Research Article

Antenna Array Calibration for a General Linear Transformation of the Array Manifold

Benjamin Friedlander

Department of Electrical Engineering, University of California, Santa Cruz, CA 95064, USA

Correspondence should be addressed to Benjamin Friedlander; benjamin.friedlander@gmail.com

Received 1 March 2018; Revised 24 September 2018; Accepted 2 October 2018; Published 9 April 2019

Academic Editor: Hervé Aubert

Copyright © 2019 Benjamin Friedlander. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider the problem of calibrating an antenna array using received signals from either known or unknown directions. A complex matrix is used to map the assumed array manifold to the true array manifold. This matrix may be completely unstructured. Algorithms are presented for joint estimation of this matrix and the directions-of-arrival, using data collected during multiple time intervals.

1. Introduction

Super-resolution direction-finding techniques, primarily those based on eigenstructure methods, have been discussed extensively in the literature since the early 1980s. These methods achieve their superior performance by making use of the so-called array manifold—the response of an antenna array to a signal received from a given direction. Direction-finding depends strongly on the accuracy of this array manifold. Therefore, the array must be carefully calibrated. It is well established by now that array calibration is the key issue in making super-resolution techniques work.

Array calibration requires a parametric model for the array manifold. During the calibration process, the parameters of this model will be adjusted or estimated to create a better fit between the assumed and true array manifolds. Different parametric models have been studied in the literature attempting to correct different sources of error. These may include imprecisely known antenna locations, unknown gains and phases in the cabling or the receivers, and unmodeled mutual coupling effects.

In this paper, we consider the case where the true array manifold is assumed to be a fixed (i.e., direction independent) linear function of some reference manifolds (i.e., the assumed array manifold). Specifically, \( \mathbf{a}_t(\phi, \psi) = \mathbf{D} \mathbf{a}_r(\phi, \psi) \), where \( \mathbf{a}_r(\phi, \psi) \) denotes the reference manifold, \( \mathbf{a}_t(\phi, \psi) \) is the true array manifold, \( \phi \) and \( \psi \) are the source azimuth and elevation, respectively, and \( \mathbf{D} \) is a complex \( M \times M \) matrix, with \( M \) being the number of antennas. No special structure is assumed for the matrix \( \mathbf{D} \), which will generally be the case if there are combined effects of gain, phase, and mutual coupling (however, in the appendix, we show how certain structures can be easily accommodated). Note that for a given array with manifold \( \mathbf{a}_c(\phi, \psi) \), the matrix \( \mathbf{D} \) depends on the selection of the reference manifold. The fact that \( \mathbf{D} \) is fixed implies that the array geometry has no uncertainties and only gain, phase, and mutual coupling uncertainties are present.

A common choice made in the literature for the reference manifold is the isolated or uncoupled array manifold—the manifold computed assuming there is no coupling between the antennas. This is usually written as \( \mathbf{a}_c(\phi, \psi) = \mathbf{C} \mathbf{a}_{uc}(\phi, \psi) \), where \( \mathbf{C} \) is the so-called mutual coupling matrix (MCM), \( \mathbf{a}_{uc}(\phi, \psi) \) is the uncoupled array manifold, and \( \mathbf{a}_c(\phi, \psi) \) is the coupled array manifold which is assumed to be the true manifold (although this is only approximately true because the matrix relating the true and uncoupled manifolds is actually direction dependent). Using this model, array calibration methods estimate the MCM, see, e.g., [1–8]. In order to reduce the number of calibration parameters which need to be estimated, the MCM is usually assumed to have some specific structures. Uniformly spaced linear arrays are assumed to have a Toeplitz structure and uniform circular arrays—a circulant structure [9]. To further reduce the number of parameters, the MCM is often also assumed to be a banded matrix [10]. Yet another simplification is to
assume that the mismatch between the assumed and true manifolds may be represented by a diagonal matrix whose elements correspond to unknown complex gains. This model is studied frequently in the array calibration literature, see, e.g., [11, 12]. In this paper, we will focus on the general case where \( \mathbf{D} \) is a full and completely unstructured matrix.

There are several issues with the models discussed above. To begin with, the assumption that the mutual coupling matrix is completely unknown seems unreasonable. Formulas (based on approximations) for mutual coupling in commonly used arrays, such as uniform linear arrays and uniform circular arrays, are readily available and have been known for a long time, as is the methodology for calculating the mutual coupling for general array configurations [13–16]. Furthermore, well-developed and widely used numerical electromagnetic codes (NECs) are available both commercially and for free to calculate the array manifold (including mutual coupling effects) for arbitrary antenna elements and array configurations [17–20]. There is no logical reason to consider the mutual coupling effect to be completely unknown and to estimate it “from scratch” nor is this being done in practice. High-resolution direction-finding systems make use of the best available array manifold, be it based on analytical formulas, NEC software, or measurements, all of which already include the effects of mutual coupling.

