USC–SIPI REPORT #310

Source Localization Using Recursively Applied and Projected (RAP) MUSIC

by

John C. Mosher and Richard M. Leahy

May 1997

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

Source Localization Using Recursively Applied and Projected (RAP) MUSIC

John C. Mosher* and Richard M. Leahy+

*Los Alamos National Laboratory
Group P-21 MS D454
Los Alamos, NM 87545
mosher@LANL.Gov, (505) 665-2175

+Signal & Image Processing Institute
University of Southern California
Los Angeles, CA 90089-2564
leahy@SIPI.USC.Edu, (213) 740-4659

Los Alamos Technical Report No. LA-UR-97-1881
and
USC-SIPI Report No. 310

Source Localization Using Recursively Applied and Projected (RAP) MUSIC[‡]

SP EDICS 3.6.2

John C. Mosher[*] and Richard M. Leahy[+]

[*]Los Alamos National Laboratory, Group P-21 MS D454,
Los Alamos, NM 87545 mosher@LANL.Gov, (505) 665-2175

[+]Signal & Image Processing Institute, University of Southern California,
Los Angeles, CA 90089-2564 leahy@SIPI.USC.Edu, (213) 740-4659

A new method for source localization is described that is based on a modification of the well known multiple signal classification (MUSIC) algorithm. In classical MUSIC, errors in the estimate of the signal subspace can make it difficult to accurately locate multiple sources using projections of the array vectors onto the signal subspace. Instead, recursively applied and projected (RAP) MUSIC finds multiple sources in a recursive fashion, by projecting both the signal subspace estimate and the array vectors against the orthogonal complement of the array gain matrix corresponding to the sources already found. Special assumptions about the array manifold structure, such as Vandermonde or shift invariance, are not required. We show through Monte-Carlo trials that this approach can provide improved performance in comparison to MUSIC and to the previously proposed "sequential" methods, S- and IES-MUSIC. This new method is described in the context of principal angles or principal correlations. Furthermore, through the use of these "subspace" correlations, we present a unified description that includes weighted subspace fitting methods. Finally, we describe extensions of RAP-MUSIC to cases of several sources which are diversely polarized or other vector sources which produce multidimensional array manifolds.

[‡]This work was supported in part by the National Institute of Mental Health Grant R01-MH53213 and by Los Alamos National Laboratory, operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

I. INTRODUCTION

Signal subspace methods in array processing encompass a range of techniques for localizing multiple sources by exploiting the eigenstructure of the measured data matrix. Multiple signal classification (MUSIC) [13] and its many variants are among the more frequently studied subspace methods. The attractions of these MUSIC methods are twofold. First, they can provide computational advantages over direct least squares methods in which all sources are located simultaneously. More importantly, they also allow exhaustive searches over the parameter space for each source, thereby avoiding potential problems with local minima encountered in searching for multiple sources over a non-convex error surface. Subspace methods have been most widely studied in application to the problem of direction of arrival estimation for narrow-band linear equally spaced arrays. Other applications involve broadband and near-field sources and arrays with arbitrary element locations. In these cases, range and azimuth may become additional parameters over which the search must be conducted. The problem can become even more involved when the sources are diversely polarized, such that the array manifolds required to model the sources with unknown polarization become multidimensional. Subspace methods can also be applied to nontraditional array processing problems, for example, the localization of quasi-static electromagnetic sources from eletrophysiological and meteorological data [6, 7, 8].

One important application of subspace methods is to the localization of equivalent current dipoles in the human brain from measurements of scalp potentials (the EEG) or external magnetic fields (the MEG) (collectively E/MEG). These current dipoles represent the foci of neural current sources in the cerebral cortex associated with neural activity in response to sensory, motor or cognitive stimuli. In this case, the current dipoles have three unknown location parameters and an unknown dipole orientation (which is equivalent to the diversely polarized case treated in [1, 13]). A direct search for the location and orientation of multiple sources involves solving a highly non-convex optimization problem. Problems with convergence to local minima have motivated other E/MEG researchers to resort to alternative search strategies such as simulated annealing and the use of genetic algorithms. As an alternative approach, we investigated a signal subspace

approach based on the MUSIC algorithm [6]. Although initial results were promising, problems are often encountered, primarily due to two factors. First, errors in the estimates of the signal subspace from noisy data can make it difficult to differentiate "true" from "false" MUSIC peaks. Second, it is often difficult to find all of the true peaks when searching a three dimensional parameter space over a discrete grid.

The method that we describe here was developed in an attempt to overcome the limitations of MUSIC. In describing this method, we also review other signal subspace methods that we show can be viewed in a unified framework. Since we are primarily interested in the E/MEG source localization problem, we have restricted our attention to methods that do not impose specific constraints on the form of the array manifold. For this reason, we do not consider methods such as ESPRIT [12] or ROOT-MUSIC [see 4], which may efficiently exploit shift invariance or Vandermonde structure in specialized arrays.

One approach to source localization that can result in significantly reduced computation over straightforward least-squares is the weighted subspace fitting (WSF) method [17, 18, 19]. We show here an interpretation of least-squares and WSF in terms of *principal correlations*, from which the MUSIC algorithm may also be developed. This framework of principal correlations can also be used to develop the "sequential" forms of MUSIC, S- and IES- MUSIC [11, 15], as well as to introduce two new methods: recursively applied (R-MUSIC) and recursively applied and projected (RAP-MUSIC).

The paper is arranged as follows. In Section II, we develop weighted subspace fitting from a least-squares perspective, then examine the principal correlations implicit in the central calculations of these metrics. We then review MUSIC in Section III from a principal correlation perspective. In Section IV, we review the related sequential forms of MUSIC, and describe our new modification, RAP-MUSIC. We then summarize the WSF and MUSIC methods in Section V. In Section VI, we describe extensions of RAP-MUSIC to the more general case of multidimensional sources and array manifolds, and in Section VII we conclude with a simulation comparing the performance of the various sequential forms of MUSIC.

II. Signal Subspace Methods

We consider the standard m sensor element array problem, for which our goal is to estimate the parameters for r sources impinging on the array. Each source is represented by an $m > r$ (possibly complex) array manifold vector $\boldsymbol{a}(\theta)$, each source parameter θ may be multidimensional, and the collection of the r manifold parameters is designated $\Theta = \{\theta_1, \ldots, \theta_r\}$. These manifold vectors collectively form an $m \times r$ array transfer matrix

$$\boldsymbol{A}(\Theta) = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_r)] \tag{1}$$

which we assume to be of full column rank r for any putative set Θ, i.e., no array ambiguities exist. Associated with each array vector is a time series $s(t)$, and the data are acquired as $\boldsymbol{x}(t) = \boldsymbol{A}(\theta)\boldsymbol{s}(t) + \boldsymbol{n}(t)$, where $\boldsymbol{s}(t)$ is the vector of r time series at time t. The data are "pre-whitened," i.e., the additive noise vector $\boldsymbol{n}(t)$ has zero mean and covariance of $E\{\boldsymbol{n}(t)\boldsymbol{n}^H(t)\} = \sigma_n^2\boldsymbol{I}$, where $E\{\bullet\}$ is the expectation operator and superscript "^{H}" denotes the Hermitian transpose.

The autocorrelation of $\boldsymbol{x}(t)$ can be decomposed into the well-known partitioning

$$\begin{aligned} \boldsymbol{R} &= E\{\boldsymbol{x}(t)\boldsymbol{x}^H(t)\} \\ &= \boldsymbol{A}(\Theta)(E\{\boldsymbol{s}(t)\boldsymbol{s}^H(t)\})\boldsymbol{A}(\Theta)^H + \sigma_n^2\boldsymbol{I} \\ &= \Phi[\Lambda + \sigma_n^2\boldsymbol{I}]\Phi^H = \Phi_s\Lambda_s\Phi_s^H + \Phi_n\Lambda_n\Phi_n^H \end{aligned} \tag{2}$$

where we have assumed that the time series $\boldsymbol{s}(t)$ are uncorrelated with the noise. We assume that the correlation of the signal time series yields a full rank matrix $\boldsymbol{P} = E\{\boldsymbol{s}(t)\boldsymbol{s}^H(t)\}$, and therefore $\boldsymbol{A}(\Theta)\boldsymbol{P}\boldsymbol{A}(\Theta)^H$ can be eigendecomposed as $\Phi_s\Lambda\Phi_s^H$, such that $\text{span}(\boldsymbol{A}(\Theta)) = \text{span}(\Phi_s)$. The r eigenvalues of the decomposition combine with the noise covariance to form the $r \times r$ diagonal matrix $\Lambda_s = \Lambda + \sigma_n^2\boldsymbol{I}$, with the eigenvalues in Λ_s arranged in decreasing order. The $(m-r) \times (m-r)$ diagonal matrix Λ_n contains the $m-r$ repeated eigenvalues σ_n^2. Thus *(2)* rep-

resents the well-known partitioning of the covariance matrix into *signal subspace* ($\text{span}(\Phi_s)$) and *noise-only subspace* ($\text{span}(\Phi_n)$) terms.

