challenge lies in advancing these methods for actual data where the noise statistics are unknown. The appeal lies in the relative simplicity in which the data are processed and the entire head region scanned.

ACKNOWLEDGMENT

We thank J. Lewine of the Los Alamos National Laboratory for providing us with his somatosensory data.

REFERENCES

[1] S. J. Williamson, G.-L. Romani, L. Kaufman, and I. Modena, Eds., *Biomagnetism: An Interdiscplinary Approach.* New York: Plenum, 1983.

[2] M. Scherg and D. von Cramon, "Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model," *Electroencephalogr. clin. Neurophysiol.,* vol. 62, pp. 32–44, 1985.

[3] J. H. Trip, "Physical concepts and mathematical models," in *Biomagnetism: An Interdisciplinary Approach,* S. J. Williamson, Ed. New York: Plenum, 1982, pp. 101–149.

[4] R. J. Ilmoniemi, M. S. Hamalainen, and J. Knuutila, "The forward and inverse problems in the spherical model," in *Biomagnetism: Applications and Theory,* H. Weinberg, G. Stroink, and T. Katila, Eds. Oxford: Pergamon, 1985, pp. 278–282.

[5] J. Sarvas, "Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem," *Phys. Med. Biol.,* vol. 32, no. 1, pp. 11–22, 1987.

[6] M. Singh, D. Doria, V. Henderson, G. Huth, and J. Beatty, "Reconstruction of images from neuro-magnetic fields," *IEEE Trans. Nucl. Sci.,* vol. NS-31, pp. 585–589, 1984.

[7] B. Jeffs, R. Leahy, and M. Singh, "An evaluation of methods for neuromagnetic image reconstruction," *IEEE Trans. Biomed. Eng.,* vol. BME-34, pp. 713–723, Sept. 1987.

[8] W. Dallas, "Space solution to the magnetostatic imaging problem," *Appl. Opt.,* vol. 24, no. 24, pp. 4543–4546, 1985.

[9] R. E. Alvarez, "Biomagnetic Fourier imaging," *IEEE Trans. Med. Imag.,* vol. 9, pp. 299–304, Sept. 1990.

[10] J. P. Wikswo, Jr., J. van Egeraat, Y. P. Ma, N. G. Sepulveda, D. J. Staton, S. Tan, and R. S. Wijesinghe, "Instrumentation and techniques for high-resolution magnetic imaging," in *Digital Image Synthesis and Inverse Optics,* A. F. Gmitro, P. S. Idell, and I. J. LaHaie, Eds., *Proc. SPIE,* 1351, July 1990, pp. 438–470.

[11] C. C. Wood, "Application of dipole localization methods to source identification of human evoked potentials," *Ann. New York Acad. Sci.,* vol. 388, pp. 139–155, 1982.

[12] M. Scherg, "Fundamentals of dipole source potential analysis," in *Auditory Evoked Magnetic Fields and Potentials, Vol. 6,* M. Hoke, F. Grandori, and G. L. Romani, Eds. Adv Audiol. Basel. Karger, 1989.

[13] M. Scherg and D. von Cramon, "A new interpretation of the generators of BAEP waves I-V: Results of a spatio-temporal dipole model," *Electroencephalogr. clin. Neurophysiol.,* vol. 62, pp. 290–299, 1985.

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

[15] A. Achim, F. Richer, and J. Saint-Hilaire, "Methods for separating temporally overlapping sources of neuroelectric data," *Brain Topography,* vol. 1, no. 1, pp. 22–28, 1988.

[16] M. Scherg and D. von Cramon, "Evoked dipole source potentials of the human auditory cortex," *Electroencephalogr. clin. Neurophysiol.,* vol. 65, pp. 344–360, 1986.

[17] J. C. Mosher, P. S. Lewis, R. Leahy, and M. Singh, "Multiple dipole modeling of spatio-temporal MEG data," in *Dig. Image Synth. Inverse Opt.,* A. F. Gmitro, P. S. Idell, and I. J. LaHaie, Eds., *Proc. SPIE,* 1351, July 1990, pp. 364–375.

[18] J. C. Mosher, P. S. Lewis, and R. Leahy, "Spatial localization of neural sources using the magnetoencephalogram," in *Fifth ASSP Workshop on Spectrum Estimation and Modeling,* Oct. 1990, pp. 289–293.

[19] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antenn. Propagat.,* vol. AP-34, pp. 276–280, Mar. 1986; reprint of the original 1979 paper from the RADC Spectrum Estimation Workshop.

[20] P. L. Nunez, "The brain's magnetic field: Some effects of multiple sources on localization methods," *Electroencephalogr. clin. Neurophysiol.,* vol. 63, pp. 75–82, 1986.

[21] S. J. Williamson and L. Kaufman, "Biomagnetism," *J. Magnetism Magnet. Mater.,* vol. 22, pp. 129–201, 1981.

