

# Recursive MUSIC: A Framework for EEG and MEG Source Localization

John C. Mosher,\* *Member, IEEE*, and Richard M. Leahy, *Member, IEEE*

**Abstract**—The multiple signal classification (MUSIC) algorithm can be used to locate multiple asynchronous dipolar sources from electroencephalography (EEG) and magnetoencephalography (MEG) data. The algorithm scans a single-dipole model through a three-dimensional (3-D) head volume and computes projections onto an estimated signal subspace. To locate the sources, the user must search the head volume for multiple local peaks in the projection metric. This task is time consuming and subjective. Here, we describe an extension of this approach which we refer to as recursive MUSIC (R-MUSIC). This new procedure automatically extracts the locations of the sources through a recursive use of subspace projections. The new method is also able to locate synchronous sources through the use of a spatio-temporal independent topographies (IT) model. This model defines a source as one or more nonrotating dipoles with a single time course. Within this framework, we are able to locate fixed, rotating, and synchronous dipoles. The recursive subspace projection procedure that we introduce here uses the metric of canonical or subspace correlations as a multidimensional form of correlation analysis between the model subspace and the data subspace. By recursively computing subspace correlations, we build up a model for the sources which account for a given set of data. We demonstrate here how R-MUSIC can easily extract multiple asynchronous dipolar sources that are difficult to find using the original MUSIC scan. We then demonstrate R-MUSIC applied to the more general IT model and show results for combinations of fixed, rotating, and synchronous dipoles.

**Index Terms**—Dipole modeling, electroencephalography, magnetoencephalography, signal subspace methods, source localization.

## I. INTRODUCTION

THE PROBLEM of localizing the sources of event related scalp potentials [the electroencephalogram (EEG)] and magnetic fields [the magnetoencephalogram (MEG)] can be formulated in terms of finding a least-squares fit of a set of current dipoles to the observed data. Early attempts at source localization were based on fitting the multiple-dipole model to a single time sample of the measurements across the E/MEG (EEG and/or MEG) array [5], [21], [31]. By noting that physiological models for the current sources typically assume

that they are spatially fixed for the duration of a particular response, researchers were able to justify fitting the multiple-dipole model to a complete spatio-temporal data set [2], [3], [17], [18]. The spatio-temporal model can result in substantial improvements in localization accuracy; however, processing the entire data set leads to a large increase in the number of unknown parameters, since the time series for each source must now be estimated in addition to the dipole location and orientation. Since these time series parameters are linear with respect to the data, they can be optimally factored out [7], [12] and the source locations found without explicit computation of their associated time series.

While factoring out the linear parameters can reduce the dimensionality of the search required to localize the sources of the measured fields, a fundamental problem remains: the least-squares cost function is highly nonconvex with respect to the locations of the dipoles. Consequently, inverse methods such as gradient-based optimization or nonlinear simplex searches often become trapped in local minima, yielding significant localization errors (cf. [9]). In an attempt to overcome this problem, we have examined the use of signal subspace methods that are common in the array signal processing literature [10]. The method that we used in [12], which was originally referred to as the MULTIPLE SIGNAL CLASSIFICATION (MUSIC) algorithm in [19], replaces the multiple-dipole directed search with a procedure in which a single dipole is scanned through a grid confined to a three-dimensional (3-D) head or source volume. At each point on this grid, the forward model for a dipole at this location is projected against a signal subspace that has been computed from the E/MEG data. The locations on this grid where the source model gives the best projections onto the signal subspace correspond to the dipole locations. We also show in [12] that we do not need to test all possible dipole orientations at each location; instead, we can solve a generalized eigenvalue problem whose solution gives us the best-fitting orientation of the dipole.

One of the major problems with the MUSIC method, and one that is addressed by the new approach described here, is how to choose the locations which give the best projection on to the signal subspace. In the absence of noise and with perfect head and sensor models, the forward model for a source at the correct location will project entirely into the signal subspace. In practice, of course, there are errors in the estimate of the signal subspace due to noise, and there are errors in the forward model due to approximations in our models of the head and data acquisition system. An additional problem is that we often

Manuscript received November 12, 1996; revised May 5, 1998. This work was supported in part by the National Institute of Mental Health under Grant ROI-MH53213, in part by the National Eye Institute under Grant ROI-EY08610-04, and in part by Los Alamos National Laboratory, operated by the University of California for the United States Department of Energy under Contract W-7405-ENG-36. *Asterisk indicates corresponding author.*

\*J. C. Mosher is with the Los Alamos National Laboratory, Group P-21 MS D454, Los Alamos, NM 87545 USA (e-mail: mosher@LANL.Gov).

R. M. Leahy is with the Signal and Image Processing Institute, University of Southern California, Los Angeles, CA 90089-2564 USA.

Publisher Item Identifier S 0018-9294(98)07779-9.

compute the metric only at a finite set of grid points. The effect of these practical limitations is that the user is faced with the problem of searching the gridded source volume for “peaks” and deciding which of these peaks correspond to true locations. It is important to note that a local peak in this metric does not necessarily indicate the location of a source. Only when the forward model projects entirely into the signal subspace—or as close as one would expect given errors due to noise and model mismatch—can we infer that a source is at that location. The effect of this limitation is that some degree of subjective interpretation of the MUSIC “scan” is required to decide on the locations of the sources. This subjective interpretation is clearly undesirable and can also lead to the temptation to incorrectly view the MUSIC scan as an image whose intensity is proportional to the probability of a source being present at each location.

Two other problems that arise with the use of MUSIC are based on the assumptions that the data are produced by a set of asynchronous dipolar sources and that the data are corrupted by additive spatially white noise. Often both of these assumptions are incorrect in clinical or experimental data. If two dipoles have synchronous activation, then the two-dimensional (2-D) signal subspace that would have been produced if they were asynchronous collapses into a one-dimensional (1-D) subspace. Scanning of a single dipole against this subspace using MUSIC will fail to localize either of the sources. The new MUSIC algorithm described here is able to localize synchronous sources through the use of a modified source representation, which we refer to as the spatio-temporal *independent topographies* (IT) model. This model is described in detail in Section II. The second problem, the issue of nonwhite noise, is not addressed in depth here. We note, however, that it is straightforward to modify both the original and R-MUSIC algorithms to cope with colored noise through standard prewhitening procedures [23]. In practice, the prewhitening could be achieved by using prestimulus data to estimate the covariance of the background noise; see [20] and [22] as recent examples of processing E/MEG data with colored noise.

We begin the paper in Section II with a combined formulation of the E/MEG forward problem in which we develop a standard matrix notation for the relationship between the source and data. We then describe the spatio-temporal IT model in which, rather than treating individual current dipoles as sources, we define a source as one or more nonrotating dipoles with a single time course. In this way our model is constrained to consist of a number of sources equal to the rank of the signal subspace. In Section III, we review the definition and properties of the signal subspace and relate these subspaces to cost functions commonly used for estimating the parameters of the model. We describe the use of canonical or “subspace” correlations as a general metric for computing the goodness of fit of putative sources to the signal subspace. We then review the MUSIC algorithm in the light of the preceding development. The new R-MUSIC algorithm is developed in Section IV. We present some examples of the application of R-MUSIC to fixed, rotating and synchronous dipolar sources in Section V.

## II. SPATIO-TEMPORAL INDEPENDENT TOPOGRAPHIES

### A. Background

Quasistatic approximations of Maxwell’s equations govern the relationship between neural current sources and the E/MEG data that they produce. For the signal subspace methods for source localization that are described here, these relationships must be expressed in matrix form. In [15], we reviewed matrix forms of the “lead field” [4], [27] for EEG and MEG measurements, for both spherical and general head models. In each case, the measurements can be expressed as an explicit function of *primary current* activity; the *passive volume currents* due to the macroscopic electric fields are implicitly embedded in the lead field formula. The lead field should also account for the sensor characteristics of the measurement modality, such as gradiometer orientation and configuration in MEG or differential pairs in EEG. The result is that our EEG or MEG measurement  $\tilde{f}(\mathbf{r})$  at sensor location  $\mathbf{r}$  may be expressed as

$$\tilde{f}(\mathbf{r}) = \int_V \mathbf{g}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{j}(\mathbf{r}') d\mathbf{r}' \quad (1)$$

where  $V$  is the volume of sources,  $\mathbf{j}(\mathbf{r}')$  represents the *primary current density* at any point  $\mathbf{r}'$  in the volume, and  $\mathbf{g}(\mathbf{r}, \mathbf{r}')$  is the *lead field* vector [4], [27] relating the sensor point to the primary current point. The scalar function  $\tilde{f}(\mathbf{r})$  represents either the voltage potential or the magnetic field component that would be observed at sensor location  $\mathbf{r}$ .

If we assume that the primary current exists only at a discrete point  $\mathbf{r}_q$ , i.e., the primary current is  $\mathbf{j}(\mathbf{r}')\delta(\mathbf{r}' - \mathbf{r}_q)$ , where  $\delta(\mathbf{r}' - \mathbf{r}_q)$  is the Dirac delta functional, then (1) simplifies in E/MEG to

$$\tilde{f}(\mathbf{r}) = \mathbf{g}(\mathbf{r}, \mathbf{r}_q) \cdot \mathbf{q} \quad (2)$$

where  $\mathbf{q}$  is the *moment of a current dipole* located at  $\mathbf{r}_q$ . We assume in this paper that our source consists of  $p$  current dipole sources. We assume simultaneous recordings at  $m$  sensors for  $n$  time instances. We can express the  $m$  by  $n$  spatio-temporal data matrix as

$$\begin{aligned} & \begin{bmatrix} \tilde{f}(\mathbf{r}_1, t_1) & \cdots & \tilde{f}(\mathbf{r}_1, t_n) \\ \cdots & \cdots & \cdots \\ \tilde{f}(\mathbf{r}_m, t_1) & \cdots & \tilde{f}(\mathbf{r}_m, t_n) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{g}(\mathbf{r}_1, \mathbf{r}_{q1})^T & \cdots & \mathbf{g}(\mathbf{r}_1, \mathbf{r}_{qp})^T \\ \cdots & \cdots & \cdots \\ \mathbf{g}(\mathbf{r}_m, \mathbf{r}_{q1})^T & \cdots & \mathbf{g}(\mathbf{r}_m, \mathbf{r}_{qp})^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1(t_1) & \cdots & \mathbf{q}_1(t_n) \\ \cdots & \cdots & \cdots \\ \mathbf{q}_p(t_1) & \cdots & \mathbf{q}_p(t_n) \end{bmatrix} \end{aligned} \quad (3)$$

or

$$\tilde{\mathbf{F}} = [\mathbf{G}(\mathbf{r}_{q1}) \cdots \mathbf{G}(\mathbf{r}_{qp})] \mathbf{Q}^T. \quad (4)$$

We refer to  $\mathbf{G}(\mathbf{r}_{qi})$  as the dipole “gain matrix” [12] that maps a dipole at  $\mathbf{r}_{qi}$  into a set of measurements. The three columns of the gain matrix represent the possible *forward fields* that may be generated by the three orthogonal orientations of the  $i$ th dipole at the  $m$  sensor locations  $\{\mathbf{r}_1, \dots, \mathbf{r}_m\}$ . Each row of the full gain matrix  $[\mathbf{G}(\mathbf{r}_{q1}) \cdots \mathbf{G}(\mathbf{r}_{qp})]$  represents the *lead field* sampled at the discrete dipole locations

$\{\mathbf{r}_{q1}, \dots, \mathbf{r}_{qp}\}$ . The matrix  $\tilde{\mathbf{F}}$  is our *spatio-temporal model matrix* of perfect measured data, i.e., the magnetic field component or scalp potential data we would observe in the absence of noise.

