Real-Time Frequency and 2-D Angle Estimation with Sub-Nyquist Spatio-Temporal Sampling

Michael D. Zoltowski, *Member, IEEE*, and Cherian P. Mathews, *Member, IEEE*

Abstract—**An algorithm is presented for real-time estimation of the frequency and azimuth and elevation angles of each signal incident on an airborne antenna array system over a very wide frequency band (2–18 GHz) commensurate with electronic signal warfare. The algorithm provides unambiguous frequency estimation despite severe temporal undersampling necessitated by cost/complexity of hardware considerations. The 2–18 GHz spectrum is decomposed into 1-GHz bands. The baseband output of each antenna is sent through two 250-MHz sampled channels where one is delayed relative to the other (prior to sampling) by 0.5 ns, which is the Nyquist interval for a 1-GHz bandwidth. Due to the high variance of the Direct ESPRIT frequency estimator, aliased frequencies are estimated via a simple formula and translated to the proper aliasing zone, utilizing eigenvector information generated by PRO-ESPRIT. The algorithm also provides unambigous 2-D angle estimate over the entire 2–18 GHz bandwidth, despite severe spatial undersampling at the higher end of this band necessitated by mutual coupling considerations and resolving power requirements at the lower end of the band. Eigenvector information generated by PRO-ESPRIT is used to facilitate computationally simple estimation of azimuth and elevation angles that are automatically paired with corresponding frequency estimates despite aliasing. Simulations are presented demonstrating the capabilities of the algorithm.**

I. Introduction

THE problem under investigation is that of real-time estimation of the frequency and azimuth and elevation angles of each signal incident on an airborne antenna array system over a very wide frequency band (2–18 GHz) commensurate with electronic warfare. The problem is complicated by severe undersampling in both the temporal and spatial domains necessitated by cost and complexity of hardware considerations [1].

To reduce the complexity of the overall receiver hardware, the bandwidth at the intermediate frequency is chosen to be quite large equal to 1 GHz. Correspondingly, the entire 2–18 GHz spectrum is decomposed into overlapping 1 GHz bands; each band is examined in succession or in parallel. The Nyquist temporal sampling rate for digitization of a 1-GHz band is 2 GHz. Although A/D converters operating at 2-GHz rate are available, they are very expensive, and processing speed following the converter may limit the overall operation of the receiver. In the prototype system pictured in Fig. 1 [1], the receiver output, after conversion to baseband, is sampled at a rate of 250 MHz, which is one eighth of the Nyquist rate. This severe undersampling leads to aliasing and attendant problems of ambiguity. The aliased frequency as a function of baseband frequency with a sampling rate of 250 MHz is plotted in Fig. 3.

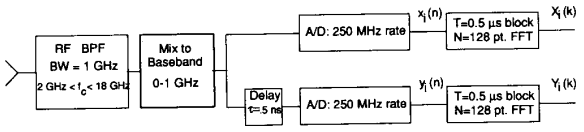


Fig. 1. Receiver module and front-end signal processing for ith antenna in the prototype system.

Note that the aliasing function plotted in Fig. 3 is for the case where only the in-phase channel is sampled. Sampling of the quadrature channel represents additional hardware costs and, overall, doubles the number of samples to be processed. Thus, in keeping with the overall goal of reduced complexity of hardware and computation, it is assumed that only the in-phase component, which is a real-valued signal, is sampled and input to the system. Note that it is typically necessary to generate the complex analytic signal in a direction finding application to resolve a 180° ambiguity in the azimuth angle estimates. Again, motivated by the desire to keep the computational complexity low, the complex analytic signal is roughly approximated by computing the DFT of the output of each antenna and throwing away the negative frequency portion of the spectrum. This approach averts the need to pass the sampled signal through an FIR digital Hilbert transformer, which could possibly lead to edge effects or a reduced number of effective time samples (depending on whether one includes all output points of the FIR digital Hilbert transformer or just those output points for which there were no zero entries in the FIR filter window.) The spatio-temporal signal model is developed in Section II.[1]

In order to estimate the base-band frequency of each signal despite aliasing, the base-band output of each antenna is sent through two 250-MHz sampled channels, where one is delayed by τ relative to the other (prior to sampling) as indicated in Fig. 1. The time-delay τ is chosen less than or equal to the Nyquist sampling interval for the baseband bandwidth W,

Manuscript received June 4, 1993; revised March 28, 1994. This work was supported by AFOSR under contract no. F49620-92-J-0198 in conjunction with Wright Laboratories. The assosciate editor coordinating the review of this paper and approving it for publication was Prof. Daniel Fuhrmann.