Another issue is that the models used for the mutual coupling matrix \( \mathbf{C} \) are almost always nonphysical. For example, it is commonly and incorrectly assumed that \( \mathbf{C} \) for a uniform linear array has a Toeplitz structure [6, 21–23]. This is only true asymptotically as the number of antennas tends to infinity. Another common assumption is that the MCM \( \mathbf{C} \) is a banded matrix. While it is true that off-diagonal coupling terms tend to be relatively small, they are generally nonnegligible and should not be assumed to be zero. The nonphysical nature of the MCM is further manifested by the choices made for the coefficients of MCM (the entries of \( \mathbf{C} \)) in simulations. With few exceptions, these are chosen in an ad hoc manner or even drawn from a random Gaussian distribution [6] and are not related to any physical model of the antenna array.

Another issue is that the mutual coupling matrix is not a unique and well-defined concept. The name is generally associated with a constant matrix \( \mathbf{C} \) which attempts to relate the uncoupled and true manifolds. Note, however, of the following. (i) Different choices can be made for the uncoupled manifold and the coupled manifold, each leading to a different MCM. (ii) The array manifolds depend on the choice of their phase center and are not uniquely defined. These choices effect the resulting MCM. (iii) The mapping of an uncoupled (isolated) array manifold to the true array manifold is generally direction dependent and can only be approximated by a constant matrix [24–26]. These ambiguities lead to a considerable confusion in the literature dealing with mutual coupling matrices.

In this paper, we take a different approach. We assume that the reference manifold \( \mathbf{a}_r(\phi, \psi) \) is the best available manifold obtained from the measurements or NEC software and it already includes the effects of mutual coupling. Therefore, \( \mathbf{D} \) is not a mutual coupling matrix but rather a way of capturing any mismatch between the true and the reference array manifolds. Such mismatch may be caused by differences between the characteristics of the antennas used to calculate \( \mathbf{a}_r(\phi, \psi) \) and the characteristics of the actual antennas, approximations made in the software used to calculate the reference manifold, and items which have not been included in the model (such as the effect of the ground or the antenna support). In general, \( \mathbf{D} \) does not possess any particular structure and it is necessary to estimate all of its \( M \times M \) entries. The signal model is discussed further in Section 2.

Having specified the calibration matrix, its parameters need to be estimated from available data. The data may originate from signals whose directions are known or unknown. The latter case is usually called self-calibration or blind calibration. Current self-calibration methods use data collected over a single time interval to form a sample covariance matrix of the array outputs and use it to jointly estimate the unknown directions and mutual coupling parameters [7, 26–31]. These techniques assume the presence of a sufficient number of cochannel signals in order to operate. The required number of signals depends on the number of unknown calibration parameters and other factors. This assumption is highly impractical because the number of active cochannel signals cannot be guaranteed to be sufficient for the self-calibration to work. This fact alone greatly limits the usefulness of most (perhaps all) self-calibration techniques presented in the literature to date.

To alleviate this fundamental difficulty, we present a novel self-calibration algorithm which operates on data collected during multiple time intervals in a given frequency channel. The directions of the active signals and their number may vary from one interval to the next. We assume that the calibration parameters change relatively slowly and may be assumed to be constant during the entire data collection time. This assumption holds in all practical direction-finding scenarios of which we are currently aware.

The probability that a sufficient number of signals is present in a given channel is dramatically increased when multiple time intervals are being used. This increases significantly the usefulness of self-calibration in practical direction-finding scenarios.

In the following sections, we describe the signal model used herein (Section 2) and develop algorithms for estimating the matrix \( \mathbf{D} \) (Section 3) and for jointly estimating the signal directions-of-arrival and the matrix \( \mathbf{D} \) (Section 4). We then discuss various aspects of the of these algorithms (Section 5). We emphasize the fact that the work presented in the following differs from prior works in two important respects: (i) the matrix \( \mathbf{D} \) may be completely unstructured and is not necessarily related to the mutual coupling matrix; (ii) the estimation of the unknown parameters is based on data collected over multiple time intervals.

2. The Signal Model

Consider the data collected at the output of a uniformly polarized array (The extension of this work to diversely
polarized arrays is nontrivial and requires a separate treatment.) with $M$ antennas at a particular frequency channel during $P$ different time intervals. During the $p$-th time interval, $K_p$ signals were active in the channel. The narrowband baseband model for these signals is given by

$$y_p(t) = \sum_{k=1}^{K_p} D_a(\phi_{k,p}, \psi_{k,p}) s_{k,p}(t) + w_p(t), \quad (1)$$

where $y_p(t)$ is the $M \times 1$ vector of received signals collected during time interval $p$, $D$ is an $M \times M$ matrix whose elements are unknown complex parameters, $a(\phi, \psi)$ is the reference array manifold, $(\phi_{k,p}, \psi_{k,p})$ are the signal directions-of-arrival, $s_{k,p}(t)$ are the impinging signal waveforms, and $w_p(t)$ is an $M \times 1$ vector of additive white complex Gaussian noise with zero mean and variance $\eta_p$. The complex elements of $D$ are assumed to be constant and independent of the signal directions-of-arrival during the entire data collection time. We assume that the signals present during different time intervals come from different directions.