[22] J. S. George, C. J. Aine, P. A. Medvick, and E. R. Flynn, "Spatial/temporal resolution of multiple sources: Paths of activation in human visual cortex," in *Advances in Biomagnetism,* S. Williamson *et al.,* Eds. New York: Plenum, 1989, pp. 197–200.

[23] 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.

[24] G. L. Romani and R. Leoni, "Localization of cerebral sources by neuromagnetic measurements," in *Biomagnetism: Applications and Theory*, H. Weinberg, G. Stroink, and T. Katila, Eds. New York: Pergamon, 1984, pp. 205-220.

[25] G. H. Golub and C. F. V. Loan, *Matrix Computations.* Baltimore, MD: Johns Hopkins Univ., 1983.

[26] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM J. Numerical Anal.*, vol. 10, pp. 413-432, Apr. 1973.

[27] I. Guttman and V. Pereyra, "Least squares estimation for a class of non-linear models," *Technometrics*, vol. 15, pp. 209-218, May 1973.

[28] J. J. Dongarra, C. B. Moler, J. R. Bunch, and G. W. Stewart, *LIN-PACK User's Guide.* Philadelphia: Soc. Industrial Appl. Math., 1979.

[29] The Mathworks, Inc., South Natick, MA 01760, *PRO-MATLAB*, 1990.

[30] G. H. Golub and C. F. V. Loan, *Matrix Computations.*, 2nd ed. Baltimore, MD: Johns Hopkins Univ., 1989.

[31] G. Strang, *Linear Algebra and its Applications.*, 2nd ed. Orlando, FL: Academic, 1980.

[32] W. Chen, K. M. Wong, and J. P. Reilly, "Detection of the number of signals: A predicted eigenthreshold approach," *IEEE Trans. Signal Process.*, vol. 39, pp. 1088-1098, May 1991.

[33] J. Mocks and R. Verleger, "Principal component analysis of event-related potentials: A note on misallocation of variance," *Electroencephalogr. clin. Neurophysiol.*, vol. 65, pp. 393-398, 1986.

[34] C. C. Wood and G. McCarthy, "Principal Component Analysis of event-related potentials: Simulation studies demonstrate misallocation of variance across components," *Electroencephalogr. clin. Neurophysiol.*, vol. 59, pp. 249-260, 1984.

[35] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, pp. 720-741, May 1989.

[36] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, pp. 1132-1143, July 1990.

[37] J. A. Cadzow, "Multiple source location—The signal subspace approach," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, pp. 1110-1125, July 1990.

[38] M. Viberg, "Subspace fitting concepts in sensor array processing," PhD thesis, Dep. Elect. Eng., Linkoping Univ., S-581 83 Linkoping, Sweden, 1989. Linkoping Studies in Science and Technology, Dissertations No. 217.

[39] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, pp. 1110-1121, May 1991.
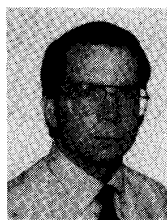
John C. Mosher (S'79-M'84-S'85-M'86) received the Bachelor's degree from the Georgia Institute of Technology, Atlanta, in 1983 and the M.S. degree from the University of Southern California, Los Angeles, in 1985, both in electrical engineering.

He is currently working toward the Ph.D. degree from USC in the Signal and Image Processing Institute of the Department of Electrical Engineering-Systems. Since 1983 he has been with TRW, Redondo Beach, as a signal processing engineer in the Electromagnetic Survivability Department. Areas of interest and research include biomedical and array signal processing, with emphasis on source localization.
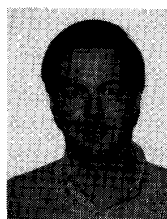
Paul S. Lewis (M'84-S'86-M'88-SM'89) received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1978, the M.S. degree in electrical engineering from the University of New Mexico in 1984, and the Ph.D. degree in electrical engineering from the University of Southern California in 1988.

He is a section leader at the Los Alamos National Laboratory, where he has worked since 1978. He has served as an adjunct professor at the University of New Mexico. His technical interests are in the area of digital signal processing, including biomedical applications, adaptive algorithms, and DSP hardware.

Dr. Lewis is a Registered Professional Engineer in the state of New Mexico.

Richard M. Leahy (M'86) was born in Surrey, England, in 1960. He received the B.Sc. and Ph.D. degrees in electrical engineering from the University of Newcastle upon Tyne, England, in 1981 and 1985, respectively.

In 1985 he joined USC where he is currently an Associate Professor in the Signal and Image Processing Institute of the Department of Electrical Engineering-Systems. He also holds a joint appointment with the Department of Radiology at USC. His research interests include medical image reconstruction and analysis, biomedical signal processing, and other applications of optimization theory and statistics in signal and image processing.