The columns of  $\mathbf{Q}$  represent the time series associated with each of the three orthogonal components of each dipole, i.e., with each column of the gain matrix. For the “fixed”-dipole model, whose moment orientation is time invariant, we can separate the orientation of each source from the moments as [12]

$$\tilde{\mathbf{F}} = [\mathbf{G}(\mathbf{r}_{q1}) \ \dots \ \mathbf{G}(\mathbf{r}_{qp})] \begin{bmatrix} \mathbf{u}_{q1} & & 0 \\ & \dots & \\ 0 & & \mathbf{u}_{qp} \end{bmatrix} \cdot \begin{bmatrix} s_{q1}(t_1) & \dots & s_{q1}(t_n) \\ \dots & \dots & \dots \\ s_{qp}(t_1) & \dots & s_{qp}(t_n) \end{bmatrix} \quad (5)$$

such that  $\mathbf{q}_i(t_k) = \mathbf{u}_{qi} s_{qi}(t_k)$ , where  $\mathbf{u}_{qi}$  is a unit norm orientation vector. The scalar time series  $s_{qi}(t_k)$  are the *linear* parameters of the  $i$ th source of our model, for time indexes  $k = 1, \dots, n$ . The corresponding dipole locations  $\mathbf{r}_{qi}$  are the *nonlinear* parameters, and the dipole orientations  $\mathbf{u}_{qi}$  are the *quasilinear* parameters.

In [12], we also considered a “rotating” dipole as one whose time series could not be decomposed into a single fixed orientation and time series and, therefore, comprised multiple “elemental” dipoles. Physically, a rotating dipole may be viewed as two nearly collocated dipoles with independent time series, such that they are indistinguishable from a model comprising a single dipole whose orientation is allowed to vary with time. The “hybrid” models in [12] comprised both fixed and rotating dipoles.

### B. Independent Topographies

A common observation in MUSIC processing is that the rank of the signal subspace should equal the number of “sources.” As defined in [12], we allowed a source to be a single dipole with one (“fixed”) or more (“rotating”) independent time courses. For example, the data model defined by (5) assumes a collection of  $p$  “fixed” dipoles. The MUSIC method applied to this model requires that all dipoles have linearly independent time courses. Stated conversely, each linearly independent time course is associated with a single dipole.

Our goal in this section is build up a new source model for the data as the sum of contributions from a fixed number of spatio-temporal *independent topographies*. Each of these IT’s is considered to be a source comprising one or more fixed dipoles which collectively have a single time course. Thus, in contrast to our models in [12], each linearly independent time course is now associated with one *or more* dipoles. By building up the model in this way, the rank of the signal subspace is, by definition, always equal to the number of sources. This representation then provides a convenient framework for describing and implementing our new variant of the MUSIC algorithm.

We cluster the  $p$  dipoles into  $r$  subsets and restate (4) as  $\tilde{\mathbf{F}} = \mathbf{G}(\rho)\mathbf{Q}^T$  where  $\rho \equiv \{\rho_1, \dots, \rho_r\}$  represents  $r$  sets of dipoles. The  $i$ th set of dipoles  $\rho_i$  comprises  $p_i$  dipoles with the location parameter set  $\rho_i \equiv \{\mathbf{r}_{q1}^{(i)}, \dots, \mathbf{r}_{qp}^{(i)}\}$ . The moments of the  $i$ th set are clustered such that, by design, they *must* result in a rank-one time series matrix. A singular value decomposition (SVD) of these moments yields

$$\mathbf{Q}_i^T = \begin{bmatrix} \mathbf{q}_1^{(i)}(t_1) & \dots & \mathbf{q}_1^{(i)}(t_n) \\ \dots & \dots & \dots \\ \mathbf{q}_{p_i}^{(i)}(t_1) & \dots & \mathbf{q}_{p_i}^{(i)}(t_n) \end{bmatrix} = \mathbf{u}_i \sigma_i \mathbf{v}_i^T \quad (6)$$

where  $\mathbf{u}_i$  and  $\mathbf{v}_i$  are, respectively, the left and right singular vectors of unit norm, and  $\sigma_i$  is the only nonzero singular value of the decomposition of  $\mathbf{Q}_i^T$ . The result is that the  $i$ th set of dipoles may be represented by the rank one spatio-temporal matrix

$$\mathbf{G}(\rho_i)\mathbf{Q}_i^T = [\mathbf{G}(\mathbf{r}_{q1}^{(i)}) \ \dots \ \mathbf{G}(\mathbf{r}_{qp}^{(i)})] \mathbf{u}_i \sigma_i \mathbf{v}_i^T = \mathbf{a}(\rho_i, \mathbf{u}_i) \mathbf{s}_i^T \quad (7)$$

where the time series follows from the decomposition in (6) as  $\mathbf{s}_i = \sigma_i \mathbf{v}_i$ .

The  $\mathbf{r}$  sets of dipoles are then collected such that (4) is now restated as

$$\tilde{\mathbf{F}} = \mathbf{G}(\rho)\mathbf{Q}^T = [\mathbf{a}(\rho_1, \mathbf{u}_1) \ \dots \ \mathbf{a}(\rho_r, \mathbf{u}_r)] \begin{bmatrix} \mathbf{s}_1^T \\ \dots \\ \mathbf{s}_r^T \end{bmatrix} \quad (8)$$

$$= \mathbf{A}(\rho, \theta) \mathbf{S}^T \quad (9)$$

where the set  $\theta \equiv \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$  contains the corresponding unit norm vectors. Our final requirement in this IT framework is that  $\tilde{\mathbf{F}}$  *must* be of rank  $r$  and, therefore, by necessity, both  $\mathbf{A}(\rho, \theta)$  and  $\mathbf{S}$  are of full column rank  $r$ . We refer to each column vector of  $\mathbf{A}(\rho, \theta)$  as a “ $p_i$ -dipolar IT,” with a corresponding time series found as the  $i$ th column of  $\mathbf{S}$ .

IT and previous spatio-temporal source models [12], [18] differ as follows. If written as (4), we can state that the model comprises  $p$  dipolar sources and corresponding time series. This model has led to the complexity of representing “fixed” and “rotating” or “regional” dipoles. The rank of this spatio-temporal dipolar model can vary from unity (all dipoles fixed and synchronous) to  $3p$  (fully rotating dipoles in an EEG model). The number of “sources” depended on the viewpoint of the researcher, since some viewed the rotating dipole as a single “regional” source or as multiple “elemental” sources. When written as (9), we have altered the spatio-temporal model by stating that the number of sources *must* equal the rank of the spatio-temporal model, where a single source may now comprise one or more dipoles. Each source generates a “topography” across the array of sensors that is spatially and temporally independent of any other source or topography.

We will conclude this section with some examples that show the relationships between IT models and common multiple-dipole models. The case of three fixed dipoles, with asynchronous time series, is represented by the rank-three spatio-temporal matrix as

$$\mathbf{A}(\rho, \theta) \mathbf{S}^T = [\mathbf{a}(\mathbf{r}_{q1}, \mathbf{u}_{q1}), \mathbf{a}(\mathbf{r}_{q2}, \mathbf{u}_{q2}), \mathbf{a}(\mathbf{r}_{q3}, \mathbf{u}_{q3})] \cdot [s_{q1}, s_{q2}, s_{q3}]^T \quad (10)$$

where  $\rho \equiv \{\mathbf{r}_{q1}, \mathbf{r}_{q2}, \mathbf{r}_{q3}\}$  is the set of dipole locations and  $\theta \equiv \{\mathbf{u}_{q1}, \mathbf{u}_{q2}, \mathbf{u}_{q3}\}$  is the set of dipole orientations as shown in (5). Our IT model comprises three “single-dipolar topographies.” Now consider the case where two of these dipoles are synchronous, but the third remains asynchronous from the others. The rank of the model is two, and our IT model must be adjusted to account for two sources

$$\mathbf{A}(\rho, \theta)\mathbf{S}^T = [\mathbf{a}(\mathbf{r}_{q1}, \mathbf{u}_{q1}), \mathbf{a}(\rho_2, \mathbf{u}_2)][\mathbf{s}_{q1}, \mathbf{s}_2]^T \quad (11)$$

where now  $\rho \equiv \{\mathbf{r}_{q1}, \rho_2\}$  is the set of source locations, and the second source comprises two dipoles,  $\rho_2 \equiv \{\mathbf{r}_{q2}, \mathbf{r}_{q3}\}$ . The set of source orientations is  $\theta \equiv \{\mathbf{u}_{q1}, \mathbf{u}_2\}$ , and we may continue to view  $\mathbf{u}_{q1}$  as the orientation of the dipole in the single-dipole topography. The orientation  $\mathbf{u}_2$ , however, is a generalized form of source orientation and is in this instance a six-dimensional (6-D) vector, as discussed in the Appendix. The first three elements of this vector relate to the orientation of dipole  $\mathbf{r}_{q2}$  and the second set of three elements relate to  $\mathbf{r}_{q3}$ . The norm of the first three elements gives the strength of dipole  $\mathbf{r}_{q2}$  relative to  $\mathbf{r}_{q3}$ , with the restriction that the combined vector of all elements in  $\mathbf{u}_2$  is of unit norm. This IT model comprises one “single-dipolar” topography and one “2-dipolar” topography.

If all three dipoles are synchronous, then our IT model comprises a single “3-dipolar” topography. Similarly, we consider the limiting case of a single time slice and a general  $p$ -dipole model; this spatio-temporal model is always rank one, and our IT model is, therefore, a single  $p$ -dipolar topography.

“Rotating” or “regional” dipoles are a simple extension of the IT framework. In the case of the MEG spherical head model, a single “rotating” dipole becomes two fixed-orientation dipoles corresponding to two “single-dipolar” topographies

$$\mathbf{A}(\rho, \theta)\mathbf{S}^T = [\mathbf{a}(\rho_1, \mathbf{u}_{q1}), \mathbf{a}(\rho_2, \mathbf{u}_{q2})][\mathbf{s}_{q1}, \mathbf{s}_{q2}]^T \quad (12)$$

with  $\rho_1 = \rho_2 = \mathbf{r}_{q1}$ . In other words, two fixed dipoles share the same location, just different orientations and time courses, which is the original intent of a “rotating” model. For other head models, the dipole may possibly rotate in three dimensions,  $\rho_1 = \rho_2 = \rho_3 = \mathbf{r}_{q1}$ . Note that the  $m \times 2$  (or  $m \times 3$ ) IT gain matrix  $[\mathbf{a}(\mathbf{r}_{q1}, \mathbf{u}_{q1})\mathbf{a}(\mathbf{r}_{q1}, \mathbf{u}_{q2})]$  (or  $[\mathbf{a}(\mathbf{r}_{q1}, \mathbf{u}_{q1})\mathbf{a}(\mathbf{r}_{q1}, \mathbf{u}_{q2})\mathbf{a}(\mathbf{r}_{q1}, \mathbf{u}_{q3})]$ ) properly remains of full column rank.

### III. SIGNAL SUBSPACE METHODS

#### A. Signal Subspaces

We will now investigate the relationship between the IT model and the signal subspace that we estimate from the spatio-temporal data. Assume that a random error matrix comprising  $n$  time slices  $\mathbf{E} \equiv [\mathbf{e}(t_1) \cdots \mathbf{e}(t_n)]$  is added to the data  $\hat{\mathbf{F}} = \mathbf{A}(\rho, \theta)\mathbf{S}^T$ , to produce an  $m \times n$  “noisy” spatio-temporal data set

$$\mathbf{F} = \mathbf{A}\mathbf{S}^T + \mathbf{E}. \quad (13)$$

For convenience, we will drop the explicit dependence of  $\mathbf{A}$  on its parameters. The goal of the inverse problem is to estimate the parameters,  $\{\rho, \theta, \mathbf{S}\}$ , given the data set  $\mathbf{F}$ . We will use

the common assumption that the noise is zero-mean and white, i.e.,  $E\{\mathbf{e}(t_i)\mathbf{e}(t_i)^T\} = \sigma_e^2\mathbf{I}$ , where  $E\{\bullet\}$  denotes the expected value of the argument, and  $\mathbf{I}$  is the identity matrix. The case for colored noise is readily treated with standard prewhitening methods [23], provided a reasonable estimate of the noise covariance is available. For event related studies, the noise covariance can probably be estimated using sufficiently long periods of prestimulus data.