We are given $N_p$ snapshots of the received signal vectors $y_p(1), \ldots, y_p(N_p)$. Writing equation (1) for all of the snapshots using a more compact matrix notation, we get

$$Y_p = DA_p S_p + W_p, \quad (2)$$

where

$$Y_p = [y_p(1), \ldots, y_p(N_p)],$$
$$W_p = [w_p(1), \ldots, w_p(N_p)]$$

are the received signal and noise $M \times N_p$ matrices,

$$A_p = [a_1(\phi_{1,p}, \psi_{1,p}), \ldots, a_{K_p}(\phi_{K,p}, \psi_{K,p})], \quad (4)$$

are an $M \times K_p$ matrix containing the array manifold vectors for the signal directions, and

$$S_p = [s_p(1), \ldots, s_p(N_p)],$$
$$s_p(t) = [s_{1,p}(t), \ldots, s_{K,p}(t)]^T \quad (5)$$

are a $K_p \times N_p$ matrix containing the sampled signal waveforms. The signals are assumed to be jointly stationary, uncorrelated, and to have a covariance matrix

$$R_{s,p} = E\{s_p(t) s_p(t)^H\}. \quad (6)$$

The covariance matrix of the received signals is therefore

$$R_p = E\{y_p(t) y_p(t)^H\} = DA_p R_{s,p} A_p^H + \eta_p I \quad (7)$$

where

$$\begin{align*}
\text{Given the measurements } Y_p, & \quad p = 1, \ldots, P, \text{ we want to estimate the unknown parameters. We consider two situations:} \\
& \text{(i) the DOAs } (\phi_{k,p}, \psi_{k,p}) \text{ are known for all active signals in all} \\
& \text{of the time intervals, but the elements of } D \text{ are unknown and} \\
& \text{need to be estimated, and (ii) neither DOAs nor } D \text{ is known} \\
& \text{and needs to be jointly estimated. Our main interest is in case} \\
& \text{(ii), but it is useful to study first the easier case (i).} \\
\end{align*}$$

### 3. Estimating the Matrix $D$ Known

#### Signal Directions

In this section, we present an algorithm for estimating the parameters of $D$ from the received signals defined in equation (2). The algorithm is based on the fact that the true array manifold evaluated at the DOA of any one of the signals is orthogonal to the noise subspace. This is the same principle which is used by the MUSIC algorithm [32–34] to estimate DOAs, except that here we assume that the DOAs are known and use them to estimate $D$.

Let $R_p$ be the covariance matrix of the received signals during time interval $p$ defined in equation (7). The noise subspace $U_{s,p}$ associated with this covariance matrix is found by performing the singular value decomposition

$$R_p = U_p R_{s,p} U_p^H, \quad (8)$$

and then separating $U_p$ into a signal subspace $U_s,p$ consisting of the first $K_p$ singular vectors and a noise subspace $U_{n,p}$ consisting of the remaining $M - K_p$ singular vectors,

$$U_p = [U_s,p, U_{n,p}]. \quad (9)$$

It is straightforward to show that the true array manifold vectors corresponding to the signal directions $D_a(\phi_{k,p}, \psi_{k,p})$ are orthogonal to the noise subspace $U_{n,p}$. Thus, when both $D$ and $D_a(\phi_{k,p}, \psi_{k,p})$ have their true values, the norm of $U_{n,p}^H D_a(\phi_{k,p}, \psi_{k,p})$ will be zero.

Since the exact covariance matrix $R_p$ is not known, it will be replaced by $\hat{R}_p$, the sample covariance matrix

$$\hat{R}_p = \frac{1}{N_p} Y_p Y_p^H, \quad (10)$$

and the exact noise subspace $U_{n,p}$ will be replaced by its approximate counterpart $\hat{U}_{n,p}$. In this case, $D_a(\phi_{k,p}, \psi_{k,p})$ will be only approximately orthogonal to the estimated noise subspace $\hat{U}_{n,p}$. These facts motivate the use of the following cost function $f(D)$ to estimate $D$. Let

$$f(D) = \sum_{p=1}^{P} \sum_{k=1}^{K_p} \beta_{k,p} \| D^H \hat{U}_{n,p} \| \hat{U}_{n,p}^H D_a(\phi_{k,p}, \psi_{k,p}) \|.$$

(11)

If the estimated noise subspaces were exact, i.e., $\hat{U}_{n,p} = U_{n,p}$, the cost function $f(D)$ will equal zero for the correct value of $D$. Using an incorrect value of the matrix $D$ will increase the value of the cost function because $D_a(\phi_{k,p}, \psi_{k,p})$ will no longer be orthogonal to the noise subspace. Thus, the solution of $f(D) = 0$ will provide the desired estimate of the matrix $D$. Note, however, that the solution is not unique. An obvious ambiguity is caused by the fact that if $D$ is a solution,
so is \( c \mathbf{D} \), where \( c \) is an arbitrary complex scalar. There may be other ambiguities related to specific choices of the array manifold \( \mathbf{a}_p(\phi_{k,p}, \psi_{k,p}) \). For example, when using a uniform linear array and choosing the reference manifold to be the omnidirectional manifold (which ignores mutual coupling), rotating all the DOAs by the same amount is equivalent to changing the phases of the columns of \( \mathbf{D} \). However, these ambiguities are not present when using the analytic manifold (which includes the effects of mutual coupling) [25].