In practice, we acquire n samples of the data to form the spatio-temporal data matrix $X = [x(1), \ldots, x(n)]$; we shall assume for simplicity that $n > m$. We may scale and decompose this data matrix using an SVD as

$$X/\sqrt{n} = \hat{\Phi}\hat{\Sigma}\hat{\Psi}^H \tag{3}$$

where the orthogonal matrix of left singular vectors $\hat{\Phi}$ and the diagonal matrix of singular values $\tilde{\Sigma}$ are each $m \times m$, and the orthogonal matrix of right singular vectors $\hat{\Psi}$ is $n \times m$. Equivalently we may eigendecompose the outer product of this scaled matrix as

$$(XX^H)/n = \hat{\Phi}\hat{\Sigma}^2\hat{\Phi}^H. \tag{4}$$

We observe that $(XX^H)/n$ is a sample estimate of $\boldsymbol{R}$ in *(2)*. Accordingly, we designate the first r left singular vectors in *(3)* as $\hat{\Phi}_s$, i.e., our estimate of a set of vectors which span the signal subspace; similarly we designate $\hat{\Phi}_n$ from the remaining eigenvectors. We designate $\hat{\Sigma}_s$ as the diagonal matrix containing the first r singular values and $\hat{\Sigma}_n$ as the diagonal matrix containing the remaining singular values. The diagonal matrix $\hat{\Lambda}_s = \hat{\Sigma}_s^2$ contains the first r eigenvalues and $\hat{\Lambda}_n = \hat{\Sigma}_n^2$ the remaining eigenvalues from *(4)*.

In least-squares fitting, we estimate the source parameters as

$$\{\hat{\Theta}, \hat{S}\}_{\text{ls}} = \arg\min \left\| X - A(\Theta)S^H \right\|_F^2 \tag{5}$$

i.e., we minimize the squared Frobenius norm of the error matrix, where $\boldsymbol{S}$ is the matrix of associated time series for all sources. Well-known optimal substitution [2] of the linear terms yields

$$\hat{\Theta}_{\text{ls}} = \arg\min \left\| X - AA^{\dagger}X \right\|_F^2 \tag{6}$$

For convenience we will not explicitly show the dependence of $A(\Theta)$ on its parameters. Once the minimizing set of parameters $\hat{\Theta}_{\text{ls}}$ has been found, the linear parameters are simply found as $\hat{S}_{\text{ls}}^H = A(\hat{\Theta}_{\text{ls}})^\dagger X$.

The product of A and its pseudoinverse $A^\dagger$ is an orthogonal projection operator $\Pi_A = AA^\dagger$, which may equivalently be represented by the outer product $\Pi_A = U_A U_A^H$, where U_A is the matrix whose columns are the left singular vectors of A that correspond to nonzero singular values. Substituting this definition and *(3)* into *(6)*, then exploiting the Frobenius norm-preserving properties of orthogonal matrices yields the equivalent least squares statements,

$$\hat{\Theta}_{\text{ls}} = arg\ min\left\{ \|X\|_F^2 - \|\Pi_A X\|_F^2 \right\} \tag{7}$$

$$= arg\ max \left\| U_A U_A^H \hat{\Phi} \hat{\Sigma} \hat{\Psi}^H \right\|_F^2 \tag{8}$$

$$= arg\ max \left\| U_A^H \hat{\Phi} \hat{\Sigma} \right\|_F^2 \tag{9}$$

Substituting in our subspace representations of X yields

$$\hat{\Theta}_{\text{ls}} = arg\ max \left\{ \left\| U_A^H \hat{\Phi}_s \hat{\Sigma}_s \right\|_F^2 + \left\| U_A^H \hat{\Phi}_n \hat{\Sigma}_n \right\|_F^2 \right\} \tag{10}$$

From *(10)*, we may observe the following. Since the singular values of X are ordered in decreasing value, maximization of these least-squares equations favor fitting the first term containing $\hat{\Sigma}_s$ rather than the second term containing $\hat{\Sigma}_n$. Similarly, as we acquire more data, our estimate of R improves, and consequently so does $\hat{\Phi}_s$ and $\hat{\Phi}_n$. By construction, the true values project as $\Pi_A \Phi_s = \Phi_s$ and $\Pi_A \Phi_n = 0$. These observations lead to an alternative maximization criterion function that focuses on just the first term,

$$\hat{\Theta}_{\text{mls}} = arg\ max \left\| U_A^H \hat{\Phi}_s \hat{\Sigma}_s \right\|_F^2 \tag{11}$$

i.e., a "modified" least-squares criterion.

Effectively, $\hat{\Sigma}_s$ in *(11)* represents a weighted sum of the projections of the estimated signal subspace eigenvectors. A more general expression replaces the diagonal matrix with an arbitrary weighting matrix $W^{1/2}$ that can be adjusted to reflect the quality of these estimates, yielding

$$\hat{\Theta}_{\text{wsf}} = arg\ max \left\| \boldsymbol{U}_A^H \hat{\Phi}_s \boldsymbol{W}^{1/2} \right\|_F^2 \tag{12}$$

This criterion function is known as *weighted subspace fitting* (WSF) [17, 18]. This terminology originates by noting that since span($\boldsymbol{A}$) = span(Φ_s) in the noiseless case, then the two matrices must be related by an invertible transformation matrix $\boldsymbol{T}$, such that

$$\Phi_s = \boldsymbol{A}(\Theta)\boldsymbol{T}. \tag{13}$$

Since we only have the estimate $\hat{\Phi}_s$, we may attempt to minimize the error in this statement using a Frobenius norm of the difference. Including an arbitrary weighting matrix $W^{1/2}$ yields the equivalent weighted subspace fitting statement,

$$\{\hat{\Theta}, \hat{\boldsymbol{T}}\}_{\text{wsf}} = arg\ min \left\| \hat{\Phi}_s \boldsymbol{W}^{1/2} - \boldsymbol{A}(\Theta)\boldsymbol{T} \right\|_F^2 \tag{14}$$

The same transformation steps presented above for least-squares *(5)* readily equate *(12)* and *(14)*.

With the appropriate statistical assumptions for the source and noise, an optimal weighting for WSF has been shown to be the diagonal matrix [4, 14]

$$\boldsymbol{W}_{\text{opt}} = (\Lambda_s - \sigma_n^2 \boldsymbol{I})^2 \Lambda_s^{-1} \tag{15}$$

Since these quantities are unknown in the inverse problem, their estimates are used instead,

$$\hat{\boldsymbol{W}}_{\text{opt}} = (\hat{\Lambda}_s - \hat{\sigma}_n^2 \boldsymbol{I})^2 \hat{\Lambda}_s^{-1} \tag{16}$$

Designating the set of signal subspace eigenvalues in the matrix $\hat{\Lambda}_s$ as $\{\hat{\lambda}_1, \ldots, \hat{\lambda}_r\}$, we observe that the jth diagonal element of $\hat{\boldsymbol{W}}_{\text{opt}}$ may be equivalently expressed as

$$\hat{w}_{jj}^2 = (\hat{\lambda}_j - \hat{\sigma}_n^2)^2 / \hat{\lambda}_j = \hat{\lambda}_j (1 - \hat{\sigma}_n^2 / \hat{\lambda}_j)^2. \tag{17}$$

Thus we "derate" those signal subspace components closest to the estimated noise floor $\hat{\sigma}_n^2$. *Signal subspace fitting* (SSF) [17] is simply WSF *(12)* with the weighting matrix set to the identity matrix,

$$\hat{\Theta}_{\text{ssf}} = arg\ max \left\| \boldsymbol{U}_A^H \hat{\Phi}_s \right\|_F^2 \tag{18}$$

In least-squares, WSF, and SSF, we note the common inner product $\boldsymbol{U}_A^H \hat{\Phi}_s$. We now decompose this product with an SVD to yield

$$\boldsymbol{U}_A^H \hat{\Phi}_s = \boldsymbol{Y} \Sigma_C \boldsymbol{Z}^H \tag{19}$$

where $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are each $r \times r$ orthogonal matrices. We designate the r ordered singular values in the diagonal matrix Σ_C as $\{c_1, \ldots, c_r\}$. Substituting *(19)* into *(12)* yields

$$\hat{\Theta}_{\text{wsf}} = arg\ max \left\{ \left\| \Sigma_C \boldsymbol{Z}^H \hat{\boldsymbol{W}}_{\text{opt}}^{1/2} \right\|_F^2 \right\} \tag{20}$$