Under the zero-mean white noise assumption, we may represent the expected value of the matrix outer product  $\mathbf{F}\mathbf{F}^T$  as

$$\begin{aligned} \mathbf{R}_F &\equiv E\{\mathbf{F}\mathbf{F}^T\} = \mathbf{A}\mathbf{S}^T\mathbf{S}\mathbf{A}^T + \sum_{i=1}^n E\{\mathbf{e}(t_i)\mathbf{e}(t_i)^T\} \\ &= \mathbf{A}\mathbf{S}^T\mathbf{S}\mathbf{A}^T + n\sigma_e^2\mathbf{I}. \end{aligned} \quad (14)$$

Here, we have assumed that our model parameters and the dipolar time series are deterministic. From our IT model, we know that  $\mathbf{A}\mathbf{S}^T\mathbf{S}\mathbf{A}^T$  is rank  $r$  and may be decomposed as  $\Phi_s\Lambda\Phi_s^T$ , where  $\Phi_s$  contain the  $m \times r$  eigenvectors such that  $\text{span}(\Phi_s) = \text{span}(\mathbf{A})$ , and  $\Lambda$  is the corresponding  $r \times r$  diagonal matrix of nonzero eigenvalues.

We can decompose the correlation matrix as

$$\mathbf{R}_F = [\Phi_s, \Phi_e] \begin{bmatrix} \Lambda + n\sigma_e^2\mathbf{I} & 0 \\ 0 & n\sigma_e^2\mathbf{I} \end{bmatrix} [\Phi_s, \Phi_e]^T \quad (16)$$

$$= \Phi_s\Lambda_s\Phi_s^T + \Phi_e\Lambda_e\Phi_e^T \quad (17)$$

where  $\Lambda_s \equiv \Lambda + n\sigma_e^2\mathbf{I}$  is the  $r \times r$  diagonal matrix combining both the model and noise eigenvalues, and  $\Lambda_e \equiv n\sigma_e^2\mathbf{I}$  is the  $(m-r) \times (m-r)$  diagonal matrix of noise-only eigenvalues. The  $m-r$  eigenvectors that are not in  $\Phi_s$  are contained in  $\Phi_e$ . In the subspace discussions below, we refer to  $\text{span}(\Phi_s)$  as the *signal subspace* and  $\text{span}(\Phi_e)$  as the *orthogonal signal subspace*, or *noise-only subspace*.

In practice, we acquire  $n$  samples of the data to form the spatio-temporal data matrix and eigendecompose the outer product of this matrix as

$$\hat{\mathbf{R}}_F = \mathbf{F}\mathbf{F}^T = \hat{\Phi}_s\hat{\Lambda}_s\hat{\Phi}_s^T + \hat{\Phi}_e\hat{\Lambda}_e\hat{\Phi}_e^T \quad (18)$$

where we designate the first  $r$  left singular vectors of the decomposition as  $\hat{\Phi}_s$ , i.e., our estimate of a set of vectors which span the signal subspace; similarly we designate  $\hat{\Phi}_e$  from the remaining eigenvectors. The diagonal matrix  $\hat{\Lambda}_s$  contains the first  $r$  eigenvalues and  $\hat{\Lambda}_e$  the remaining eigenvalues. Alternatively, we may decompose this data matrix using an SVD as  $\mathbf{F} = \hat{\Phi}\hat{\Sigma}\hat{\Psi}^T$  and partition similarly, where we note that  $\hat{\Sigma}_s^2 = \hat{\Lambda}_s$ .

#### B. Parameter Estimation

In *least-squares fitting*, we estimate the source parameters as the arguments which minimize the cost function,

$$\{\hat{\rho}, \hat{\theta}, \hat{\mathbf{S}}\}_{\text{ls}} = \arg \min \|\mathbf{F} - \mathbf{A}(\rho, \theta)\mathbf{S}^T\|_{\text{F}}^2 \quad (19)$$

i.e., we minimize the squared Frobenius norm of the error matrix. As we reviewed in [12], well-known optimal substitution [7] of the linear terms yields the modified cost function

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ls}} = \arg \min \|\mathbf{F} - \mathbf{A}\mathbf{A}^\dagger\mathbf{F}\|_{\text{F}}^2 \quad (20)$$

where  $\mathbf{A}^\dagger$  is the Moore–Penrose pseudoinverse of  $\mathbf{A}$  [8]. Once the minimizing set of parameters  $\{\hat{\rho}, \hat{\theta}\}_{\text{ls}}$  has been found, the linear parameters are simply found as  $\hat{\mathbf{S}}_{\text{ls}}^T = \mathbf{A}(\{\hat{\rho}, \hat{\theta}\}_{\text{ls}})^\dagger \mathbf{F}$ .

The product of  $\mathbf{A}$  and its pseudoinverse  $\mathbf{A}^\dagger$  is an orthogonal projection operator  $\Pi_{\mathbf{A}} = \mathbf{A}\mathbf{A}^\dagger$ , which may equivalently be represented by the outer product  $\Pi_{\mathbf{A}} = \mathbf{U}_A \mathbf{U}_A^T$ , where  $\mathbf{U}_A$  is the matrix whose columns are the left singular vectors of  $\mathbf{A}$  that correspond to its nonzero singular values. Substituting this definition yields

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ls}} = \arg \min \{ \|\mathbf{F}\|_F^2 - \|\Pi_{\mathbf{A}}\mathbf{F}\|_F^2 \} \quad (21)$$

into which we substitute the SVD of the data matrix and readily restate as a maximization

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ls}} = \arg \max \|\mathbf{U}_A \mathbf{U}_A^T \hat{\Phi} \hat{\Sigma} \hat{\Phi}^T\|_F^2. \quad (22)$$

Using the Frobenius norm-preserving properties of orthogonal matrices, we can restate (22) as

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ls}} = \arg \max \|\mathbf{U}_A^T \hat{\Phi} \hat{\Sigma}\|_F^2. \quad (23)$$

Finally, substituting in our signal and noise subspace representations of  $\mathbf{F}$  yields

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ls}} = \arg \max \left\{ \|\mathbf{U}_A^T \hat{\Phi}_s \hat{\Sigma}_s\|_F^2 + \|\mathbf{U}_A^T \hat{\Phi}_e \hat{\Sigma}_e\|_F^2 \right\}. \quad (24)$$

From (24), we may observe the following. Since the singular values of  $\mathbf{F}$  are ordered in decreasing value, maximization of this least-squares cost function favors fitting the first term containing  $\hat{\Sigma}_s$  rather than the second term containing  $\hat{\Sigma}_e$ . Similarly, as we acquire more data, our estimate of  $\mathbf{R}_F$  improves and, consequently, so does  $\hat{\Phi}_s$  and  $\hat{\Phi}_e$ . By construction, the true values project as  $\Pi_{\mathbf{A}}\Phi_s = \Phi_s$  and  $\Pi_{\mathbf{A}}\Phi_e = 0$ . These observations lead to an alternative maximization criterion function that focuses on just the first term

$$\{\hat{\rho}, \hat{\theta}\}_{\text{mls}} = \arg \max \|\mathbf{U}_A^T \hat{\Phi}_s \hat{\Sigma}_s\|_F^2 \quad (25)$$

i.e., a “modified” least-squares criterion [10], [24], [28]. Effectively,  $\hat{\Sigma}_s$  in (25) represents a weighted sum of the projections of the estimated signal subspace eigenvectors. *Signal subspace fitting* (SSF) [28] is (25) with the weighting matrix set to the identity matrix

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ssf}} = \arg \max \|\mathbf{U}_A^T \hat{\Phi}_s\|_F^2. \quad (26)$$

In these cost functions, we note the common inner product  $\mathbf{U}_A^T \hat{\Phi}_s$ . We can decompose this product with an SVD to yield

$$\mathbf{U}_A^T \hat{\Phi}_s = \mathbf{Y} \Sigma_C \mathbf{Z}^T \quad (27)$$

where  $\mathbf{Y}$  and  $\mathbf{Z}$  are each  $r \times r$  orthogonal matrices. We designate the  $r$  ordered singular values in the diagonal matrix  $\Sigma_C$  as  $\{c_1, \dots, c_r\}$ . Substituting (27) into (26) yields

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ssf}} = \arg \max \{ \|\Sigma_C\|_F^2 \} \quad (28)$$

to yield simply an unweighted sum

$$\{\hat{\rho}, \hat{\theta}\}_{\text{ssf}} = \arg \max \left\{ \sum_{k=1}^r c_k^2 \right\}. \quad (29)$$

The decomposition shown in (27) yields the *subspace correlations* between the subspaces spanned by  $\mathbf{A}$  and  $\hat{\Phi}_s$  [8]. In the Appendix, we summarize the steps for calculating these subspace correlations to yield a function we designate as

$$\{c_1, c_2, \dots\} = \text{subcorr}\{\mathbf{A}, \hat{\Phi}_s\} \quad (30)$$

which returns the ordered set of subspace correlations  $\{c_1, \dots, c_r\}$ ,  $1 \geq c_1 \geq \dots \geq c_r \geq 0$ , between the two subspaces spanned by  $\mathbf{A}$  and  $\hat{\Phi}_s$ , where  $r$  is the minimum of the ranks of  $\mathbf{A}$  and  $\hat{\Phi}_s$ . We use the notation  $\text{subcorr}\{\mathbf{A}, \hat{\Phi}_s\}_i$  to denote the  $i$ th correlation  $c_i$ .

Summarizing, we have shown a sequence of steps which relate least-squares to modified least-squares to signal subspace fitting. Each of these cost functions represents a somewhat different approach to estimating the unknown parameters  $\{\rho, \theta\}$  in our spatio-temporal model. By substituting (27) into each cost function, each estimation technique can be viewed as a different method of weighting and summing the subspace correlations between the model matrix  $\mathbf{A}(\rho, \theta)$  and the signal subspace estimate. More formal derivation and analysis of these various cost functions in relation to the general array signal processing problem may be found in [24], [25], and [28]–[30]. We will now use the novel framework of subspace correlations to redevelop the MUSIC algorithm and introduce our new variant, R-MUSIC.

### C. MUSIC

The least-squares and SSF methods reviewed above require nonlinear multidimensional searches to find the unknown parameters  $\{\rho, \theta\}$ . MUSIC was introduced by Schmidt [19] as a means to reduce the complexity of this nonlinear search. Here we review MUSIC in terms of the subspace correlations, which in turn leads to our proposed R-MUSIC approach.