In practice, the estimated noise subspace \( \hat{\mathbf{U}}_{n,p} \) rather than its exact value \( \mathbf{U}_{n,p} \) will be used in which case \( \mathbf{D} \) can still be estimated by minimizing this cost function, but the minimum will no longer equal zero. To do this, note that (11) can be rewritten as

\[
 f(\mathbf{d}) = \mathbf{d}^H \mathbf{Q} \mathbf{d},
\]

where \( \mathbf{d} = \text{vec}(\mathbf{D}^T) \), and

\[
 \mathbf{Q} = \sum_{p=1}^{K_p} \sum_{k=1}^{K_p} \left( (1 \otimes \mathbf{a}^T(\phi_{k,p}, \psi_{k,p}))^H \hat{\mathbf{U}}_{n,p} \hat{\mathbf{U}}_{n,p}^H (1 \otimes \mathbf{a}^T(\phi_{k,p}, \psi_{k,p})) \right).
\]

This can be shown by rewriting each of the terms in equation (11) as follows (dropping the indices \( k \) and \( p \) for notational convenience):

\[
 \mathbf{D} \mathbf{a}_k(\phi, \psi) = \begin{bmatrix} d_1^T \mathbf{a}_k(\phi, \psi) \\ \vdots \\ d_{M}^T \mathbf{a}_k(\phi, \psi) \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T(\phi, \psi) \\ \vdots \\ \mathbf{a}_{M}^T(\phi, \psi) \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_{M} \end{bmatrix}.
\]

\[
 \mathbf{D} \mathbf{a}_k(\phi, \psi)^T \mathbf{D} = \begin{bmatrix} \mathbf{a}_1^T(\phi, \psi) & \cdots & \mathbf{a}_{M}^T(\phi, \psi) \end{bmatrix} \begin{bmatrix} d_1^T \\ \vdots \\ d_{M}^T \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T(\phi, \psi) \\ \vdots \\ \mathbf{a}_{M}^T(\phi, \psi) \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_{M} \end{bmatrix}.
\]

where \( d_m^T \) are the rows of \( \mathbf{D} \), or equivalently, and \( d_m \) are the columns of \( \mathbf{D}^T \).

We assume that \( \mathbf{d} \) has unit norm (or some other nonzero norm), i.e., \( \mathbf{d}^H \mathbf{d} = 1 \) to avoid the trivial solution \( \mathbf{d} = 0 \). This cost function will be minimized when \( \mathbf{d} \) is the eigenvector of the matrix \( \mathbf{Q} \) corresponding to its smallest eigenvalue.

In order to get a unique solution \( \mathbf{d} \), the matrix \( \mathbf{Q} \) should have a unique smallest eigenvalue. In the case of the exact noise subspace (i.e., \( \hat{\mathbf{U}}_{n,p} = \mathbf{U}_{n,p} \)), this eigenvalue will be zero. In other words, to guarantee uniqueness, the rank of \( \mathbf{Q} \) should be \( M - 1 \), so that it has only one zero eigenvalue. Note that each of the terms comprising the matrix \( \mathbf{Q} \) has the rank of the subspace \( \mathbf{U}_{n,p} \) which is \( M - K_p \). Therefore, the rank of \( \mathbf{Q} \) is no more than the minimum of \( r_Q = \sum_{m=1}^{K_p} (M - K_p) K_p \) and \( M^2 - 1 \). Therefore, a necessary (but not sufficient) condition for uniqueness is that \( r_Q \geq M^2 - 1 \).

If \( K_p = 1 \), the rank condition becomes \( P(M - 1) \geq M^2 - 1 \) or \( P \geq M + 1 \). If \( K_p = K \), we have \( P \geq (M^2 - 1)/(K(M - K)) \). For sufficiently large values of \( M \), this translates into \( P > M/K \). This reinforces the fact mentioned earlier that estimating the unstructured calibration matrix \( \mathbf{D} \) requires using data from multiple time intervals. This cannot be done by current techniques which rely on data collected in a single time interval.

4. Joint Estimation of the Matrix \( \mathbf{D} \) and Signal Directions

So far, we assumed that the DOAs of the signals used to estimate the matrix \( \mathbf{D} \) are known. Next, we consider the case where the DOAs are unknown and need to be estimated jointly with the calibration parameters \( \mathbf{D} \). Note that if \( \mathbf{D} \) is known, the estimation of the DOAs is relatively straightforward. Since the DOAs are assumed to be different at different time intervals, their estimation must be done separately for each time interval. This can be accomplished by any standard DOA estimation algorithm, such as MUSIC [32–34]. The joint estimation of \( \mathbf{D} \) and the DOAs is a much more difficult problem. The maximum likelihood estimator can be easily derived but is computationally intractable because it involves the optimization of a high-dimensional nonlinear function. One may try instead to use the cost function in equation (11) which can be written as \( f(\mathbf{D}, \Theta) \), where \( \Theta \) is the set of DOAs of all of the signals during all of the observation periods. Attempting to jointly minimize this function over \( \mathbf{D} \) and \( \Theta \) again involves a computationally intractable high-dimensional optimization problem. This cost function lends itself to a computationally efficient solution by iterating between the estimation of the calibration parameters \( \mathbf{D} \) and the estimation of the DOAs. It should be noted that there is no formal proof or guarantee of convergence for this iterative algorithm. However, experiments show that the iteration always converges to a local minimum and therefore yields useful estimates provided that the initial conditions of the iteration are close enough to the true values. This situation, while less than satisfactory, is typical in iterative parameter estimation techniques.