We designate the r columns of $\boldsymbol{Z}$ as $[z_1, \ldots, z_r]$ and the jth element of the kth vector as $z_{k,j}$ (i.e., the j,kth element of the matrix $\boldsymbol{Z}$). We may therefore express *(20)* as

$$\hat{\Theta}_{\text{wsf}} = arg\ max \left\{ \sum_{k=1}^{r} c_k^2 \sum_{j=1}^{r} z_{k,j}^2 \hat{w}_{jj}^2 \right\} \tag{21}$$

i.e., each c_k^2 has an associated weight. By comparison, SSF ($\boldsymbol{W}$ set to identity) yields simply an unweighted sum

$$\hat{\Theta}_{\text{ssf}} = arg\ max \left\{ \sum_{k=1}^{r} c_k^2 \right\}. \tag{22}$$

since $\sum_j z_{k,j}^2$ is unity by construction in *(19)*. We thus see that (weighted) subspace fitting fits those source parameters Θ which maximize a (weighted) sum of the singular values found in *(19)*.

The decomposition shown in *(19)* yields the *principal correlations* between the subspaces spanned by $\boldsymbol{A}$ and $\hat{\Phi}_s$ [3], also known as the *canonical correlations* (e.g. [16]). In the appendix,

we summarize the steps for calculating these "subspace correlations" to yield a function we designate as

$$[c_1, \ldots, c_r, \tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{Z}}] = subcorr\{\boldsymbol{A}, \boldsymbol{B}\} \tag{23}$$

which returns the ordered set of principal correlations $\{c_1, \ldots, c_r\}$, $1 \geq c_1 \geq \ldots \geq c_r \geq 0$, between the two subspaces spanned by $\boldsymbol{A}$ and $\boldsymbol{B}$, where r is the minimum of the ranks of $\boldsymbol{A}$ and $\boldsymbol{B}$. The appendix contains the definitions of the auxiliary data products $\tilde{\boldsymbol{Y}}$ and $\tilde{\boldsymbol{Z}}$, which are used below. We may thus interpret SSF (WSF) as simply minimizing the (weighted) sum of the squared principal correlations between $\boldsymbol{A}(\Theta)$ and $\hat{\Phi}_s$.

III. MUSIC

The least-squares and WSF methods reviewed above require nonlinear multidimensional searches to find the unknown parameters. MUSIC was introduced by Schmidt [13] as a means to reduce the complexity of the nonlinear search. Here we review MUSIC in terms of the principal correlations, which in turn leads to recently proposed modifications to MUSIC [10, 11, 15] and our proposed RAP-MUSIC approach.

Given that the rank of $\boldsymbol{A}(\Theta)$ is r and the rank of $\hat{\Phi}_s$ is at least r, the smallest principal correlation value,

$$min\, subcorr\{\boldsymbol{A}(\Theta), \hat{\Phi}_s\} \equiv c_r = \boldsymbol{u}_r^H \boldsymbol{v}_r = (\boldsymbol{A}(\Theta)\tilde{\boldsymbol{y}}_r)^H \hat{\Phi}_s \tilde{\boldsymbol{z}}_r, \tag{24}$$

represents the minimum principal correlation (maximum principal angle) between principal vectors in the column space of $\boldsymbol{A}(\Theta)$ and the signal subspace $\hat{\Phi}_s$; see the appendix for the definitions of $\tilde{\boldsymbol{y}}_r$ and $\tilde{\boldsymbol{z}}_r$. The subspace correlation of any individual column $\boldsymbol{a}(\theta) \in \boldsymbol{A}(\Theta)$ with the signal subspace must therefore equal or exceed this minimum principal correlation,

$$subcorr\{\boldsymbol{a}(\theta_i), \hat{\Phi}_s\} \geq min\, subcorr\{\boldsymbol{A}(\Theta), \hat{\Phi}_s\}, \; i = 1, \ldots, r \tag{25}$$

As the quality of our signal subspace estimate improves (either by improved signal to noise

ratios or longer data acquisition), then $\hat{\Phi}_s$ will approach Φ_s, and the minimum correlation approaches unity when the correct parameter set Θ is identified, such that the r distinct sets of parameters θ_i have principal correlations approaching unity. Thus a search strategy for identifying the multiple parameter set $\Theta = \{\theta_1, \ldots, \theta_r\}$ is to identify r peaks of the metric

$$subcorr^2\{a(\theta), \hat{\Phi}_s\} = \frac{a^H(\theta)\hat{\Phi}_s\hat{\Phi}_s^H a(\theta)}{\|a(\theta)\|^2} \tag{26}$$

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 orthonormal. We recognize this form as the *multiple signal classification* (MUSIC) metric [13], with the minor difference of using the signal subspace projector $\hat{\Phi}_s\hat{\Phi}_s^H$ rather than the more commonly used noise-only subspace projector $\hat{\Phi}_n\hat{\Phi}_n^H$. If our estimate of the signal subspace is perfect, then we will find r global maxima equal to unity. If we assume scalar parameters θ, then the advantage of MUSIC is obvious: the least-squares or WSF search over an r-dimensional space Θ is replaced by a one dimensional search. The MUSIC advantage for higher dimensional θ is even greater.

Errors in our estimate $\hat{\Phi}_s$ typically reduce *(26)* to a function with a single global maximum and at least $(r-1)$ local maxima. Identifying the local maxima becomes more difficult, since nonlinear search techniques may miss shallow or adjacent peaks and return to a previous peak. We also need to locate the r best peaks, not simply a peak associated with a simple local extremum of the measure. As illustrated in Fig. 1, at low SNR the other sources may not even exhibit an adequate "peak-like" structure, but rather exhibit only a deflection in the side of an adjacent peak. This "peak-picking" problem becomes more difficult as the dimension of each individual source rises, such as in two and three-dimensional source localization.

IV. Sequential Forms of MUSIC

A. R-MUSIC

As noted, least-squares, WSF, and SSF require a single multidimensional search, while MUSIC consists of a series of low-dimensional searches for r peaks. From the principal correlation viewpoint, we recognize that several intermediate correlations are possible between the cases represented by MUSIC and SSF. In the case of perfect estimation of the signal subspace, we already note that a single set $\Theta = \{\theta_1, \ldots, \theta_r\}$ yields $c_r = 1$,

$$[c_1, \ldots, c_r] = subcorr\{\boldsymbol{A}(\Theta), \Phi_s\}. \tag{27}$$

We further note that there are r possible subsets each containing $(r-1)$ parameters, each subset yielding

$$[c_1, \ldots, c_{r-1}] = subcorr\{\boldsymbol{A}(\{\theta_1, \ldots, \theta_{r-1}\}), \Phi_s\} \tag{28}$$

where $c_{r-1} = 1$. For subsets comprising k sources from the r total sources, we observe that the combinatorics yield $\binom{r}{k}$ sets, each set yielding $c_k = 1$.

The case $k = 1$ is MUSIC. For $k = 2$, there are $r(r-1)/2$ possible sets yielding $c_2 = 1$. An alternative approach, however, is to fix one of these pairs of sources using a peak from $k = 1$, leaving just $(r-1)$ source locations such that $c_2 = 1$. If we fix the second source parameter at one of these peaks, then for the case $k = 3$ there remain $(r-2)$ source locations such that $c_3 = 1$. This sequential search through the subspace correlations in the perfect subspace case Φ_s suggests the design of an algorithm that bypasses the peak-picking problem of MUSIC.