Given that the rank of  $\mathbf{A}(\rho, \theta)$  is  $r$  and the rank of  $\hat{\Phi}_s$  is at least  $r$ , the smallest subspace correlation value,  $c_r \equiv \text{subcorr}\{\mathbf{A}(\rho, \theta), \hat{\Phi}_s\}_r$ , represents the minimum subspace correlation (maximum principal angle) between principal vectors in the column space of  $\mathbf{A}(\rho, \theta)$  and the signal subspace  $\hat{\Phi}_s$ . The subspace correlation of any individual column  $\mathbf{a}(\rho_i, \theta_i)$  with the signal subspace must therefore equal or exceed this minimum subspace correlation

$$\text{subcorr}\{\mathbf{a}(\rho_i, \theta_i), \hat{\Phi}_s\}_1 \geq c_r, \quad i = 1, \dots, r. \quad (31)$$

As the quality of our signal subspace estimate improves (either by improved signal to noise ratios (SNR's) or longer data acquisition), then  $\hat{\Phi}_s$  will approach  $\Phi_s$  and the minimum correlation approaches unity when the correct parameter set  $\{\rho, \theta\}$  is identified, such that the  $r$  distinct sets of parameters  $\{\rho_i, \theta_i\}$  have subspace correlations approaching unity. Thus, a search strategy for identifying the parameter set  $\{\rho, \theta\} = \{\{\rho_1, \theta_1\}, \dots, \{\rho_r, \theta_r\}\}$  is to identify  $r$  peaks of the metric

$$\text{subcorr}^2\{\mathbf{a}(\rho, \theta), \hat{\Phi}_s\}_1 = \frac{\mathbf{a}^T(\rho, \theta) \hat{\Phi}_s \hat{\Phi}_s^T \mathbf{a}(\rho, \theta)}{\|\mathbf{a}(\rho, \theta)\|^2} \quad (32)$$

where the squared subcorr operation is readily equated with the right hand side, since the first argument is a vector and

the second argument is already a matrix with orthonormal columns. We recognize this as the MUSIC metric [19], with the minor difference of using the signal subspace projector  $\hat{\Phi}_s \hat{\Phi}_s^T$  rather than the more commonly used noise-only subspace projector  $\hat{\Phi}_e \hat{\Phi}_e^T$ . If our estimate of the signal subspace is perfect, then we will find  $r$  global maxima equal to unity.

#### D. Quasilinear Solution

Before proceeding to a description of MUSIC, we first address the problem of finding the orientation vector  $\mathbf{u}_i$ . The dipole parameters in the set  $\{\rho_i, \mathbf{u}_i\}$  are chosen to maximize  $\text{subcorr}\{\mathbf{a}(\rho_i, \mathbf{u}_i), \hat{\Phi}_s\}_1$ , implying that we must search over all possible combinations of location and orientation. The parameter vector  $\mathbf{u}_i$ , however, represents a linear combination of the columns of the gain matrix, such that  $\mathbf{a}(\rho_i, \mathbf{u}_i) = \mathbf{G}(\rho_i)\mathbf{u}_i$  [see (7) and preceding discussion]. As described in the Appendix, the first subspace correlation value  $\text{subcorr}\{\mathbf{G}(\rho_i), \hat{\Phi}_s\}_1$  implicitly gives us the best way of combining the columns of  $\mathbf{G}(\rho_i)$ . We, therefore, first find the dipole parameters  $\rho_i$  which maximize  $\text{subcorr}\{\mathbf{G}(\rho_i), \hat{\Phi}_s\}_1$ . As described in the Appendix, we then readily extract the corresponding quasilinear  $\mathbf{u}_i$  which maximize this subspace correlation. This avoids explicitly searching for these quasilinear parameters, reducing the overall complexity of the nonlinear search.

#### E. Classical MUSIC

In [12], we adapted a “diversely polarized” form of Schmidt’s original MUSIC algorithm [6], [19] to the problem of multiple point dipoles. We briefly review and update that presentation here to include our discussion of subspace correlations. The steps are as follows.

- 1) Obtain a spatio-temporal data matrix  $\mathbf{F}$ , comprising information from  $m$  sensors and  $n$  time slices. Decompose  $\mathbf{F}$  or  $\mathbf{F}\mathbf{F}^T$  and select the rank of the signal subspace to obtain  $\hat{\Phi}_s$ . Overspecifying the true rank by a couple of dimensions usually has little effect on performance. Underspecifying the rank can dramatically reduce the performance.
- 2) Create a relatively dense grid of dipolar source locations. At each grid point, form the gain matrix  $\mathbf{G}$  for the dipole. At each grid point, calculate the subspace correlations  $\text{subcorr}\{\mathbf{G}, \hat{\Phi}_s\}$ .
- 3) As a graphical aid, plot the inverse of  $\sqrt{1 - c_1^2}$ , where  $c_1$  is the maximum subspace correlation. Correlations close to unity will exhibit sharp peaks. Locate  $r$  or fewer peaks in the grid. At each peak, refine the search grid to improve the location accuracy, and check the second subspace correlation. A large second subspace correlation is an indication of a “rotating dipole.”

### IV. R-MUSIC

Problems with the use of MUSIC arise when there are errors in the estimate of the signal subspace and the subspace correlation is computed at only a finite set of grid points. The

largest peak is easily located by searching over the grid for the largest correlation; however, the second and subsequent peaks must be located by means of a 3-D “peak-picking” routine. The R-MUSIC method overcomes this problem by recursively building up the IT model and comparing this full model to the signal subspace.

#### A. Development

In the following we assume that our IT’s each comprise one or more dipoles. We search first for the single dipolar topographies, then the two-dipolar topographies, and so forth. As we discover each topography model, we add it to our existing IT model and continue the search. We build the source model by recursively applying the subspace correlation measure, the key metric of MUSIC, to successive subspace correlations.

For exemplary purposes, we first assume that the  $r$  IT’s each comprise a single dipole. Single-dipole locations are readily found by scanning the head volume. At each point in the volume, we calculate

$$\{c_1, c_2, \dots\} = \text{subcorr}\{\mathbf{G}(\mathbf{r}_q), \hat{\Phi}_s\} \quad (33)$$

where  $\{c_1, c_2, \dots\}$  is the set of subspace correlations. We find the dipole location  $\hat{\mathbf{r}}_{q1}$ , which maximizes the subspace correlation  $c_1$ . As described in the Appendix, the corresponding dipole orientation  $\hat{\mathbf{u}}_1$  is easily obtained from  $\text{subcorr}\{\mathbf{G}(\hat{\mathbf{r}}_{q1}), \hat{\Phi}_s\}$ , and we designate our topography model comprising this first dipole as

$$\hat{\mathbf{A}}^{(1)} = \mathbf{a}(\hat{\mathbf{r}}_{q1}, \hat{\mathbf{u}}_1). \quad (34)$$

To search for the second dipole, we again search the head volume; however, at each point in the head, we first form the model matrix  $\mathbf{M} = [\hat{\mathbf{A}}^{(1)}, \mathbf{G}(\mathbf{r}_q)]$ . We then calculate

$$\{c_1, c_2, \dots\} = \text{subcorr}\{\mathbf{M}, \hat{\Phi}_s\} \quad (35)$$

but now we find the dipole point that maximizes the *second* subspace correlation,  $c_2$ ; the first subspace correlation should already account for  $\mathbf{a}(\hat{\mathbf{r}}_{q1}, \hat{\mathbf{u}}_1)$  in the model. The corresponding dipole orientation  $\hat{\mathbf{u}}_2$  may be readily obtained by projecting this second topography against the subspace,  $\text{subcorr}\{\mathbf{G}(\hat{\mathbf{r}}_{q2}), \hat{\Phi}_s\}$ , and we append this to our model to form

$$\hat{\mathbf{A}}^{(2)} = [\mathbf{a}(\hat{\mathbf{r}}_{q1}, \hat{\mathbf{u}}_1), \mathbf{a}(\hat{\mathbf{r}}_{q2}, \hat{\mathbf{u}}_2)]. \quad (36)$$

We repeat the process  $r$  times, maximizing the  $k$ th subspace correlation at the  $k$ th pass,  $k = 1, \dots, r$ . The final iteration is effectively attempting to minimize the subspace “distance” [8] between the full  $r$  topographies matrix and the signal subspace estimate.

If the  $r$  topographies comprise  $r_1$  single-dipolar topographies and  $r_2$  2-dipolar topographies, then R-MUSIC will first extract the  $r_1$  single dipolar models. At the  $(r_1 + 1)$ th iteration, we will find no single dipole location that correlates well with the subspace. We then increase the number of dipole elements

per topography to two. We must now search simultaneously for two dipole locations, such that

$$\{c_1, c_2, \dots\} = \text{subcorr}\left\{\left[\hat{\mathbf{A}}^{(r_1)}, \mathbf{G}(\rho)\right], \hat{\Phi}_s\right\} \quad (37)$$

is maximized for the subspace correlation  $c_{r_1+1}$ , where  $\rho = \{\mathbf{r}_{q1}, \mathbf{r}_{q2}\}$  comprises two dipoles. If the combinatorics are not impractical, we can exhaustively form all pairs on our grid and compute maximum subspace correlations for each pair. The alternative is to begin a two-dipole nonlinear search with random initialization points to maximize this correlation (cf. [9]). This low-order dipole search can be easily performed using standard minimization methods.

We proceed in this manner to build the remaining  $r_2$  2-dipolar topographies. As each pair of dipoles is found to maximize the appropriate subspace correlation, the corresponding pair of dipole orientations may be readily obtained from  $\text{subcorr}\{\mathbf{G}(\rho), \hat{\Phi}_s\}$ , as described in the Appendix. Extensions to more dipoles per IT are straightforward, although the complexity of the search obviously increases. In any event, the complexity of the search will always remain less than or equal to the least-squares search required for finding all dipoles simultaneously.

### B. Algorithm

To summarize, we assume that our forward model has been corrupted by additive noise, and that this noise is zero mean with a known spatial covariance matrix  $\sigma_c^2 \mathbf{I}$ . We decompose  $\mathbf{F}$  or  $\mathbf{F}\mathbf{F}^T$  and select the rank  $r$  of the signal subspace to form  $\hat{\Phi}_s$ , which is our estimate of a set of vectors that span the signal subspace. If the rank is uncertain, we should err toward overspecifying the signal subspace rank. If we overselect the rank, the additional subspace vectors should span an arbitrary subspace of the noise-only subspace, and the probability that these vectors correlate with our model is small. However, as the overspecification of the signal subspace increases, so does the probability that we may inadvertently include a noise-only subspace that correlates with our models, so some prudence is called for in rank selection. We demonstrate examples of overselection of the subspace in Section V.

We design a sufficiently dense grid in our volume of interest, and at each grid point  $\mathbf{l}_i$  we form the head model for the single-dipole gain matrix  $\mathbf{G}(\mathbf{l}_i)$ . We initialize the topography complexity as “1-dipolar topography,” i.e., each topography comprises a single dipole. We then proceed as follows:

- 1) For *index* from 1 to rank  $r$ :
- 2) Let  $\hat{\mathbf{A}} = [\mathbf{a}_1, \dots, \mathbf{a}_{(\text{index}-1)}]$  be the model extracted as of the previous loop ( $\hat{\mathbf{A}}$  is a null matrix for the first loop).
- 3) Form sets of grid points  $\lambda_i$ , where for a 1-dipolar topography each set consists of the location of a single grid point  $\mathbf{l}_i$ . For a 2-dipolar topography,  $\lambda_i$  contains the locations of pairs of grid points, and so on for higher order dipolar topographies. If the combinatorics make it impractical to consider all possible combinations of grid points, choose a random subset of the possible combinations.