To start the iteration, an initial value for \( \mathbf{D} \) must be selected. In the absence of prior information, we let \( \mathbf{D} = \mathbf{I} \) be the initial value. Next, the MUSIC algorithm is applied to estimate the signal DOAs in each period. (If only one signal or several well-separated signals are active at any given time, it is possible, even preferable, to use beamforming rather than the MUSIC algorithm.) The MUSIC "spectrum" is then given by

\[
 S_p(\phi, \psi) = \frac{1}{\mathbf{a}_p^H(\phi, \psi) \mathbf{D}^H \mathbf{U}_{n,p} \mathbf{U}_{n,p}^H \mathbf{D} \mathbf{a}_p(\phi, \psi)}. \tag{15}
\]

The direction estimates are obtained by searching for the peaks of \( S(\phi, \psi) \). The estimates \( \hat{\phi}_{k,p}, \hat{\psi}_{k,p} \) correspond to the values of \( \phi \) and \( \psi \) at the peak locations. Given the DOA estimates, we recompute the matrix \( \mathbf{Q} \) (13) and obtain a new estimate for the calibration parameters \( \mathbf{D} \). This process is repeated until convergence. As is usually the case with nonlinear optimization, the algorithm converges to a local minimum which may or may not be the desired global minimum. Convergence to the global minimum requires that the initial value of \( \mathbf{D} \) be sufficiently close to its true value. This issue is discussed further in Section 5.
The iterative algorithm is easily modified to incorporate prior knowledge of signal directions. If some of the signals arrive from known directions their estimates will be replaced by their known values before estimating $D$. If all the signals in a given period have known DOAs, the DOA estimation step for that period may be skipped.

In the discussion so far we considered the case where the matrix $D$ is completely unstructured. In the appendix, we present a version of the algorithm which incorporates certain linear equality constraints. Using these constraints, it is possible to accommodate a variety of structures of the matrix $D$.

5. Performance Evaluation

In this section, we study the performance of the calibration algorithms described earlier, mainly by using numerical examples. In all of these examples, we use an $M = 8$ element circular array with radius of one wavelength. The antennas are identical center-fed vertical dipoles, one wavelength long, and one thousandths of a wavelength thick. The antenna ports are loaded by a 50 Ω impedance. The MATLAB™Antenna Toolbox was used to compute the array manifold $a(\phi, \psi)$. All the antennas and signal sources were assumed to lie in the horizontal plane where $\psi = 0^\circ$. This was done to simplify the plotting and interpretation of the results and does not represent a restriction of the algorithms. In these examples, the directions of arrival were chosen at random in the range $-90^\circ \leq \phi \leq 90^\circ$. To simplify the discussion, all periods had the same number of $K$ of signals (the algorithm can handle different numbers of signals). The signals were assumed to be received with equal power and have a common signal-to-noise ratio (SNR). The matrix was specified as $D = I + \Delta D$, where the elements of $\Delta D$ were selected independently from a complex Gaussian random distribution with mean 0 and standard deviation $\sigma_p$. The initial value of $D$ used by the iterative algorithm was $D = I$.

5.1. The Estimation of $D$. We start by demonstrating the performance of the algorithm for estimating the matrix $D$, see equations (12) and (13). Consider first the case where the signal directions are perfectly known and the exact covariance matrices $R_p$ are used (i.e., the infinite data case). We calculate the error between the true matrix $D$ and the its estimate $\hat{D}$. The error is given by the normalized Frobenius norm of the difference,

$$\epsilon_D = \frac{\|D - \hat{D}\|_F}{\|D\|_F}. \quad (16)$$

Figure 1 depicts $\epsilon_D$ as a function of the number of periods $P$ used to estimate the matrix $D$, for three cases: having $K = 1$, $K = 2$, and $K = 3$ signals present during each period. As expected, when a sufficient number of periods is being used, the error is very small, ideally zero. When the number of periods is insufficient to provide the proper rank for the matrix $Q$, the estimation fails and $\epsilon_D$ will be close to unity. Recall the rank condition introduced earlier which stated that the smallest value of $P$ needed is

$$(M^2 - 1)/(K(M - K)).$$

Evaluating this condition for $M = 8$ and $K = 1, 2, 3$, we get $P_{\text{min}} = 9, 6, 5$, respectively. This matches the results seen in the figure. Recall that the rank condition is only necessary but not sufficient. Thus, in general, $P_{\text{min}}$ represents only a lower bound on the number of periods required for correctly estimating $D$, as will be apparent in some of the following examples.