We refer to this method as recursive MUSIC (R-MUSIC). The first recursion of R-MUSIC is identical to MUSIC,

$$\hat{\theta}_1 = arg\ max\, subcorr\{\boldsymbol{a}(\theta), \hat{\Phi}_s\} \tag{29}$$

This maximization is easily carried out by first searching over a grid of putative θ, then initializing

a nonlinear refinement from the best grid point. For an r dimensional signal subspace, we anticipate at least r peaks, but we retain only the source location associated with the maximum peak. With the best single source parameter identified, we form the preliminary array transfer matrix

$$\hat{A}^{(1)} = a(\hat{\theta}_1). \tag{30}$$

To find the second source, we maximize

$$\hat{\theta}_2 = arg\ max\{c_2\} \tag{31}$$

where c_2 is the second principal correlation of

$$[c_1, c_2] = subcorr\{[\hat{A}^{(1)}, a(\theta)], \hat{\Phi}_s\} \tag{32}$$

i.e., the concatenation of the array vectors for the first identified source with those for our putative locations for a second source. A search through source locations θ should reveal at least $r-1$ peaks of c_2, but in general not at $\hat{\theta}_1$, since the subspace associated with this source has already been identified with $\hat{A}^{(1)}$. With the second source identified at the global maximum of c_2, we form the trial array transfer matrix as

$$\hat{A}^{(2)} = [a(\hat{\theta}_1), a(\hat{\theta}_2)]. \tag{33}$$

We repeat this recursion r times, such that on the final iteration we are searching for the single global maximum of

$$\hat{\theta}_r = arg\ max\{c_r\} \tag{34}$$

where

$$[c_1, \ldots, c_r] = subcorr\{[\hat{A}^{(r-1)}, a(\theta)], \hat{\Phi}_s\} \tag{35}$$

Thus the kth recursion generates a measure with at least $(r+1-k)$ peaks, but we need only search for the maximum peak which is considerably simpler than looking for multiple local peaks. If the dimensionality of each source parameter θ is low, then a search for the global maximum at each recursion may readily be carried out on a dense grid of putative source locations. Once a set

of r sources is identified, we can refine these peaks in a full nonlinear maximization using weighted subspace fitting or a least-squares approach. If the final correlation c_r is close to unity, then this refinement may be unnecessary.

In [5, 9, 10], we implement this R-MUSIC algorithm for the case of localizing current dipoles in the human brain, using E/MEG data acquired over an array of sensors. We discuss some of the special considerations for this form of array manifold in Section VI. In a subsequent search of the literature for similar MUSIC approaches, we learned of the "sequential" forms S-MUSIC [11] and IES-MUSIC [15], with the latter presented as an extension of the former. Our further investigations of these approaches, in combination with our needs in E/MEG processing, lead in this section to an improved variation of R-MUSIC that we call recursively applied and projected MUSIC (RAP-MUSIC).

B. S- and IES-MUSIC

S-MUSIC [11] and IES-MUSIC [15] are two sequential modifications of MUSIC that are closely related to the R- and RAP-MUSIC methods introduced here. Both methods find the first source by maximizing the MUSIC metric for a single source (we will continue to express the metrics using the signal subspace eigenvectors $\hat{\Phi}_s$ rather than the noise-only eigenvectors). Using the first source, we form the orthogonal projection operators

$$\Pi_{a(\hat{\theta}_1)} = (a(\hat{\theta}_1)a^H(\hat{\theta}_1))/\|a(\hat{\theta}_1)\|^2 \tag{36}$$

$$\Pi^{\perp}_{a(\hat{\theta}_1)} = I - \Pi_{a(\hat{\theta}_1)} \tag{37}$$

In S-MUSIC [11], we apply the projection operator *(37)* to the array manifold and find the second source as

$$\{\hat{\theta}_2\}_S = arg\ max\ \ g(\theta) \tag{38}$$

where

$$g(\theta) = \frac{(\boldsymbol{a}^H(\theta)\Pi^{\perp}_{\boldsymbol{a}(\hat{\theta}_1)}\hat{\Phi}_s\hat{\Phi}_s^H\Pi^{\perp}_{\boldsymbol{a}(\hat{\theta}_1)}\boldsymbol{a}(\theta))}{\left\|\Pi^{\perp}_{\boldsymbol{a}(\hat{\theta}_1)}\boldsymbol{a}(\theta)\right\|^2} \tag{39}$$

(or $g(\theta) = 0$ equals zero if $\left\|\Pi^{\perp}_{\boldsymbol{a}(\hat{\theta}_1)}\boldsymbol{a}(\theta)\right\| = 0$). Remark: In the alternative noise-only subspace minimization form as presented in [11], $g(\theta)$ is at a minimum for both the first and second source estimates, therefore increasing the complexity of the search algorithm. In the signal subspace maximization form presented here *(39)*, we simply look for the maximum to find the second source.

In IES-MUSIC [15], the denominator of *(39)* is dropped as inconsequential to overall performance, and the following modification is suggested,

$$g(\theta, \hat{\rho}) = \boldsymbol{a}^H(\theta)(\boldsymbol{I} - \hat{\rho}^*\Pi_{\boldsymbol{a}(\hat{\theta}_1)})\hat{\Phi}_s\hat{\Phi}_s^H(\boldsymbol{I} - \hat{\rho}\Pi_{\boldsymbol{a}(\hat{\theta}_1)})\boldsymbol{a}(\theta) \tag{40}$$

This measure is effectively S-MUSIC for $\hat{\rho} = 1$ and is MUSIC for $\hat{\rho} = 0$. IES-MUSIC designs an optimal scalar ρ for the case of two sources, but this scalar requires knowledge of the two sources θ_1 and θ_2. Since these parameters are unknown, IES-MUSIC first obtains the estimated locations $\hat{\theta}_1$ and $\hat{\theta}_2$ from another approach, such as MUSIC. Using the same estimated optimal weighting matrix from WSF fitting, $\hat{\boldsymbol{W}}_{\text{opt}}$ *(16)*, the approach forms the matrix $\boldsymbol{W}_{\text{IES}}$ as

$$\boldsymbol{W}_{\text{IES}} = [\boldsymbol{a}(\hat{\theta}_1), \boldsymbol{a}(\hat{\theta}_2)]^H\hat{\Phi}_s\hat{\boldsymbol{W}}_{\text{opt}}^{-1}\hat{\Phi}_s^H[\boldsymbol{a}(\hat{\theta}_1), \boldsymbol{a}(\hat{\theta}_2)] \tag{41}$$

The estimated optimal weighting scalar $\hat{\rho}$ is then formed as [15]

$$\hat{\rho} = \frac{m\boldsymbol{W}^*_{\text{IES}}(2, 1)}{(\boldsymbol{a}^H(\hat{\theta}_1)\boldsymbol{a}(\hat{\theta}_2))\boldsymbol{W}_{\text{IES}}(1, 1)} \tag{42}$$

where $\boldsymbol{W}_{\text{IES}}(i, j)$ is the i, j th element of the matrix. The location of the second source is then refined as the point which maximizes *(40)*. Extensions of S- and IES- MUSIC to more than two sources are possible, as noted in [15]. For S-MUSIC the extension is straightforward: the projection operator in *(37)* is replaced with one that is generated from the concatenation of the array vectors for all sources that have been found (cf. *(33)* to *(35)*). The extension for IES-MUSIC is not so

straightforward, since the weighting scalar $\hat{\rho}$ becomes a function of more than two sources.

C. RAP-MUSIC

The projection using $\Pi^{\perp}_{a(\hat{\theta}_1)}$ in the sequential forms of MUSIC described above and the recursive procedure of R-MUSIC can be combined into an alternative form of R-MUSIC, which we call recursively applied and projected (RAP) MUSIC[‡]. In this method, as with all the previous sequential forms, the first source is found using a traditional MUSIC search. The second source is then found by applying the projector operator $\Pi^{\perp}_{a(\hat{\theta}_1)}$ to both arguments of *subcorr* in *(32)* to yield

$$c_1(2) = subcorr\left\{\Pi^{\perp}_{a(\hat{\theta}_1)}\boldsymbol{a}(\theta), \Pi^{\perp}_{a(\hat{\theta}_1)}\hat{\Phi}_s\right\}, \tag{43}$$

where we note that $\Pi^{\perp}_{a(\hat{\theta}_1)}[\boldsymbol{a}(\hat{\theta}_1), \boldsymbol{a}(\theta)] = [\mathbf{0}, \Pi^{\perp}_{a(\hat{\theta}_1)}\boldsymbol{a}(\theta)]$ to yield *(43)*. The notation $c_1(2)$ denotes the principal correlation of the second stage of the source localization procedure. The need for this notation will become clear in the section on extensions of the method.