- 4) For each set of grid points  $\lambda_i$ , form the grid model  $\mathbf{M}_i = [\hat{\mathbf{A}}, \mathbf{G}(\lambda_i)]$ , i.e., concatenate the set of grid point models to the present extracted model.
- 5) Calculate the set of subspace correlations,  $\{c_1, c_2, \dots\} = \text{subcorr}\{\mathbf{M}_i, \hat{\Phi}_s\}$ , using the algorithm described in the Appendix.
- 6) Find the maximum over all sets of grid points  $\lambda_i$  for  $c_{\text{index}}$ , e.g., for *index* = 2, find the maximum *second* subspace correlation.
- 7) Optionally, if the set of grid points  $\{\lambda_i\}$  is not particularly dense or complete, then use a nonlinear optimization method (e.g., Nelder–Meade simplex) to maximize  $c_{\text{index}}$ , beginning the optimization at the best  $\lambda_i$ . If the grid is dense and our sets in Step 2) complete, this step may not be necessary.
- 8) Is the correlation at the location of the maximum “sufficient,” i.e., does  $c_{\text{index}}$  indicate a good correlation? If the correlation is adequate, proceed to Step 11). If it is not, proceed to Step 9).
- 9) [Insufficient correlation in Step 8)] We have two situations to consider. We may have overspecified the true rank of the signal subspace, in which case we are now attempting to fit a topography into a noise-only subspace component. We can test for this condition by forming the projection operator  $\mathbf{P}_{\hat{\mathbf{A}}} = \hat{\mathbf{A}}\hat{\mathbf{A}}^\dagger$  (where  $\hat{\mathbf{A}}^\dagger$  is the *pseudoinverse* [12]) from the existing estimated model, then forming the residual  $\mathbf{F}_{\text{res}} = \mathbf{F} - \mathbf{P}_{\hat{\mathbf{A}}}\mathbf{F}$ . Inspection and testing of the residual should reveal whether or not we believe a signal is still present. If we believe the residual is simply “noise,” break this loop. Otherwise, proceed to Step 10).
- 10) (Signal still apparent in the residual) Increase the complexity of the topography (e.g., from one to two dipolar) and return to Step 3) without increasing the loop *index*.
- 11) [Good correlation in Step 8)] We have found the best set of locations  $\hat{\rho}_{\text{index}}$  of the next independent topography, with corresponding gain matrix  $\mathbf{G}_{\text{index}}(\hat{\rho}_{\text{index}})$ . We need the best-fitting orientation. Calculate the principal orientation vector  $\mathbf{x}_1$  (see the Appendix) from  $\text{subcorr}\{\mathbf{G}_{\text{index}}(\hat{\rho}_{\text{index}}), \hat{\Phi}_s\}$ , normalize  $\hat{\mathbf{u}}_{\text{index}} = \mathbf{x}_1 / \|\mathbf{x}_1\|$ , and form the topography vector  $\hat{\mathbf{a}}_{\text{index}} = \mathbf{G}_{\text{index}}(\hat{\rho}_{\text{index}})\hat{\mathbf{u}}_{\text{index}}$ .
- 12) Increment the *index* and loop to Step 1) for the next independent topography.

In Step 8), we have used a minimum correlation of 95%. In [14], we discussed some of the means for determining if a MUSIC peak represents “adequate” or “sufficient” correlation. Our recommendation of 95% reflects the empiricism that a “good” solution should generate a topography which explains at least 90% (the square of the correlation, i.e., the “R-squared” statistic) of the variance of the topography identified in the data. If we overselect the rank of the signal subspace, then we will in general break out of the loops at Step 9), once we have found the true number of sources and have only noise left in the residual. We will not address the determination of statistical “sufficiency” of

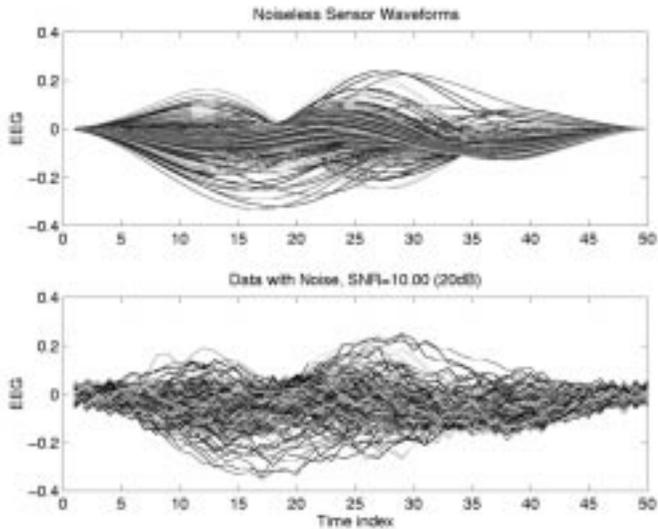


Fig. 1. The upper plot is the overlay of the response of all 255 EEG sensors to the simulated three dipolar sources. The sources were given independent, overlapping time series. White Gaussian noise was added such that the SNR was 20 dB, and the percent variance explained by the true solution was 91% of the total variance.

the model in this paper. See [1] and [26] for discussions on the testing of the residual for remnant signals.

If the grid is dense or we performed Step 7) for each topography, we may find the R-MUSIC set of parameters is already a good solution. The R-MUSIC algorithm has maximized a set of subspace correlations, a metric different from the least-squares approach. We may refine this solution with a least-squares search:

- 13) Our R-MUSIC search has yielded an estimate of the full spatial topographies gain matrix  $\hat{\mathbf{A}} = [\mathbf{a}_1, \dots, \mathbf{a}_r]$ , which is a function of the estimated full set of dipole locations  $\hat{\rho}$  and orientations  $\hat{\theta}$ . Beginning with these parameters, initialize a nonlinear search using the cost functions (24) or (25).

Step 13) represents an increase in the complexity in the nonlinear search over that of R-MUSIC, at possibly diminishing returns in terms of improvement in the solution. Note that each iteration of the nonlinear search must now adjust the parameters of all of the dipoles, not just a single topography as in R-MUSIC.

## V. COMPUTER SIMULATIONS

We present two simulations to illustrate some of the features of our proposed IT model and the R-MUSIC algorithm. In the first simulation, we arranged 255 EEG sensors about the upper region of an 8.8-cm single-shell sphere, with a nominal spacing between sensors of 1 cm. For illustrative purposes, we arranged three dipolar sources in the same plane,  $z = 7$  cm, and the three sources were given independent, overlapping time courses. The overlay of the responses of all sensors is given in the upper plot of Fig. 1. We then added white Gaussian noise to all data points, scaled such that the squared Frobenius norm of the noise matrix was one-tenth that of the

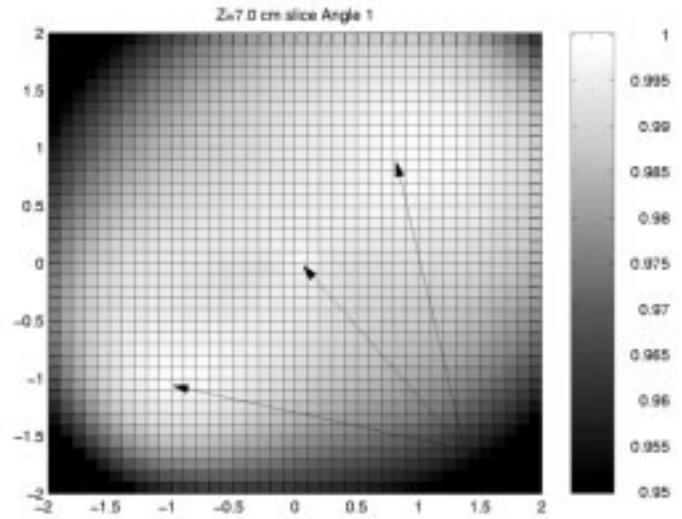


Fig. 2. Subspace correlation between an EEG dipolar model and the rank five signal subspace extracted from the data. The correlations were calculated in the  $z = 7$ -cm plane on a 1-mm grid. Each grid point was then scaled in intensity using the color bar on the right-hand side of the figure. All correlations below 95% are scaled as black. The largest correlation of 99.8% is correctly at  $[-1, -1, 7]$  cm, as rounded to the nearest mm; however, this peak and the other two peaks, as indicated by the arrows, are not readily discernible, either graphically or computationally.

squared Frobenius norm of the noiseless signal matrix, for an SNR of 20 dB. The lower plot of Fig. 1 shows the overlay of all sensors for the signal plus noise data.

The singular value spectrum was clearly rank three, but we selected rank five to illustrate robustness to rank over-selection. We created a 1-mm grid in the  $z = 7$ -cm plane and calculated the correlation between a single-dipole model and the signal subspace. Fig. 2 displays these correlations as an image whose intensities are proportional to the primary correlation  $c_1$ . We have defined the gray scale in Fig. 2 such that the subspace correlation must exceed 95% in order to be visible. In Fig. 3, we have replotted the same data, but in this case we plot  $1/(\sqrt{1 - c_1^2})$  in order to graphically intensify the appearance of the peaks. This image is the original MUSIC scan proposed in [12]. The measure  $\sqrt{1 - c_1^2}$  is equivalent to the correlation with the noise-only subspace, the original proposal by Schmidt [19].

The largest subspace correlation of 99.8% is easily found at  $[-1, -1, 7]$  cm, as rounded to the nearest mm. The peak at  $[1, 1, 7]$  is apparent in Fig. 3, but the peak at  $[0, 0, 7]$  is obscured in both figures (see caption). Graphically or computationally declaring the location of these other two peaks is not obvious without subjective interpretation by the observer.

We generated the forward field for this first dipole, then rescanned the subspace correlation on the same grid with the combined model. Fig. 4 displays the *second* subspace correlation; in this and subsequent figures, we will resume plotting the correlation value directly, rather than the inverted metric. We can now more clearly see the peaks corresponding to the two remaining sources, and the first source has been suppressed. The maximum peak of this image at 99.7% is easily located at  $[1, 1, 7]$ . We then generated the forward field

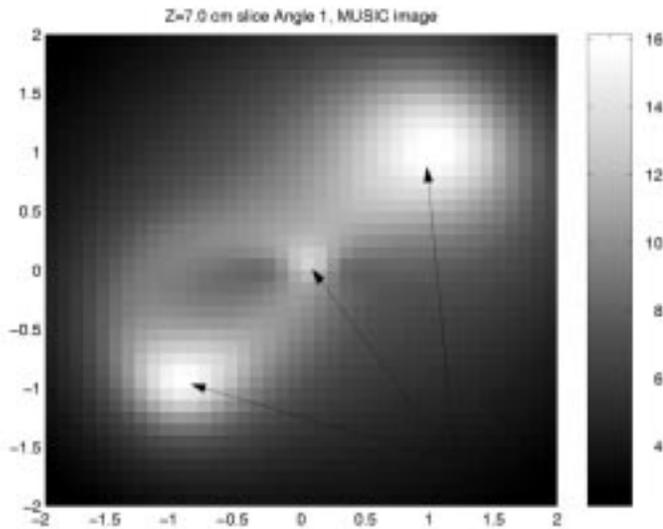


Fig. 3. Rather than plot each pixel as a scaled version of a correlation value as in Fig. 2, here we plot  $1/(\sqrt{1-c_1^2})$ . This image metric is equivalent to the MUSIC metric discussed in [12], [19]. We see two of the three peaks rather clearly, but the central peak is somewhat obscured, since its graphical "intensity" is about 75% of the intensity of the other two peaks. Although the peaks are better defined than in Fig. 2, interpretation of the intensity scale is now ambiguous, and the user must still "peak pick" graphically or algorithmically in three dimensions. Subsequent figures resume the correlation scaling of Fig. 2.