Next, we consider the case where $D$ is computed using the sample covariance matrices $\hat{R}_p$, rather than exact covariances. Figure 2 depicts $\epsilon_D$ averaged over 100 noise realizations, as a function of the number of periods $P$ with $K = 2$ signals in each period, for SNRs of 10, 20, and 30 dB. The DOAs are assumed to be known. In this case, a nonzero estimation error $\epsilon_D$ is expected. Note also that the number of periods required to minimize $\epsilon_D$ is somewhat greater than the value $P_{\text{min}} = 6$ calculated using the rank condition.

The error in estimating $D$ is a function of the errors in the estimated DOAs. To check the sensitivity of $\epsilon_D$ to DOA errors, we added to each of the DOAs an error term drawn from a Gaussian distribution with zero mean and variance $\sigma_\theta$ and calculated the resulting $D$ estimation error. Results were averaged over 300 independent random azimuth errors. The exact covariance matrices $R_p$ corresponding to the erroneous DOAs were used. Figure 3 depicts the average $\epsilon_D$ as a function of $\sigma_\theta$ using $P = 25, 35, 45$ periods, with $K = 2$ signals per period. As is to be expected, the sensitivity of $\epsilon_D$ to azimuth errors (measured by the slope of the curve) reduces as the number of periods being used increases.

5.2. The Impact of $D$ on DOA Estimation. The objective of the calibration procedure is to estimate $D$ in order to reduce the

![Figure 1: The error in estimating D as a function of the number of periods P, for different numbers of signals. Known DOAs, exact covariance matrices R_p.](image)
mismatch between the true and assumed array manifolds, so as to improve the DOA estimation. This raises the question: how much of an effect does the size of the D mismatch have on the estimated DOAs? To study this, we consider the case where the true manifold is \( D = I + \Delta D \), with \( \Delta D \) being a random complex Gaussian matrix as described earlier. The DOAs are estimated using an assumed array manifold with \( D = I \). In other words, \( \Delta D \) represents the mismatch. The DOAs are estimated for many random realizations of \( \Delta D \), for different values of \( \sigma_D \), the variance of the elements of \( D \).

Figure 4 depicts the standard deviation of the DOA estimates as a function of the size of the mismatch \( \sigma_D \). The parameters of this experiment were 100 independent realizations of \( \Delta D \) for each value of \( \sigma_D \), \( P = 10 \) periods, and with one signal in each period. The exact covariances \( R_p \) were used.

Figure 5 depicts the probability of resolving two equal power signals separated by an angle \( \Delta \phi \) as a function of (uncalibrated) mismatch \( D \). The parameters of this experiment were 1000 independent realizations of \( \Delta D \) for each value of \( \sigma_D \), \( P = 1 \) periods, and with two signals in each period. The exact covariances \( R_p \) were used, so the DOA errors and loss of resolution are due solely to the \( D \) mismatch. In each case, the locations of the peaks of the MUSIC spectrum were determined. If two peaks were found within a specified angular range containing the true DOAs, the signals were considered to be resolved. The range was \([\phi_1 - 2 \Delta \phi, \phi_2 + 2 \Delta \phi]\), where \( \phi_1 \) and \( \phi_2 \) are the signal directions where \( \phi_1 < \phi_2 \), and \( \Delta \phi = \phi_2 - \phi_1 \) is the angular separation. If less than two peaks were in the range, the signals are considered to be not resolved. The probability of resolution is defined as the fraction of the 1000 runs in which resolution was observed. The probability of resolution decreases as the angular separation of the two signals becomes smaller. As is to be expected, the impact of mismatch on performance is larger for closely spaced signals than on a single signal (or widely spaced signals). For example, a mismatch of size \( \sigma_D = 0.1 \) causes a modest DOA error of about 1° when one signal is present but makes it almost impossible to resolve when two signals are separated by 4°.
5.3. The Joint Estimation of \( D \) and the DOAs. In the examples above, we considered the estimation of DOAs and \( D \) separately. Next, we consider the joint estimation of the matrix \( D \) and signal directions using the iterative approach described in Section 4.

Figure 6 demonstrates the convergence of the estimated parameters. This figure depicts the average error in \( D \) and the estimated DOAs as a function of the iteration number. The error was calculated by taking the Frobenius norm of \( \hat{D} - D \) and normalizing it by the Frobenius norm of \( D \). The vertical lines are error bars corresponding to \( \pm 1 \) standard deviation of the error. The parameters of this experiment were 100 independent realizations of \( \Delta D \) for \( \sigma_D = 0.05, P = 40 \) periods, with two signals in each period. The exact covariances \( R_p \) were used. The errors are seen to converge within a few iterations.