The authors are with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907-1285 USA

IEEE Log Number 9403751.

[1]The procedure for frequency estimation with subNyquist temporal sampling developed within may be easily adopted for narrow-band direction-of-arrival estimation with two identical, collinear uniform linear arrays (ULA's). In this application, the displacement between the two arrays should be less than a half wavelength, but the interelement spacing for either array may be much greater than a half wavelength to achieve a large aperture and, hence, increased resolution capability relative to a ULA of the same total number of elements but with half-wavelength spacing.

i.e., $\tau \leq 1/(2W)$. In the prototype system depicted in Fig. 1, $W = 1$ GHz, and $\tau = .5$ ns $= .5 \times 10^{-9}$ s. ESPRIT [2], [3] may then be applied to estimate the baseband frequencies in any 1-GHz base-band bandwidth. To facilitate real-time implementation, ESPRIT is applied in DFT space. In this mode of processing the steps are

i) Compute an FFT of a block of samples.
ii) Locate peaks via a simple peak-picking algorithm.
iii) Apply ESPRIT to a small set of DFT values around each peak.

In Section III, we show that the *Direct* ESPRIT frequency estimator has a variance several orders of magnitude greater than the Cramer-Rao lower bound (CRB). An alternative approach referred to as *Indirect ESPRIT* is presented that is computational simple and achieves performance very close to the CRB. Indirect ESPRIT makes novel use of eigenvector information generated by the PRO-ESPRIT algorithm [3] to estimate the aliased frequency of each source via a simple formula and correctly translate it to the proper aliasing zone, where it is added to or subtracted from the appropriate integer of the sampling rate in accordance with Fig. 3.

Once the frequency of each signal is estimated, the next goal is to estimate the corresponding azimuth and elevation angles. There are two problems here. First, each angle estimate must be correctly paired with the proper frequency estimate. Second, in general, 2-D angle estimation is significantly more computationally complex than 1-D angle estimation. Again, real-time implementation is an overriding factor. Now, since the sources are at different frequencies, the filtering inherent in selecting only those DFT values around a spectral peak should ideally be sufficient to isolate single source contributions and avoid the frequency-angle pairing problem. However, aside from sidelobe leakage effects, this is not the case as sources well separated in analog frequency may be aliased to very nearly the same digital frequency. In Section IV, eigenvector information generated by PRO-ESPRIT is used to facilitate computationally simple estimation of azimuth and elevation angles automatically paired with corresponding frequency estimates despite aliasing.

In the case of a uniformly spaced linear array, half-wavelength spacing between antennas is required to avoid ambiguities in estimating the arrival angle of a signal. With half-wavelength spacing at the upper end of the 2–18 GHz spectrum, the elements are too closely spaced at the lower end of the spectrum, leading to problems of mutal coupling and poor resolution. The resolution capability and estimator accuracy of any arrival angle estimation algorithm is proportional to the aperture length measured in units of wavelengths. To achieve a high degree of resolution power and estimator accuracy and yet avoid mutual coupling, the elements must be spaced nonuniformly over a large aperture.

The prototype system employs an L-shaped antenna array having nonuniformly spaced elements along each leg as pictured in Fig. 2. The interelement spacings along either axis is much greater than a half wavelength, particularly at 18 GHz. In Section IV, we develop i) a prescription for interelement spacings for nonambiguous angle estimation and

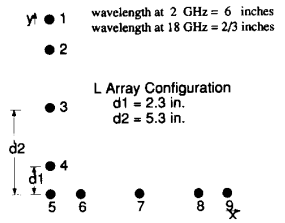


Fig. 2. L-shaped antenna array employed in simulations for azimuth/elevation angle estimation over 2–18 GHz spectrum. Each leg is symmetric about its center.