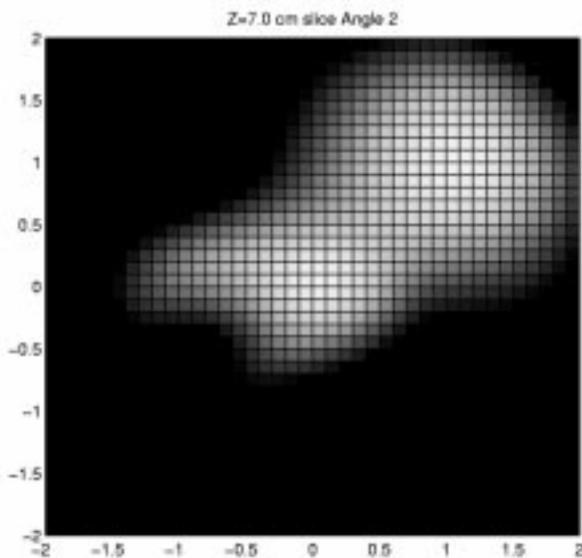


Fig. 4. The forward field from the first dipolar solution was formed and concatenated with the forward field generated by each point on this grid. The subspace correlations were again computed between this combined model and the signal subspace. The *second* subspace correlation is displayed here as an image. The original peak has been suppressed, and we more clearly see two peaks in this image. The maximum correlation in this images is found at [1, 1, 7] of 99.7%.

for this second dipole and appended it to the first dipole's forward field. We then rescanned the subspace correlations on the same grid with the combined model. Fig. 5 displays the *third* subspace correlation, where we now readily observe the single remaining peak for the third source, 99.6% at [0, 0, 7].

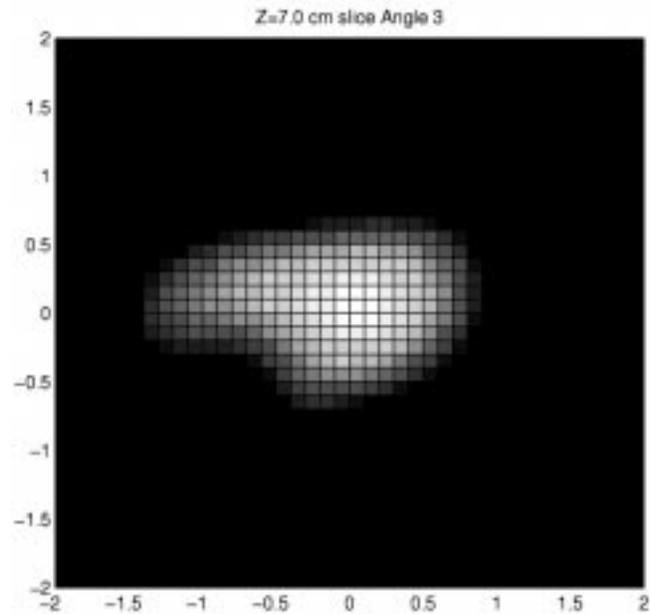


Fig. 5. We generated the forward fields for the first two dipolar locations and concatenated that set with the field generated by each point in the grid. We then re-ran the subspace correlations, then imaged the *third* subspace correlation. The remaining source solution is now easily observed at [0, 0, 7], with a correlation of 99.6%. The percent variance explained by the combined three dipoles was 91%.

Visual examination of the residual at this point indicated no remaining signal, and subspace correlations of multiple-dipole models yielded no substantial correlations.

The second simulation was designed to demonstrate the localization of a "rotating" dipole and a pair of synchronous dipoles, as well as to illustrate the use of a directed search algorithm to refine these locations. In this simulation, we arranged 240 MEG planar gradiometer sensors about the upper hemisphere, with a nominal spacing of about 2 cm and a baseline separation of 1 cm. A "rotating" dipole was located at [0, 0, 7] cm, and a pair of dipoles with synchronous activation was located at [-2, -2, 7] and [2, 2, 7] cm. We then created a 1.5-mm grid in the  $z = 6.5$ -cm plane, i.e., in a plane displaced from the true source plane, and the gridding was slightly coarser than the first simulation. The noise level was again set to 20 dB. The true rank of the signal subspace was three, with the rotating dipole comprising two single-dipolar topographies, and the third topography comprising a 2-dipolar topography. Fig. 6 displays the overlay of the noiseless and noisy sensor responses.

We again overselected the rank of the signal subspace to be five, then scanned the one-dipole model against the signal subspace. We found a single good peak at 99.3%, as displayed in Fig. 7. Note the absence of any other peaks; the remaining "rotating" dipolar topography is obscured by this peak, and the other topography is not a single dipole. The peak observed in the grid was at [0.1, -0.2, 6.5] cm. We initiated a directed search from this point to maximize the correlation to 99.8% at [0.0, 0.0, 7.0] cm, the correct solution for the single-dipole topography, rounded to the nearest mm.

As in the previous example, we then scanned for a second dipole, observing the second subspace correlation. The maxi-

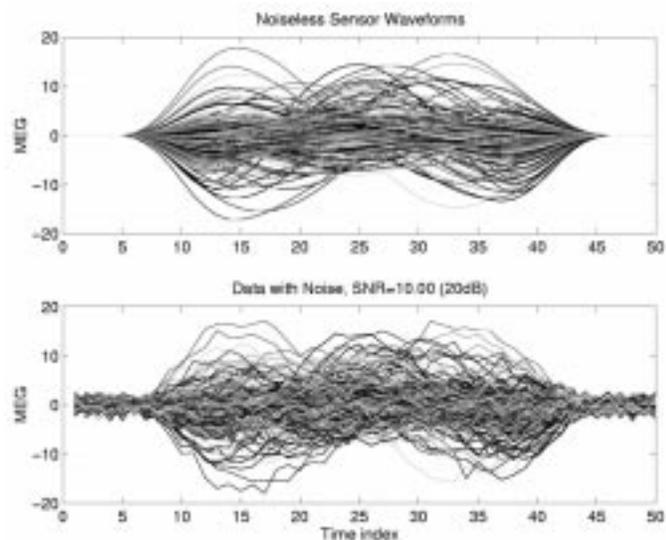


Fig. 6. MEG Simulation comprising 240 planar gradiometers with a 1-cm baseline spaced about 2 cm apart, arranged about the upper region of a 12-cm sphere. A rotating dipole was located at  $[-2, -2, 7]$  cm, and a pair of dipoles with synchronous activation was located at and  $[2, 2, 7]$  cm. The white Gaussian noise was scaled such that the squared Frobenius norm of the signal was ten times that of the noise (20 dB). The true rank of the signal subspace was three, but we again overselected it to be five.

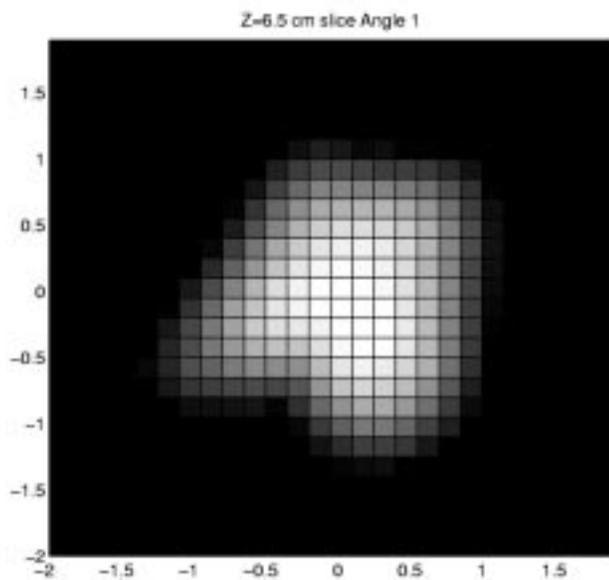


Fig. 7. Subspace correlation between the 1-dipole model and the rank five MEG subspace. A single peak is observed at the location of the rotating dipole; the second dipolar source is effectively buried by this peak. This plane was scanned at  $z = 6.5$  cm, but the true solution lay above this scanning plane. The peak in this plane was used to initiate a directed search for the maximum correlation, which was located correctly at  $[0, 0, 7]$  cm, as rounded to the nearest mm.

imum correlation in the grid was again high, 99.2%, at  $[-0.1, -0.2, 6.5]$ , as shown in Fig. 8. A directed search initiated at this point maximized the second subspace correlation at 99.7% at  $[0, 0, 7]$ , the same dipole location as the first solution. The dipole orientations of the two solutions were nearly orthogonal,  $[0.94, 0.34, 0]$  and  $[-0.3, 0.95, 0]$ , indicating we had correctly identified the simulated “rotating” dipole at this point.

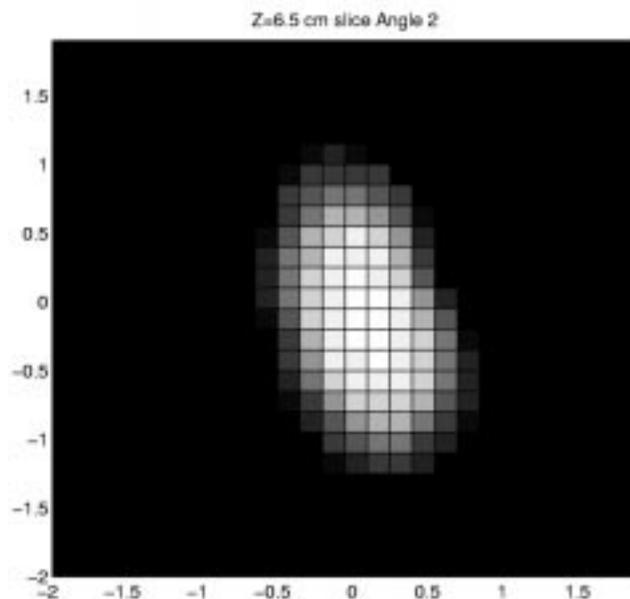


Fig. 8. Second subspace correlation between the 1-dipole model and the rank five MEG subspace, using the solution from Fig.7. The peak in this plane was used to initiate a directed search for the maximum second correlation, which was located correctly, again at  $[0, 0, 7]$  cm as in Fig.7. We have, thus, located a “rotating” dipole. The remaining IT source was not a single dipole and, thus, not observed in these single-dipole scans.

We then scanned for a third single-dipole solution, but only a peak of 88.8% was found, and a directed search maximization only improved this correlation to 88.9%. Thus this third dipole could only account for  $(88.9\%)^2 = 79.0\%$  of the variance of the third topography, and we rejected this third single-dipolar topography solution.

Since one dipole was inadequate to describe the third topography, we shifted to our next putative solution, that of two dipoles. Our grid comprised 729 dipole locations, and all combinations of two dipoles yielded 265 356 sets. Rather than exhaustively search all set combinations, we randomly selected a small subset for a total of about 3000 sets. We then concatenated each of these 3000 pairs with the first two-dipole solutions, calculated the subspace correlation of the combined model and observed the third subspace correlation. The maximum third correlation of 98.4% corresponded to the pair at  $[-2, -1.9, 6.5]$ ,  $[1.8, 1.8, 6.5]$ . We initiated a 2-dipole directed search from this set and achieved a maximum correlation of 99.7% at  $[-2.0, -2.0, 7.0]$ ,  $[2.0, 2.0, 7.0]$  cm, the correct solution. In Fig. 9, we plot 2-D cross-slices of this 6-D function, holding constant the correct  $z$  plane and the true location of one of the two dipoles. We clearly observe the correlation metric peaking at the correct solution. As in the first example, visual examination of the residual from this model revealed that no signal was present, and further correlations with multiple-dipole models yielded no substantial correlations.