As mentioned earlier, global convergence of this iterative algorithm is not guaranteed. If the initial mismatch between the assumed and true values of \( D \) is sufficiently large, the algorithm will fail to converge. To evaluate the range of value over which convergence will occur, we repeat the previous experiment (Figure 6) for different values of \( \sigma_D \) and evaluate the probability of convergence \( P_c \), defined as the fraction of the 100 independent runs in which convergence was observed. The result is depicted in Figure 7. In this example, reliable convergence is observed for \( \sigma_D \leq 0.08 \) for \( P = 20 \) periods and \( \sigma_D \leq 0.95 \) for \( P = 40 \). For larger values of \( \sigma_D \), the algorithm fails to converge. As can be seen, there is a sharp transition between convergence and nonconvergence. As seen from the figure, increasing the number of periods used in the joint estimation increases somewhat the capture range of the algorithm (i.e., the range of initial errors for which it converges).

If some of the signals have known DOAs, the algorithm can make use of them with a consequent improvement in performance (as explained in Section 4). In particular, the capture range of the algorithm increases, as demonstrated
in Figure 8. Here, the probability of convergence is plotted for two cases: one in which all the DOAs are unknown and the other in which the signal in 5 of the \( P = 20 \) periods are known. The presence of some known DOAs more than doubled the capture range. If there is a sufficient number of periods with known DOAs (according to the rank condition), it is possible to compute the matrix \( D \) directly, without iteration. Figure 9 illustrates the positive effect of the calibration process on the resolution of closely spaced signals. The presence of some known DOAs more than doubled the capture range. If there is a sufficient number of periods with known DOAs (according to the rank condition), it is possible to compute the matrix \( D \) directly, without iteration. Figure 9 illustrates the positive effect of the calibration process on the resolution of closely spaced signals. The figure depicts the MUSIC spectrum for one of the periods containing two closely spaced signals, before and after the estimation of \( D \) (using the same parameters as in Figures 6 and 7 with \( \sigma_D = 0.09 \)).

The signals were not resolvable with the initial mismatch but are clearly resolved once \( D \) has been estimated. All of the previous figures were generated using exact covariance matrices, corresponding to the infinite data case. Figure 10 depicts the convergence of the algorithm when the sample covariance matrices are used. These were calculated based on \( N = 100 \) snapshots of the array output. The plot shows the standard deviation of the DOA estimation errors as a function of the iteration number for different signal-to-noise ratios. Unlike in Figure 6, \( \sigma_{az} \) converges to a nonzero value, because noisy covariance matrices \( \hat{R} \) are used here. The value to which \( \sigma_{az} \) converges decreases when the SNR increases, as is to be expected.

6. Conclusions

We presented an algorithm for estimating the mismatch between the true and assumed array manifolds, i.e., calibrating the array given measurements of signals impinging on the antenna array. Unlike previously published work on calibration, we consider a general form of mismatch represented by a matrix \( D \) which may have no particular structure. In order to estimate all of the unknown entries of this matrix, we develop an algorithm which uses data collected during multiple observation intervals. The algorithm may be used to estimate the matrix \( D \) given signals whose directions are known or to jointly estimate the matrix \( D \) and the signal directions. To do the latter in a computationally feasible way, the algorithm iterates between estimating directions and estimating the matrix \( D \). The convergence characteristics of the algorithm were studied by computer simulation.

Appendix

A. Estimation of \( D \) with Linear Constraints

Consider the case where the elements of the matrix \( D \) are subject to a set of linear constraints. In this case, we want to minimize \( d^H Qd \) (see equation (12)) subject to the constraint \( Bd = c \), where \( B \) is an \( r \times M^2 \) matrix, \( c \) is \( r \times 1 \), and \( r < M^2 \) is the number of linear constraints. Assume that \( d \) has a nonzero norm. These linear constraints can represent a variety of structures which \( D \) may have. Some examples are as follows:

(i) \( D \) is a Banded Matrix. Each row of \( B \) consists of zeros except for a single \(+1\) corresponding to one of the zero elements of \( D \), \( c = 0 \), and \( r \) equals the number of zero elements in \( D \).

(ii) \( D \) is a Diagonal Matrix. This is a special case of a banded matrix with \( r = M(M - 1) \).

(iii) \( D \) is a Symmetric Matrix. Each row of \( B \) consists of zeros except for two entries \(+1\) and \(-1\) corresponding to the \((i,j)\) and \((j,i)\) elements of \( D \), \( c = 0 \), and \( r = M(M - 1)/2 \).
(iv) **D is a Hermitian Matrix.** In this case, \( d \) and \( Q \) need to be written using real numbers only, so that the real and imaginary parts of each complex number appear separately. So, \( d = [d^T_r, d^T_i]^T \), where \( d_r \) is the vectorized version of the real part of \( D \) and \( d_i \) is the vectorized version of the imaginary part of \( D \). Then, each row of \( B \) consists of zeros except for two entries +1 and −1 corresponding to the \((i,j)\) and \((j,i)\) elements of the real part of \( D \) or two entries +1 and −1 corresponding to the \((i,j)\) and \((j,i)\) elements of the imaginary part of \( D \). Here again \( c = 0 \).

(v) **D is a Toeplitz Matrix.** Each row of \( B \) consists of zeros except for two entries +1 and −1 corresponding to the \((i,j)\) and \((i+1, j+1)\) elements of \( D \), where \( 1 \leq i, j \leq M - 1 \), \( c = 0 \), and \( r = (M - 1)^2 \).