By explicitly forming the normalizations required in *(43)*, we may express it in a form comparable to the other MUSIC methods as

$$\begin{aligned} &subcorr^2\left\{\Pi^{\perp}_{a(\hat{\theta}_1)}\boldsymbol{a}(\theta), \Pi^{\perp}_{a(\hat{\theta}_1)}\hat{\Phi}_s\right\} \\ &\quad = \frac{(\boldsymbol{a}^H(\theta)\Pi^{\perp}_{a(\hat{\theta}_1)}\hat{\Phi}_s(\hat{V}_{a(\hat{\theta}_1)}\hat{\Sigma}^{-2}_{a(\hat{\theta}_1)}\hat{V}^H_{a(\hat{\theta}_1)})\hat{\Phi}_s^H\Pi^{\perp}_{a(\hat{\theta}_1)}\boldsymbol{a}(\theta))}{\left\|\Pi^{\perp}_{a(\hat{\theta}_1)}\boldsymbol{a}(\theta)\right\|^2} \end{aligned} \tag{44}$$

where we have orthogonalized the projected signal subspace using an SVD as

$$\Pi^{\perp}_{a(\hat{\theta}_1)}\hat{\Phi}_s = \hat{U}_{a(\hat{\theta}_1)}\hat{\Sigma}_{a(\hat{\theta}_1)}\hat{V}^H_{a(\hat{\theta}_1)} \tag{45}$$

and, retaining only the nonzero singular value components of *(45)*, we note that

[‡]In [5, 10], we referred to R-MUSIC as RAP-MUSIC; we have since adopted the terminology used here to differentiate the two methods.

$$\hat{U}_{a(\hat{\theta}_1)} = \Pi^{\perp}_{a(\hat{\theta}_1)} \hat{\Phi}_s \hat{V}_{a(\hat{\theta}_1)} \hat{\Sigma}^{-1}_{a(\hat{\theta}_1)}. \tag{46}$$

Substitution using *(46)* and the idempotent property of projection operators yields *(44)*, a form comparable to *(39)* and *(40)*.

The second source is then found by maximizing the principal correlation $c_1(2)$. Additional sources are found by repeating the procedure, i.e. to find the k^{th} source, we form the projector $\Pi^{\perp}_{\hat{A}^{(k-1)}}$ on to the orthogonal complement of the column space of the array transfer matrix for the first $k-1$ locations:

$$\hat{A}^{(k-1)} = [a(\hat{\theta}_1), a(\hat{\theta}_2) \ldots a(\hat{\theta}_{k-1})] \tag{47}$$

The k^{th} source location is then found as that which maximizes the principal correlation:

$$c_1(k) = subcorr\left\{ \Pi^{\perp}_{\hat{A}^{(k-1)}} a(\theta), \Pi^{\perp}_{\hat{A}^{(k-1)}} \hat{\Phi}_s \right\} \tag{48}$$

As we will demonstrate in the simulations below, we found RAP-MUSIC to be numerically superior to the other forms of MUSIC, including our own R-MUSIC. We also found RAP-MUSIC to be more computationally efficient than R-MUSIC, and hence in Section VII we will focus our performance comparisons on RAP-MUSIC only.

V. Summary of Signal Subspace Methods

Using the preceding development, we can now describe the various subspace fitting and MUSIC methods in a common framework. The modified least-squares and subspace fitting methods in Section II can be stated in the general form:

$$\hat{\Theta} = arg\ max \left\| U^H_{A(\Theta)} \hat{\Phi}_s W^{1/2} \right\|^2_F \tag{49}$$

where $W^{1/2} = \hat{\Sigma}_s$ yields the modified least squares method*(11)*, $W^{1/2} = \hat{W}^{1/2}_{\text{opt}}$ yields WSF *(12)*, and $W^{1/2} = I$ yields SSF *(18)*. We may restate *(49)* as (cf. [4, 18, 19]),

$$\hat{\Theta} = arg\ max\{ Tr\{ \Pi_{A(\Theta)} \hat{\Phi}_s W \hat{\Phi}_s^H \} \}, \tag{50}$$

where $Tr\{\bullet\}$ is the trace of the matrix. Alternatively, similar subspace-based estimators, such as MODE [14], can be generalized as

$$\hat{\Theta} = arg\ min\{ Tr\{ A(\Theta)^H \hat{\Phi}_n \hat{\Phi}_n^H A(\Theta) W \} \} \tag{51}$$

where the weights W,as described in [14], correspond to several different subspace estimators. Both of these estimators, *(50)* and *(51)*, require multidimensional searches over the parameter set Θ and can be shown to be asymptotically related (cf. [4]).

The various MUSIC methods presented in this paper may be placed in this same framework. For the case of two scalar-valued source parameters as a pair of one dimensional searches, the first source is found as the solution of

$$\hat{\theta}_1 = arg\ max\left\{ \frac{a^H(\theta) \hat{\Phi}_s \hat{\Phi}_s^H a(\theta)}{\|a(\theta)\|^2} \right\} \tag{52}$$

and the second source is found as

$$\hat{\theta}_2 = arg\ max\left\{ \frac{a^H(\theta;\hat{\theta}_1) \hat{\Phi}_s W \hat{\Phi}_s^H a(\theta;\hat{\theta}_1)}{\|a(\theta;\hat{\theta}_1)\|^2} \right\} \tag{53}$$

$$= arg\ max\left\{ Tr\left\{ \Pi_{a(\theta;\hat{\theta}_1)} \hat{\Phi}_s W \hat{\Phi}_s^H \right\} \right\} \tag{54}$$

where $a(\theta;\hat{\theta}_1)$ is an altered form of the array manifold, based on the location of the first source, and $\Pi_{a(\theta;\hat{\theta}_1)}$ is the projection operator onto the spaced spanned by this altered array manifold. In Table 1, we summarize the specific forms of $a(\theta;\hat{\theta}_1)$ and W for the projected methods presented in Section IV. Extensions of these methods to more than two sources are discussed in Section IV and Section VI.

Finally, we may also restate these sequential MUSIC methods, using our *subcorr* function, as follows:

MUSIC:

$$\hat{\theta}_2 = arg\ max\, subcorr\{\boldsymbol{a}(\theta), \hat{\Phi}_s\} \tag{55}$$

S-MUSIC:

$$\hat{\theta}_2 = arg\ max\, subcorr\left\{\Pi^{\perp}_{\boldsymbol{a}(\hat{\theta}_1)}\boldsymbol{a}(\theta), \hat{\Phi}_s\right\} \tag{56}$$

IES-MUSIC:

$$\begin{aligned}\hat{\theta}_2 = arg\ max &\left\|\boldsymbol{I} - \hat{\rho}^{*}\Pi_{\boldsymbol{a}(\hat{\theta}_1)}\boldsymbol{a}(\theta)\right\|^2 \cdot \\ &subcorr^2\{(\boldsymbol{I} - \hat{\rho}^{*}\Pi_{\boldsymbol{a}(\hat{\theta}_1)})\boldsymbol{a}(\theta), \hat{\Phi}_s\}\end{aligned} \tag{57}$$

RAP-MUSIC:

$$\hat{\theta}_2 = arg\ max\, subcorr\left\{\Pi^{\perp}_{\boldsymbol{a}(\hat{\theta}_1)}\boldsymbol{a}(\theta), \Pi^{\perp}_{\boldsymbol{a}(\hat{\theta}_1)}\hat{\Phi}_s\right\} \tag{58}$$

VI. Extensions

Some source localization problems involve multidimensional array manifolds and source parameters. For example, the multidimensional manifold can represent "diverse polarization," [1, 13]. In [5, 9, 10] we show how to extend R-MUSIC to the E/MEG problem, where each source location is a three dimensional vector, and each source is represented by a three-dimensional manifold. We specifically describe in these references how the additional matrices $\tilde{\boldsymbol{Y}}$ and $\tilde{\boldsymbol{Z}}$ from the subspace correlation function are useful in estimating the "quasi-linear" polarization parameters, such that the RAP-MUSIC search remains a function of just the nonlinear location parameters.