This relatively simple pair of simulations has illustrated some of the key concepts of the R-MUSIC algorithm and the IT model. Both simulations used relatively dense grids of EEG or MEG sensors, such that sensor spacing was not an issue; see [13] for analysis of the effects of EEG

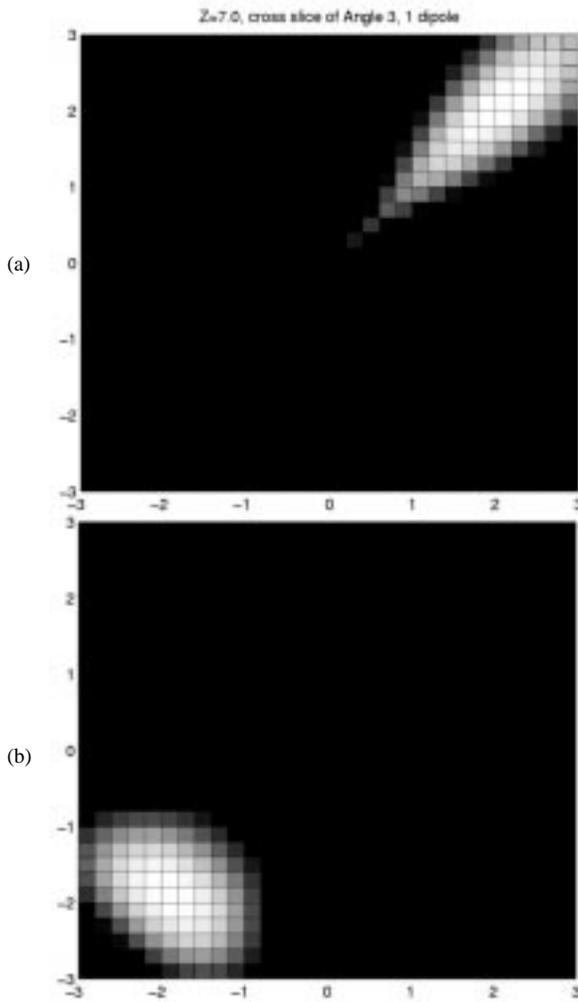


Fig. 9. Third subspace correlation of the MEG simulation, computed in a 6-D space comprising the synchronous two-dipole topography. Pairs of dipoles were concatenated to the model identified in the first two topographies, then the third subspace correlation computed. (a) is a 2-D slice through this 6-D space, holding fixed the true location of one of the dipoles and the true  $z = 7$ -cm value of the second dipole. We observe the correlation peaking correctly at  $[2, 2, 7]$ . (b) holds fixed this dipole and the same value, and peaks correctly at the other dipole  $[-2, -2, 7]$ .

and MEG sensor spacing on dipole localization performance. In both simulations, we overselected the true rank of the signal subspace to illustrate the robustness to such an error; we repeated the localization results with the true rank and achieved nearly identical results to those presented here. In these simulations, as in [12], the subspace scans were presented as images to highlight the MUSIC peaks; however, the R-MUSIC algorithm readily extracts these peaks without the need for the user to manually observe and select these solutions. A “blind test” simulation study using R-MUSIC can be found in [11].

In practice, after we have scanned on a discrete grid for any of the single- or multiple-dipolar solutions, we always then initiate a directed search from these points to maximize the correlation. By optimizing the correlation in this manner, we bypass some of the concerns of coarse or inadequate gridding. In the first simulation, each of the three dipoles was located with a single-dipole search of three location parameters; in

contrast, a full nonlinear least-squares would have required nine parameters. In the second simulation, we performed two single-dipole searches, followed by a two-dipole search of six nonlinear parameters. A full nonlinear least-squares search would have required a 12-parameter search.

## VI. CONCLUSION

In the E/MEG inverse problem, our goal is to estimate a set of parameters that represent our source. In the multiple-dipole model, an issue that complicates the least-squares problem is that it requires a multidimensional search over a highly nonconvex cost function. Here we have described a new algorithm, R-MUSIC, which uses subspace correlations between the model subspace and the data subspace to reduce the problem to a sequential search. By identifying one source at a time we reduce the computational complexity of the search. The search for topographies comprising a few dipoles can then be performed over the entire source volume, largely avoiding the local minima problem.

The IT model that was also presented here is a new framework in which to view the concept of a source. We often encounter dipolar sources that are effectively fully correlated in their time courses due to either bisynchronous activation or strong noise. The IT model allows a straightforward interpretation of these correlated dipoles as a single source topography comprising multiple dipoles. Combining the R-MUSIC method with the IT source model keeps the complexity of the parameter search simple relative to more traditional multidimensional cost functions while bypassing the “peak-picking” problem of the classical MUSIC algorithm. While determining multiple peaks in a single parameter case (the common presentation in much of the array signal processing literature on MUSIC) is possible, we found the problem confounding in even our simplest case of single-dipolar topographies, where we must search for peaks in three dimensions. Graphically searching for multiple peaks in two-dipolar topographies (a 6-D space) is generally not practical.

In this paper, we have used multiple dipoles as our source model, increasing the independent topography complexity by simply increasing the number of synchronous dipoles. The IT model and R-MUSIC algorithm are readily extended to include source models that can represent more distributed current activity, as we will address in a future publication.

## APPENDIX SUBSPACE CORRELATION

### A. Definitions and Computation

From [8], we summarize the definition and method for computation of the canonical or “subspace” correlation. Given two matrices,  $\mathbf{A}$  and  $\mathbf{B}$ , where  $\mathbf{A}$  is  $m \times p$ , and  $\mathbf{B}$  is  $m \times q$ , let  $r$  be the minimum of the ranks of the two matrices. We wish to calculate a function  $\{c_1, c_2, \dots, c_r\} = \text{subcorr}\{\mathbf{A}, \mathbf{B}\}$ , where the set of scalars  $c_k$  are defined as follows:

$$c_k = \max_{\mathbf{a} \in \mathbf{A}} \max_{\mathbf{b} \in \mathbf{B}} \mathbf{a}^T \mathbf{b} = \mathbf{a}_k^T \mathbf{b}_k \quad (38)$$

subject to

$$\begin{aligned} \|\mathbf{a}\| &= \|\mathbf{b}\| = 1 \\ \mathbf{a}^T \mathbf{a}_i &= 0, \quad i = 1, \dots, k-1 \\ \mathbf{b}^T \mathbf{b}_i &= 0, \quad i = 1, \dots, k-1. \end{aligned} \quad (39)$$

The vectors  $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$  and  $\{\mathbf{b}_1, \dots, \mathbf{b}_r\}$  are the *principal vectors* between the subspaces spanned by  $\mathbf{A}$  and  $\mathbf{B}$ , and by construction, each set of vectors represents an orthonormal basis. Note that  $1 \geq c_1 \geq c_2 \geq \dots \geq c_r \geq 0$ . The angles  $\theta_k$ , where  $\cos \theta_k = c_k$ , are the *principal angles*, representing the geometric angle between  $\mathbf{a}_k$  and  $\mathbf{b}_k$ , or analogously,  $c_k$  is the *subspace correlation* between these two vectors. The steps to compute the subspace correlations are as follows [8, p. 585].

- 1) If  $\mathbf{A}$  and  $\mathbf{B}$  are already orthogonal matrices, we redesignate them as  $\mathbf{U}_A$  and  $\mathbf{U}_B$  and skip to Step 2). Otherwise, perform a SVD of  $\mathbf{A}$ , such that  $\mathbf{A} = \mathbf{U}_A \Sigma_A \mathbf{V}_A^T$ . Similarly decompose  $\mathbf{B}$ . Retain only those components of  $\mathbf{U}_A$  and  $\mathbf{U}_B$  that correspond to nonzero singular values, i.e., the number of columns in  $\mathbf{U}_A$  and  $\mathbf{U}_B$  correspond to their ranks, and the other matrices are square, with dimension equal to the ranks.
- 2) Form  $\mathbf{C} = \mathbf{U}_A^T \mathbf{U}_B$ .
- 3) If only the correlations are desired, then compute only the singular values of  $\mathbf{C}$  (the extra computation for the singular vectors is not required). The  $r$  ordered singular values  $1 \geq c_1 \geq c_2 \geq \dots \geq c_r \geq 0$  are the subspace correlations between  $\mathbf{A}$  and  $\mathbf{B}$ .
- 4) If the principal vectors are also desired, then compute the full SVD,  $\mathbf{C} = \mathbf{U}_C \Sigma_C \mathbf{V}_C^T$ . The  $r$  ordered singular values are extracted from the diagonal of  $\Sigma_C$ . Form the sets of principal vectors  $\mathbf{U}_a = \mathbf{U}_A \mathbf{U}_C$  and  $\mathbf{U}_b = \mathbf{U}_B \mathbf{V}_C$  for sets  $\mathbf{A}$  and  $\mathbf{B}$ , respectively.

The matrices  $\mathbf{U}_a$  and  $\mathbf{U}_b$  are each orthogonal, and the columns comprise the ordered sets of principal vectors for matrices  $\mathbf{A}$  and  $\mathbf{B}$ , respectively. If both matrices are of the same subspace dimension, the measure  $\sqrt{1 - c_r^2} = \sin \theta_r$  is called the *distance* between spaces spanned by  $\mathbf{A}$  and  $\mathbf{B}$  [8]. When the distance is zero, we see that  $\mathbf{A}$  and  $\mathbf{B}$  are parallel subspaces. A maximum distance of unity ( $c_r = 0$ ) indicates at least one basis of  $\mathbf{A}$  is orthogonal to  $\mathbf{B}$  or vice versa; if the maximum subspace correlation is  $c_1 = 0$ , then all bases are orthogonal. We see that minimizing the distance is equivalent to maximizing the minimum subspace correlation between  $\mathbf{A}$  and  $\mathbf{B}$ .

We may also readily compute the specific linear combinations of  $\mathbf{A}$  and  $\mathbf{B}$  that yielded these principal vectors and angles. By construction, we know that  $\mathbf{A}\mathbf{X} = \mathbf{U}_a$  for some  $\mathbf{X}$ , and  $\mathbf{X}$  can be simply found using the pseudoinverse of  $\mathbf{A}$ . If we have used the SVD to decompose  $\mathbf{A}$ , then the calculation of  $\mathbf{X}$  reduces to  $\mathbf{X} = \mathbf{V}_A \Sigma_A^{-1} \mathbf{U}_C$ ; similarly, we compute  $\mathbf{Y} = \mathbf{V}_B \Sigma_B^{-1} \mathbf{V}_C$ .

The best way to linearly combine the columns of  $\mathbf{A}$  (i.e., the combination that minimizes the angle of the resulting vector with  $\mathbf{B}$ ) is found in the first column of  $\mathbf{X} \equiv [\mathbf{x}_1, \dots, \mathbf{x}_r]$  (similarly define  $\mathbf{Y}$ ),  $\mathbf{a}_1 = \mathbf{A}\mathbf{x}_1$ , which is best correlated with  $\mathbf{B}$  when it is arranged as  $\mathbf{b}_1 = \mathbf{B}\mathbf{y}_1$ . In other words, there is no other  $\mathbf{x}$  (excepting a scale factor of  $\mathbf{x}_1$ ) for which

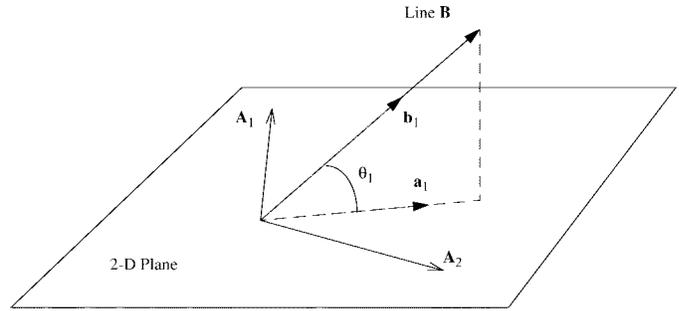


Fig. 10. Geometric interpretation of principal angles. The 2-D plane is spanned by the two columns,  $\mathbf{A}_1$  and  $\mathbf{A}_2$ , of the matrix  $\mathbf{A}$ . These vectors and the vector  $\mathbf{b}$  pass through the coordinate origin. The function  $\text{subcorr}\{\mathbf{A}, \mathbf{B}\}$  returns the *subspace correlation*  $\cos \theta_1$ , where  $\theta_1$  is the *principal angle* between the line and the plane. The principal vectors are  $\mathbf{a}_1$  and  $\mathbf{b}_1$ , which are unit length vectors in the plane and line, respectively.

a corresponding best-fitting  $\mathbf{y}$  will yield a better correlation between  $\mathbf{a}$  and  $\mathbf{b}$ . The first columns of  $\mathbf{U}_a$  and  $\mathbf{U}_b$  are  $\mathbf{a}_1$  and  $\mathbf{b}_1$ .