(vi) **D is a Circulant Matrix.** Each row of \( B \) consists of zeros except for two entries +1 and −1 corresponding to the equal elements of \( D \) in two adjacent rows (or columns), \( c = 0 \) and \( r = M(M - 1) \).

Note that all of these examples have \( c = 0 \). Therefore, we will solve the constrained minimization problem for this special case and mention the general case of \( c \neq 0 \) only briefly.

A.1. **Special Linear Constraints** \( c = 0 \). Consider the case where the linear constraints are given by the following:

\[
Bd = 0. \tag{A.1}
\]

Let

\[
\tilde{B} = \begin{bmatrix} B \\ B^c \end{bmatrix}, \tag{A.2}
\]

where \( \tilde{B} \) is a full rank \( M^2 \times M^2 \) matrix and \( B^c \) is the orthogonal complement of \( B \). In other words, the rows of \( B^c \) are orthogonal to the rows of \( B \). We can now rewrite the cost function to be minimized as follows:

\[
d^H Q d = d^H B^H B^{-H} B^c \tilde{B} d = \tilde{d}^H \tilde{Q} \tilde{d}. \tag{A.3}
\]

Note that

\[
\tilde{B} d = \begin{bmatrix} Bd \\ B^c d \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{d} \end{bmatrix}, \tag{A.4}
\]

where

\[
\tilde{d} = B^c d \tag{A.5}
\]

is a low dimensional version of \( d \). Without loss of generality, we assume that \( d \) has a unit norm. Then, \( \tilde{d}^H \tilde{Q} \tilde{d} \) will be minimized if \( \tilde{d} \) is the eigenvector of \( \tilde{Q} \) corresponding to its smallest eigenvalue. Once \( \tilde{d} \) is computed, we get \( d \) from

\[
d = \tilde{B}^{-1} \begin{bmatrix} 0 \\ \tilde{d} \end{bmatrix}. \tag{A.6}
\]

To illustrate the effect of the constraint, we repeated the experiment shown in Figure 1 with \( D \) constrained to be a
function case, we use Lagrange multipliers. Thus, we minimize the cost

\[ \mathbf{D} = \text{banded matrix} \]

for different numbers of signals, when \( \mathbf{D} \) is constrained to be a banded matrix. All parameters are the same as in Figure 1.

A banded matrix with 5 nonzero diagonals (the main diagonal and the two adjacent diagonals on each side). The result is shown in Figure 11. Note that \( \mathbf{D} \) can now be estimated with fewer periods than in the unconstrained case.

A.2. General Linear Constraints \( \epsilon \neq 0 \). To solve the general case, we use Lagrange multipliers. Thus, we minimize the cost function

\[ f(\mathbf{d}) = \mathbf{d}^H \mathbf{Qd} - \lambda^T (\mathbf{Bd} - \mathbf{c}) - \beta (\mathbf{d}^H \mathbf{d} - 1), \]  

where \( \lambda \) is an \( r \times 1 \) and \( \beta \) is a scalar. The solution is found by setting the derivative of the cost function to zero,

\[ \frac{\partial f}{\partial \mathbf{d}} = (\mathbf{Q} - \beta \mathbf{I}) \mathbf{d} - \mathbf{B}^H \lambda = 0, \]

so

\[ \mathbf{d} = (\mathbf{Q} - \beta \mathbf{I})^{-1} \mathbf{B}^H \lambda. \]

Inserting the solution back into the linear constraints, we get

\[ \mathbf{Bd} = \mathbf{B}(\mathbf{Q} - \beta \mathbf{I})^{-1} \mathbf{B}^H \lambda = \mathbf{c}, \]

and thus

\[ \lambda = (\mathbf{B}(\mathbf{Q} - \beta \mathbf{I})^{-1} \mathbf{B}^H)^{-1} \mathbf{c}. \]

From the quadratic constraint, we get

\[ \mathbf{d}^H \mathbf{d} = \lambda^H \mathbf{B}(\mathbf{Q} - \beta \mathbf{I})^{-2} \mathbf{B}^H \lambda \]  

or

\[ \mathbf{d}^H \mathbf{d} = \mathbf{c}^H ((\mathbf{B}(\mathbf{Q} - \beta \mathbf{I})^{-1} \mathbf{B}^H)^{-1} \mathbf{B}(\mathbf{Q} - \beta \mathbf{I})^{-1} \mathbf{B}^H)^{-1} \mathbf{c} = 1. \]

There does not seem to be a closed-form solution to this nonlinear equation in \( \beta \). However, since \( \beta \) is a scalar quantity, it can be easily found by straightforward search. The minimizing value of \( \beta \) is then inserted into the equation for \( \mathbf{d} \)

\[ \mathbf{d} = (\mathbf{Q} - \beta \mathbf{I})^{-1} \mathbf{B}^H (\mathbf{B}(\mathbf{Q} - \beta \mathbf{I})^{-1} \mathbf{B}^H)^{-1} \mathbf{c}, \]

to provide the final constrained solution.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

References