The extensions of RAP-MUSIC to multidimensional array manifold matrices and multiple sources are straightforward. For instance, in E/MEG [6, 9], the array manifold matrix for a single source is typically represented as a three column matrix,

$$\boldsymbol{G}(\boldsymbol{r}) = [\boldsymbol{g}_x(\boldsymbol{r}), \boldsymbol{g}_y(\boldsymbol{r}), \boldsymbol{g}_z(\boldsymbol{r})] \tag{59}$$

expressed in Cartesian coordinates. Generally, $\boldsymbol{G}(\boldsymbol{r})$ is of full column rank, but in special cases it

may be of rank two [6]. An additional parameter is the source orientation $\boldsymbol{q}$, assumed of unit magnitude, such that the combined source parameter set is $\theta = \{\boldsymbol{r}, \boldsymbol{q}\}$ and the array manifold vector may be expressed as

$$a(\theta) = \boldsymbol{G}(\boldsymbol{r})\boldsymbol{q} \tag{60}$$

These orientation parameters $\boldsymbol{q}$ are effectively linear (i.e., "quasi-linear") and need not be explicitly included in the search. Instead, in the first recursion of RAP-MUSIC we compute

$$[c_1(1), c_2(1), c_3(1)] = subcorr\{\boldsymbol{G}(\boldsymbol{r}), \hat{\Phi}_s\} \tag{61}$$

for each location $\boldsymbol{r}$, and we set the first location as

$$\hat{\boldsymbol{r}}_1 = arg\ max\{c_1(1)\} \tag{62}$$

Once this optimal location has been determined, we then extract the optional parameters $\tilde{\boldsymbol{Y}}$ and $\tilde{\boldsymbol{Z}}$ (see appendix),

$$[c_1(1), c_2(1), c_3(1), \tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{Z}}] = subcorr\{\boldsymbol{G}(\hat{\boldsymbol{r}}_1), \hat{\Phi}_s\} \tag{63}$$

We note that by construction, the linear combination of $\boldsymbol{G}(\hat{\boldsymbol{r}}_1)$ most correlated with $\hat{\Phi}_s$ is found as the principal vector $\boldsymbol{G}(\hat{\boldsymbol{r}})\tilde{\boldsymbol{y}}_1$, i.e, we use the first column of $\tilde{\boldsymbol{Y}}$. We therefore set $\hat{\boldsymbol{q}}_1 = \tilde{\boldsymbol{y}}_1 / \|\tilde{\boldsymbol{y}}_1\|$ to yield the first estimate $\hat{\theta}_1 = \{\hat{\boldsymbol{r}}_1, \hat{\boldsymbol{q}}_1\}$.

By extension, the kth recursion of RAP-MUSIC for multiple sources, a d-dimensional array manifold matrix, and an r dimensional signal subspace is

$$\hat{\theta}_k = \{\hat{\boldsymbol{r}}_k, \hat{\boldsymbol{q}}_k\}, \tag{64}$$

where

$$\hat{\boldsymbol{r}}_k = arg\ max\{c_1(k)\} \tag{65}$$

$$[c_1(k), \ldots, c_{min(d,r)}(k), \tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{Z}}] = subcorr\left\{\Pi^{\perp}_{\hat{A}^{(k-1)}}\boldsymbol{G}(\hat{\boldsymbol{r}}_k), \Pi^{\perp}_{\hat{A}^{(k-1)}}\hat{\Phi}_s\right\} \tag{66}$$

$$\hat{A}^{(k-1)} = [a(\hat{\theta}_1), \ldots, a(\hat{\theta}_{k-1})] \tag{67}$$

$\boldsymbol{a}(\hat{\theta}_i) = \boldsymbol{G}(\hat{\boldsymbol{r}}_i)\hat{\boldsymbol{q}}_i$, $i = 1, \ldots, k-1$, where $\hat{\boldsymbol{q}}_k = \tilde{\boldsymbol{y}}_1 / \|\tilde{\boldsymbol{y}}_1\|$, and $\tilde{\boldsymbol{y}}_1$ is the first column of $\tilde{\boldsymbol{Y}}$ in *(66)*, as defined in the appendix.

VII. Simulation

We have followed the simulations in [15] in order to draw performance comparisons between the various forms of MUSIC. The sensor array is the conventional uniform linear array of sensors spaced a half-wavelength apart. The sources are far field narrow-band and impinging on the array from scalar direction θ. The array manifold vector may therefore be specified as

$$\boldsymbol{a}(\theta) = [1, e^{i\pi\sin\theta}, \ldots, e^{i\pi(m-1)\sin\theta}]^T, \tag{68}$$

where $\theta = 0$ is broadside to the array, and $\|\boldsymbol{a}(\theta)\|^2 = m$. The source time series are assumed to be complex zero-mean Gaussian sequences with covariance matrix

$$E\{\boldsymbol{s}(t)\boldsymbol{s}(t)^H\} = \boldsymbol{P}. \tag{69}$$

We assume fifteen sensor elements and two sources at 25 and 30 degrees. The source covariance matrix is specified as

$$\boldsymbol{P} = \begin{bmatrix} 1 & \gamma \\ \gamma^* & 1 \end{bmatrix} \tag{70}$$

where $|\gamma| \leq 1$ determines the degree of correlation between these two sources of equal power. The variance of the noise is set to unity, such that the signal to noise power ratio is also unity.

We simulate n samples of both the signal and noise, form the estimated data covariance matrix, then extract the matrix $\hat{\Phi}_s$ comprising the two estimated signal subspace vectors. The noise variance is estimated as the mean of the noise-only subspace eigenvalues. For each realization, we find the maxima of the MUSIC measure in a region about each of the true solutions. The source with the better correlation was considered source θ_1. The second source θ_2 was then found by maximizing the appropriate measure, *(55)* - *(58)*. Since IES-MUSIC is a "two-pass" algorithm, i.e., it requires an initial estimate of both source parameters, we used the RAP-MUSIC source estimates

for the initial estimate, as the RAP-MUSIC solution was on average superior to the MUSIC and S-MUSIC estimates. We also ran as a comparison IES-MUSIC with ρ set to the true optimal value using the true source angles. Our preliminary studies showed R-MUSIC to be performing only as well as S-MUSIC, yet at a greater computational cost than RAP-MUSIC; we therefore did not perform further analysis of R-MUSIC in order to focus the presentation on RAP-MUSIC.

In [15], closed-form formulae are presented for calculating the theoretical error variance of MUSIC, S-MUSIC, and IES-MUSIC. For each estimator, we also calculated a numerical root mean squared (RMS) error,

$$\mathrm{RMS} = \left(\frac{1}{\mathrm{Runs}} \sum_{i=1}^{\mathrm{Runs}} \left(\hat{\theta}_2(i) - \theta_2 \right)^2 \right)^{1/2} . \tag{71}$$

where $\hat{\theta}_2(i)$ represents the estimate from the i th Monte Carlo run. In each of these runs, we determined which of the two MUSIC peaks in the regions about the true answer was greater and declared this source as $\hat{\theta}_1$. We then estimated the second source, then tabulated the actual number of runs used for both $\theta_2 = 25$ or 30 degrees, which is approximately evenly split at about 250 Monte Carlo runs each.

In Table 2, we held the number of time samples constant at $n = 1000$ and varied the degree of correlation between the two sources. For uncorrelated sources, $\gamma = 0$, all measures performed similarly, as also demonstrated in [15]. The differences in performance begin to arise at $\gamma = 0.7$, as tabulated in our table, where we observe that IES-MUSIC and RAP-MUSIC have RMS error about 25% better than MUSIC and S-MUSIC. At $\gamma = 0.925$, we see that RAP-MUSIC continues to have performance comparable to that of perfect IES-MUSIC, but that estimated IES-MUSIC is beginning to degrade in comparison; MUSIC and S-MUSIC have RMS error almost twice that of IES-MUSIC and RAP-MUSIC at this point. By $\gamma = 0.975$, all methods are experiencing comparable difficulty in estimating the sources. MUSIC is particularly poor at this correlation, since in many trials an adequate peak-like structure did not occur in the region around the true answer, as

we illustrate in Fig. 1.

In general, the RMS error of MUSIC and S-MUSIC match the theoretical bounds established in [15] quite well, and our RMS errors agree well with those presented in [15] for their comparable cases. RAP-MUSIC consistently maintains an improved RMS error over that of IES-MUSIC, and we note again that IES-MUSIC depends on some other technique in order to arrive at an initial set of source estimates. IES-MUSIC performance using the optimally designed ρ agrees quite well with the theoretical bounds, but this performance obviously requires prior knowledge of the true solution.

These RMS errors were calculated at a relatively large number of time samples. We also tested small sample performance. In Table 3, we held the correlation constant at $\gamma = 0.9$ and varied the number of time samples. At lower numbers of time samples, we generally had a difficult task determining a second MUSIC peak, and the MUSIC results were unreliable. RAP-MUSIC consistently maintained improved performance over the other methods, and the performances were generally in good agreement with the theoretical bounds established by [15].

VIII. Conclusion

We have presented a novel framework, based on the principal correlations between subspaces, in which to view least-squares, weighted subspace fitting, MUSIC, and their variations. The MUSIC methods replace the search for multiple sources with procedures for separately identifying each source. For multiple sources, classical MUSIC requires the identification of multiple local maxima in a single metric. While it is straightforward to identify the first source using the global maximum of this metric, finding subsequent sources requires a peak-picking procedure and can lead to errors, particularly when these sources are weak or strongly correlated with the first source. The other sequential MUSIC forms presented here are measures designed to make localization of the second source more straightforward. Our modifications, R-MUSIC and RAP-MUSIC, are derived from a principal correlations perspective. Our original R-MUSIC algorithm, derived for E/MEG research, had performance comparably to S-MUSIC, but the numerical studies presented

here show our newer RAP-MUSIC to yield improved simulation performance over the other forms of MUSIC. Extensions of the RAP-MUSIC approach to many sources and higher dimensionality of the manifold are also more straightforward than the other sequential forms.