Similarly, the worst way to linearly combine  $\mathbf{A}$  is  $\mathbf{a}_r = \mathbf{A}\mathbf{x}_r$ . The best fit to this particular  $\mathbf{x}$  is  $\mathbf{b}_r = \mathbf{B}\mathbf{y}_r$ , with a correlation of only  $c_r$ . No other  $\mathbf{x}$  will yield a *best fitting*  $\mathbf{y}$  such that the subspace correlation is *lower*.

If two correlations are identical, for instance  $c_1 = c_2 = 1$ , then the two corresponding vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are, themselves, arbitrary, but they form a plane such that any linear combination of the two vectors yields a vector whose corresponding correlation is  $c_1 = c_2$ .

## B. Geometric Example

To give an intuitive geometric insight into these subspace correlations, consider  $\text{subcorr}\{\mathbf{A}, \mathbf{B}\}$ , where we define the columns of a  $3 \times 2$  matrix  $\mathbf{A}$  to represent two vectors that form a basis for a 2-D plane in a 3-D space. Similarly, let the  $3 \times 1$  vector  $\mathbf{B}$  represent a 1-D vector (line). The subspaces  $\mathbf{A}$  and  $\mathbf{B}$  both pass through the origin. In this case,  $\text{subcorr}\{\mathbf{A}, \mathbf{B}\}$  yields a single correlation coefficient, representing the cosine of the angle between the line and the plane. We can directly form  $\mathbf{a}_1 = \mathbf{A}\mathbf{x}_1$ , which is the unit length vector in the plane of  $\mathbf{A}$  closest to  $\mathbf{B}$ . We illustrate this case in Fig. 10. If the correlation is unity, then  $\mathbf{B}$  lies in the plane of  $\mathbf{A}$ ; if the correlation is zero, then  $\mathbf{B}$  is perpendicular to the plane, and  $\mathbf{x}_1$  is arbitrary.

Next, consider a second 2-D plane spanned by a  $3 \times 2$  matrix  $\mathbf{B}$ , and again the planes formed by the columns of both  $\mathbf{A}$  and  $\mathbf{B}$  pass through the origin. We find that the first (maximum) subspace correlation of  $\text{subcorr}\{\mathbf{A}, \mathbf{B}\}$  is always unity, since two such planes always intersect along a line, namely the line found by  $\mathbf{A}\mathbf{x}_1$  or  $\mathbf{B}\mathbf{y}_1$ . The second subspace correlation is the cosine of the angle between the planes, the angle we intuitively picture when visualizing two intersecting planes.

## C. Subspace Correlations and E/MEG

In E/MEG MUSIC processing, we may compute the subspace correlations between a dipole model and the signal subspace, e.g.,  $\text{subcorr}\{\mathbf{G}(\mathbf{r}_q), \Phi_s\}$ . In this case, the vectors in  $\mathbf{X}$  relate to the dipole orientations. By scaling the first orientation to unity,  $\mathbf{u}_1 \equiv \mathbf{x}_1 / \|\mathbf{x}_1\|$ , we obtain the unit-

dipole orientation that best correlates the dipolar source at  $\mathbf{r}_q$  with the signal subspace. For a two-dipolar topography,  $\text{subcorr}\{\{\mathbf{G}(\mathbf{r}_{q1}), \mathbf{G}(\mathbf{r}_{q2})\}, \Phi_s\}$ , then  $\mathbf{u}_1$  represents the concatenation of the two dipole orientations,  $\mathbf{u}_1^T = [\tilde{\mathbf{q}}_1^T, \tilde{\mathbf{q}}_2^T]$ , such that the two-dipolar topography

$$[\mathbf{G}(\mathbf{r}_{q1}), \mathbf{G}(\mathbf{r}_{q2})]\mathbf{u}_1 = \mathbf{G}(\mathbf{r}_{q1})\tilde{\mathbf{q}}_1 + \mathbf{G}(\mathbf{r}_{q2})\tilde{\mathbf{q}}_2 \quad (40)$$

best correlates with the signal subspace. Consistent with our IT model description, we note that the dipole orientations  $\tilde{\mathbf{q}}_1$  and  $\tilde{\mathbf{q}}_2$  in (40) are themselves not unit vectors, but that their concatenation into the vector  $\mathbf{u}_1$  is constrained to unity norm.

## REFERENCES

- [1] A. Achim, "Signal detection in averaged evoked potentials: Monte Carlo comparison of the sensitivity of different methods," *Electroenceph., Clin. Neurophys.*, vol. 96, pp. 574–584, 1995.
- [2] ———, "Cerebral source localization paradigms: Spatiotemporal source modeling," *Brain, Cognition*, vol. 27, pp. 256–287, 1995.
- [3] A. Achim, F. Richer, and J. Saint-Hilaire, "Methods for separating temporally overlapping sources of neuroelectric data," *Brain Topogr.*, vol. 1, no. 1, pp. 22–28, 1988.
- [4] R. C. Barr, T. C. Pilkington, J. P. Boineau, and M. S. Spach, "Determining surface potentials from current dipoles, with application to electrocardiography," *IEEE Trans. Biomed. Eng.*, vol. BME-23, pp. 88–92, Apr. 1966.
- [5] D. Brenner, J. Lipton, L. Kaufman, and S. J. Williamson, "Somatically evoked magnetic fields of the human brain," *Sci.*, vol. 199, pp. 81–83, 1978.
- [6] E. Ferrara and T. Parks, "Direction finding with an array of antennas having diverse polarizations," *IEEE Trans. Antennas Propagat.*, vol. AP-31, pp. 231–236, Mar. 1983.
- [7] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM J. Numer. Anal.*, vol. 10, pp. 413–432, Apr. 1973.
- [8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1984.
- [9] M. Huang, C. J. Aine, S. Supek, E. Best, D. Ranken, and E. R. Flynn, "Multi-start downhill simplex method for spatio-temporal source localization in magnetoencephalography," *Electroenceph., Clin. Neurophysiol.*, vol. 108, no. 1, pp. 32–44, Jan. 1998.
- [10] H. Krim and M. Viberg, "Two decades of signal processing: The parametric approach," *IEEE Signal Processing, Mag.*, vol. 13, no. 4, pp. 67–94, July 1996.
- [11] B. Luetkenhoener, R. Greenblatt, M. Hämäläinen, J. C. Mosher, M. Scherg, C. Tesche, and S. P. Valdes, "Comparison between different approaches to the biomagnetic inverse problem—Workshop report," in *Biomag'96: Advances in Biomagnetism Research*, C. J. Aine, E. R. Flynn, Y. Okada, G. Stroink, S. J. Switheyby, and C. C. Wood, Eds. New York: Springer-Verlag, to be published.
- [12] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 541–557, June 1992.
- [13] J. C. Mosher, M. E. Spencer, R. M. Leahy, and P. S. Lewis, "Error bounds for EEG and MEG dipole source localization," *Electroenceph., Clin. Neurophysiol.*, vol. 86, pp. 303–321, June 1993.
- [14] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Coherence and MUSIC in biomagnetic source localization," in *Biomagnetism: Fundamental Research and Clinical Applications*, C. Baumgartner, L. Deecke, G. Stroink, and S. J. Williamson, Eds. Amsterdam, the Netherlands: Elsevier/IOS, 1995, pp. 330–334.
- [15] J. C. Mosher, R. M. Leahy, and P. S. Lewis, "Matrix kernels for MEG and EEG source localization and imaging," in *Proc. IEEE Acoust., Speech, Signal Processing Conf.*, 1995, vol. 5, pp. 2943–2946.
- [16] J. C. Mosher and R. M. Leahy, "Source localization using recursively applied and projected (RAP) MUSIC," presented at *31st Ann. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2–5, 1997.
- [17] M. Scherg, "Fundamentals of dipole source potential analysis," in *Auditory Evoked Magnetic Fields and Potentials*, M. Hoke, F. Grandori, and G. L. Romani, Eds. Basel, Switzerland: Karger, 1989, vol. 6.
- [18] M. Scherg and D. von Cramon, "Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model," *Elec., Clin. Neurol.*, vol. 62, pp. 32–44, 1985.
- [19] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 276–280, Mar. 1986; reprint of the original paper presented at RADC Spectrum Estimation Workshop, 1979.
- [20] K. Sekihara, S. Miyachi, and H. Koizumi, "Covariance incorporated MEG-MUSIC algorithm and its application to detect SI and SII when large background brain activity exists," *Neuroimage*, vol. 3, no. 3, p. S29, June 1996.
- [21] J. C. Shaw and M. Roth, "Potential distribution analysis I: A new technique for the analysis of electrophysiological phenomena," *Electroenceph., Clin. Neurophysiol.*, pp. 273–284, 1955.
- [22] A. C. K. Soong and A. J. Koles, "Principal-component localization of the sources of the background EEG," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 59–67, Jan. 1995.
- [23] H. Sorenson, *Parameter Estimation, Principles and Problems*. New York: Marcel Dekker, 1980.
- [24] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Signal Processing*, vol. 38, pp. 1132–1143, July 1990.
- [25] P. Stoica, P. Handel, and A. Nehorai, "Improved sequential MUSIC," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, pp. 1230–1239, Oct. 1995.
- [26] S. Supek and C. J. Aine, "Simulation studies of multiple dipole neuromagnetic source localization: Model order and limits of source resolution," *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 529–540, June 1993.
- [27] J. H. Tripp, "Physical concepts and mathematical models," in *Biomagnetism: An Interdisciplinary Approach*, S. J. Williamson, Ed. New York: Plenum, 1982, pp. 101–149.
- [28] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 1110–1121, May 1991.
- [29] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 2436–2449, Nov. 1991.
- [30] M. Viberg and A. L. Swindlehurst, "Analysis of the combined effects of finite samples and model errors on array processing performance," *IEEE Trans. Signal Processing*, vol. 42, pp. 3073–3083, Nov. 1994.
- [31] 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.



**John C. Mosher** (S'80–M'82) received the B.S. degree in electrical engineering with highest honors from the Georgia Institute of Technology, Atlanta, in 1983. From 1979–1983, he was also a cooperative education student with Hughes Aircraft Company in Fullerton, CA. He received the M.S. (1985) and Ph.D. (1993) degrees in electrical engineering from the Signal and Image Processing Institute of the University of Southern California.

From 1983–1993, he worked at TRW in Los Angeles, CA, researching signal analysis procedures for electromagnetic pulse effects. Upon graduation, he accepted a staff position at the Los Alamos National Laboratory, Los Alamos, NM, where he researches the forward and inverse modeling problems of electrophysiological recordings. A member of the Biophysics Group, his interests also include the general source localization and imaging problems, both in neuroscience work and in other novel applications of sensor technology.



**Richard M. Leahy** (M'85) was born in Surrey, England, in 1960. He received the B.Sc. and Ph.D. degrees in electrical engineering from the University of Newcastle, Newcastle upon Tyne, England, in 1981 and 1985, respectively.

In 1985, he joined the University of Southern California where he is currently a Professor in the Department of Electrical Engineering-Systems and Director of the Signal and Image Processing Institute. He holds joint appointments with the Departments of Radiology and Biomedical Engineering at USC. His research interests lie in the application of signal and image processing theory to biomedical inverse problems. His current research involves the reconstruction and analysis of medical images with an emphasis on PET, MEG, and MRI.