APPENDIX: PRINCIPAL CORRELATION

Here we summarize the definition and method for computation of the principal or "subspace" correlation – the definition and computation method are equivalent to those in [3]. Given two matrices, $\boldsymbol{A}$ and $\boldsymbol{B}$, where $\boldsymbol{A}$ is $m \times p$, and $\boldsymbol{B}$ is $m \times q$, let r be the minimum of the ranks of the two matrices. We define the function

$$[c_1, \ldots, c_r, \tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{Z}}] = subcorr\{\boldsymbol{A}, \boldsymbol{B}\}, \tag{72}$$

as follows. The scalars c_k are the set of ordered principal correlations which follow from the recursion,

$$c_k = \max_{\boldsymbol{u} \in \boldsymbol{A}} \max_{\boldsymbol{v} \in \boldsymbol{B}} \boldsymbol{u}^H \boldsymbol{v} = \boldsymbol{u}_k^H \boldsymbol{v}_k, \; k = 1, \ldots, r \tag{73}$$

subject to

$$\|\boldsymbol{u}_k\| = \|\boldsymbol{v}_k\| = 1 \tag{74}$$

$$\boldsymbol{u}^H \boldsymbol{u}_i = 0 \qquad i = 1, \ldots, k-1 \tag{75}$$

$$\boldsymbol{v}^H \boldsymbol{v}_i = 0 \qquad i = 1, \ldots, k-1 \tag{76}$$

The vectors $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ are the *principal vectors* between the subspaces spanned by $\boldsymbol{A}$ and $\boldsymbol{B}$, and by construction, each set of vectors represents an orthonormal basis. Note that $1 \geq c_1 \geq c_2 \geq \ldots \geq c_r \geq 0$. The scalar c_k is the *principal correlation* between $\boldsymbol{u}_k$ and $\boldsymbol{v}_k$, or analogously, the angles θ_k, where $\cos\theta_k = c_k$, are the *principal angles*, representing the geometric angle between these two vectors. The steps to compute the subspace correlations are as follows [3 (p. 585)]:

1. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are already orthogonal matrices, we redesignate them as $\boldsymbol{U}_A$ and $\boldsymbol{U}_B$ and

skip to Step 2. Otherwise, perform a singular value decomposition (SVD) of $\boldsymbol{A}$, such that $\boldsymbol{A} = \boldsymbol{U}_A \Sigma_A \boldsymbol{V}_A^T$. Similarly decompose $\boldsymbol{B} = \boldsymbol{U}_B \Sigma_B \boldsymbol{V}_B^T$. Retain only those components of $\boldsymbol{U}_A$ and $\boldsymbol{U}_B$ that correspond to nonzero singular values, i.e., the number of columns in $\boldsymbol{U}_A$ and $\boldsymbol{U}_B$ correspond to their ranks.

2. Form $\boldsymbol{C} = \boldsymbol{U}_A^T \boldsymbol{U}_B$.

3. If only the correlations are needed, then compute only the singular values of $\boldsymbol{C}$ (the extra computation for the singular vectors is not required). The r ordered singular values $1 \geq c_1 \geq \ldots \geq c_r \geq 0$ are the principal correlations between $\boldsymbol{A}$ and $\boldsymbol{B}$.

4. If the principal vectors are also required, then compute the full singular value decomposition, $\boldsymbol{C} = \boldsymbol{Y} \Sigma_C \boldsymbol{Z}^T$. The r ordered singular values are extracted from the diagonal of Σ_C. Form the sets of principal vectors $\boldsymbol{U}_a = \boldsymbol{U}_A \boldsymbol{Y}$ and $\boldsymbol{U}_b = \boldsymbol{U}_B \boldsymbol{Z}$ for sets $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively.

The matrices $\boldsymbol{U}_a$ and $\boldsymbol{U}_b$ are each orthogonal, and the columns comprise the ordered sets of principal vectors for matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ 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 $\boldsymbol{A}$ and $\boldsymbol{B}$ [3]. When the distance is zero, $\boldsymbol{A}$ and $\boldsymbol{B}$ are parallel subspaces. A maximum distance of unity ($c_r = 0$) indicates at least one basis vector of $\boldsymbol{A}$ is orthogonal to $\boldsymbol{B}$; if the maximum principal correlation is $c_1 = 0$, then two subspaces spanned are orthogonal. We see that minimizing the distance is equivalent to maximizing the minimum principal correlation between $\boldsymbol{A}$ and $\boldsymbol{B}$.

We may also readily compute the specific linear combinations of $\boldsymbol{A}$ and $\boldsymbol{B}$ that yield these principal vectors and angles. By construction, we know that $\boldsymbol{A}\tilde{\boldsymbol{Y}} = \boldsymbol{U}_a$ for some $\tilde{\boldsymbol{Y}}$, and $\tilde{\boldsymbol{Y}}$ can be simply found using the pseudoinverse of $\boldsymbol{A}$. If we have used the SVD to decompose $\boldsymbol{A}$, then the calculation of $\tilde{\boldsymbol{Y}}$ reduces to

$$\tilde{\boldsymbol{Y}} = \boldsymbol{V}_A \Sigma_A^{-1} \boldsymbol{Y}; \tag{77}$$

similarly, we compute

$$\tilde{Z} = V_B \Sigma_B^{-1} Z. \tag{78}$$

The best way to linearly combine the columns of A (i.e. the combination that minimizes the principal angle of the resulting vector with B) is found in the first column of $\tilde{Y} \equiv [\tilde{y}_1, \ldots, \tilde{y}_r]$ (similarly define $\tilde{Z}$). This linear combination yields the principal vector $u_1 = A\tilde{y}_1$, which is maximally correlated with B when B is linearly combined as $v_1 = B\tilde{z}_1$. In other words, there is no other $\tilde{y}$ (excepting a scale factor of $\tilde{y}_1$) for which a corresponding best fitting $\tilde{z}$ will yield a better correlation between u and v.

IX. References

[1] Ferrara E, Parks T, "Direction finding with an array of antennas having diverse polarizations," *IEEE Trans. Anten. Prop.* Vol. AP-31, pp. 231–236, Mar. 1983.

[2] Golub GH, Pereyra V, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal Numerical Analysis,* vol. 10, pp. 413–432, April 1973.

[3] Golub GH, Van Loan CF, *Matrix Computations,* second edition, Johns Hopkins University Press, 1984.

[4] Krim H, Viberg M, "Two decades of signal processing: The parametric approach," *IEEE Signal Processing Magazine,* July 1996, Vol. 13, No. 4, pp. 67–94.

[5] Luetkenhoener, B, Greenblatt, R, Hamalainen, M, Mosher, J, Scherg, M, Tesche, C, Valdes Sosa, P, "Comparison between different approaches to the biomagnetic inverse problem – workshop report," to appear in Aine, C.J., Flynn, E.R., Okada, Y., Stroink, G., Swithenby, S.J., and Wood, C.C. (Eds.) *Biomag96: Advances in Biomagnetism Research*, Springer-Verlag, New York, 1997.

[6] Mosher JC, Lewis PS, and Leahy RM, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans. Biomedical Eng*, Jun 1992, Vol. 39, pp. 541 – 557.

[7] Mosher JC, "Localization from near-source quasi-static electromagnetic fields," Ph.D. Dissertation, University of Southern California, Los Angeles, CA, March 1993 (Los Alamos Technical Report LA–12622–T; Signal and Image Processing Institute Report #233, University of Southern California).

[8] Mosher JC, Rynne TM, Lewis PS, "MUSIC for thunderstorm localization," *IEEE Proc. 27th Asilo-*

mar Conf. on Signals, Systems, and Computers, Pacific Grove, CA, November 1993, pp. 986–990.

[9] Mosher, JC, Leahy, RM, "Recursively applied MUSIC: A framework for EEG and MEG source localization," Los Alamos National Laboratory Technical Report LA-UR-96-3829, October 1996, submitted for review and possible publication.

[10] Mosher, JC, Leahy, RM, "EEG and MEG source localization using recursively applied (RAP) MUSIC," *Proceedings Thirtieth Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov 3–6, 1996.

[11] Oh, SK, Un, CK, "A sequential estimation approach for performance improvement of eigenstructure-based methods in array processing," *IEEE Trans. Signal Processing*, Jan. 1993, Vol. 41, No. 1, pp. 457–463.

[12] Roy, R, and Kailath, T, "ESPRIT – Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. on ASSP*, ASSP-37, No. 7, pp. 984–995, July 1989.

[13] Schmidt, RO "Multiple emitter location and signal parameter estimation," *IEEE Trans. on Ant. and Prop.* vol. AP-34, pp. 276–280, March 1986. Reprint of the original 1979 paper from the *RADC Spectrum Estimation Workshop.*

[14] Stoica P, Sharman KC, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Signal Processing*, July 1990, Vol. 38, No. 7, pp. 1132–1143.

[15] Stoica, P, Handel, P, Nehorai, A, "Improved sequential MUSIC," *IEEE Trans. Aero. Elect. Sys*, Oct. 1995, Vol. 31, No. 4, pp. 1230–1239.

[16] Vandewalle J, De Moor B, "A variety of applications of singular value decomposition in identification and signal processing," in *SVD and Signal Processing, Algorithms, Applications, and Architectures*, E.F. Deprettere (Editor), Elsevier Science Publishing, Holland, 1988.

[17] Viberg M, Ottersten B, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Processing*, May 1991, Vol. 39, No. 5, pp. 1110–1121.

[18] Viberg M, Ottersten B, Kailath T, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Processing*, Nov. 1991, Vol. 39, No. 11, pp. 2436–2449.

[19] Viberg M, Swindlehurst AL, "Analysis of the combined effects of finite samples and model errors on array processing performance," *IEEE Trans. Signal Processing*, Nov. 1994, Vol. 42, No. 11, pp. 3073–3083

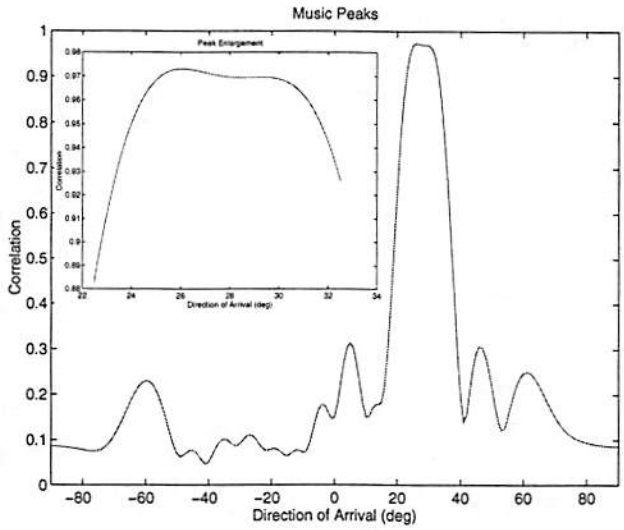

Fig. 1: In this simulation (described in more detail in the sequel), two sources arrive at 25 and 30 degrees at a uniform linear array. The MUSIC correlation is scanned on a fine grid over all putative angles of arrival. The two peaks are not readily discernible, as shown in the inset enlargement of the overall peak. An algorithm must be "trained" to "peak-pick" the second source, here shown at 30 degrees. Such algorithms must also distinguish between the two best peaks and all other "local" peaks, as illustrated in this figure. The projected forms of MUSIC presented in this paper make detection of the second peak more obvious, as well as improve the statistical performance in locating the sources.

Table 1: Comparison of the forms of MUSIC examined in this paper, for estimation of the second source $\hat{\theta}_2 = arg\ max\left\{ Tr\left\{ \Pi_{a(\theta;\hat{\theta}_1)} \hat{\Phi}_s W \hat{\Phi}_s^H \right\} \right\}$, given the first source θ_1 has been estimated using MUSIC. Each technique varies in its modification of the array manifold vector $a(\theta;\hat{\theta}_1)$ and its selection of a weighting matrix W.

	$a(\theta;\hat{\theta}_1)$	W	where
MUSIC	$a(\theta)$	I	$\theta \neq \hat{\theta}_1$
S-MUSIC	$(I - \Pi_{a(\hat{\theta}_1)})a(\theta)$	I	$\Pi_{a(\hat{\theta}_1)} = (a(\hat{\theta}_1)a^H(\hat{\theta}_1))/(a^H(\hat{\theta}_1)a(\hat{\theta}_1))$
IES-MUSIC	$(I - \hat{\rho}\Pi_{a(\hat{\theta}_1)})a(\theta)$	I	$\hat{\rho}$ from *(42)*
RAP-MUSIC	$(I - \Pi_{a(\hat{\theta}_1)})a(\theta)$	$\hat{V}_{a(\hat{\theta}_1)}\hat{\Sigma}^{-2}_{a(\hat{\theta}_1)}\hat{V}^H_{a(\hat{\theta}_1)}$	$\Pi^{\perp}_{a(\hat{\theta}_1)}\hat{\Phi}_s = \hat{U}_{a(\hat{\theta}_1)}\hat{\Sigma}_{a(\hat{\theta}_1)}\hat{V}^H_{a(\hat{\theta}_1)}$

Table 2: Comparison of Analytic Std. Devs. and RMS error. The number of time samples remains constant at 1000, and the correlation γ between the two sources is varied. For each of the 500 Monte Carlo realizations, source 1 (either 25 or 30 degrees) was selected as the source with the highest MUSIC peak. The theoretical and root mean squared (RMS) error of the second source is tabulated. IES-MUSIC is shown both with its scalar ρ set using the true parameters and with its scalar estimated. MUSIC was unreliable in locating the second peak for $\gamma = 0.975$, as illustrated in Fig. 1.

	n	1000		1000		1000		1000		1000		1000	
	γ	0.975		0.950		0.925		0.900		0.800		0.700	
	θ_2 (deg)	25	30	25	30	25	30	25	30	25	30	25	30
	Runs	254	246	236	264	270	230	242	258	241	259	231	269
MUSIC (deg)	Theoretical	0.531	0.555	0.294	0.308	0.215	0.224	0.174	0.182	0.110	0.116	0.088	0.092
	RMS err	- -	- -	0.494	0.389	0.214	0.232	0.165	0.171	0.105	0.115	0.087	0.088
S-MUSIC	Theoretical	0.534	0.559	0.297	0.310	0.216	0.226	0.175	0.183	0.111	0.116	0.089	0.093
	RMS err	0.834	0.818	0.278	0.283	0.184	0.202	0.146	0.160	0.101	0.112	0.082	0.086
IES-MUSIC	Theoretical	0.083	0.087	0.065	0.069	0.062	0.064	0.060	0.063	0.059	0.062	0.059	0.061
	RMS err, ρ	0.461	0.484	0.074	0.077	0.070	0.069	0.066	0.065	0.066	0.071	0.062	0.067
	RMS err, $\hat{\rho}$	0.946	0.919	0.189	0.184	0.110	0.116	0.084	0.084	0.071	0.075	0.063	0.068
RAP-MUSIC	RMS err	0.879	0.863	0.150	0.153	0.083	0.093	0.070	0.070	0.067	0.069	0.061	0.065

Table 3: The number of time samples is now varied, while the correlation γ between the two sources is held constant at 0.9. As in Table 2, for each of the 500 Monte Carlo realizations, source 1 (either 25 or 30 degrees) was selected as the source with the highest MUSIC peak. The theoretical root mean squared error of the second source is tabulated. MUSIC was unreliable in locating the second peak for $N = 100$. We observe RAP-MUSIC yielding consistently superior RMS error.

	n	100		200		400		1000	
	γ	0.9		0.9		0.9		0.9	
	θ_2 (deg)	25	30	25	30	25	30	25	30
	Runs	262	238	240	260	250	250	253	247
MUSIC (deg)	Theoretical	0.550	0.575	0.389	0.407	0.275	0.288	0.174	0.182
	RMS err	- -	- -	1.247	1.174	0.446	0.464	0.163	0.165
S-MUSIC	Theoretical	0.554	0.579	0.391	0.410	0.277	0.290	0.175	0.183
	RMS err	0.889	0.932	0.522	0.515	0.267	0.301	0.148	0.161
IES-MUSIC	Theoretical	0.190	0.199	0.134	0.141	0.095	0.100	0.060	0.063
	RMS err, ρ	0.358	0.404	0.158	0.176	0.096	0.097	0.061	0.068
	RMS err, $\hat{\rho}$	0.966	1.056	0.516	0.495	0.211	0.245	0.081	0.084
RAP-MUSIC	RMS err	0.785	0.870	0.370	0.397	0.146	0.187	0.071	0.073