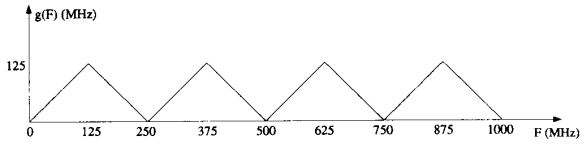


Fig. 3. Aliasing function: Aliased frequency as a function of baseband frequency after sampling at 250 MHz with real processing (no I and Q).

ii) an attendant algorithm for angle estimation that is computationally simple for real-time implementation. Although there is a plethora of previous work on the design of nonuniform linear arrays [6]–[8], the development in Section IV assumes a small number of antenna elements due to cost and space limitations on the antenna platform attached to the aircraft. In addition, high sidelobes is not as much a problem since we are able to isolate the individual contribution of each source. In contrast to previous work [6]–[8], the prescription for interelement spacings is developed synergistically with a simple integer-based search algorithm for angle estimation. Secton V presents simulations that demonstrate the power of the overall frequency and 2-D angle estimation algorithm summarized in the flowchart presented in Fig. 5.

II. Spatio-Temporal Sampling and Data Model

The parameters for the prototype subNyquist spatio-temporal sampling system are indicated in Fig. 1. We concentrate on signal parameter estimation for a particular 1-GHz base-band bandwidth. For the sake of simplicity, the signals are modeled as RF pulsed waveforms. The development to follow, though, holds as long as each signal satisfies the standard narrowband assumption $\frac{B}{f_c}\frac{L}{\lambda}\cos\theta \ll 1$. For a given signal

B bandwidth
f_c carrier frequency
L length of the array
λ wavelength
$\cos\theta$ direction cosine relative to the array axis.

Since the carrier frequencies here lie somewhere between 2 and 18 GHz, the narrowband assumption is satisfied almost always except for some extremely wide-band signals. We also assume that no two signals are at exactly the same RF frequency. Even if there is multipath propagation between a given source and the airborne antenna array, the Doppler shift each multipath signal undergoes is distinct as long as each

multipath signal has its own distinct azimuth and elevation coordinates [9].

Let the sampling rate be denoted F_s. We are here assuming that F_s is well below the Nyquist rate leading to aliasing. For our prototype system, $F_s = 250$ MHz is equal to one eighth of the Nyquist rate (2 GHz for a 1-GHz baseband bandwidth). Consider sampling a single sinusoid of the form $\cos(2\pi F_j t + \phi)$, where F_j is the baseband frequency ($0 \leq F_j \leq 1$ GHz).

$$\begin{aligned}
&\cos(2\pi F_j t + \phi)|_{t=n/F_s} \\
&\quad = \cos\left(2\pi \frac{F_j}{F_s} n + \phi\right) \\
&\quad = \cos[2\pi f_j n + \phi] \qquad \text{for } 0 < F_j < \frac{F_s}{2} \\
&\quad = \cos\left[2\pi\left(\frac{F_j}{F_s} - 1\right)n + \phi\right] \\
&\quad = \cos\left[2\pi\left(1 - \frac{F_j}{F_s}\right)n - \phi\right] \\
&\quad = \cos[2\pi f_j n - \phi] \qquad \text{for } \frac{F_s}{2} < F_j < F_s \\
&\quad = \cos\left[2\pi\left(1 - \frac{F_j}{F_s}\right)n + \phi\right] \\
&\quad = \cos[2\pi f_j n + \phi] \qquad \text{for } F_s < F_j < 3\frac{F_s}{2} \\
&\quad = \cos\left[2\pi\left(\frac{F_j}{F_s} - 2\right)n + \phi\right] \\
&\quad = \cos\left[2\pi\left(2 - \frac{F_j}{F_s}\right)n - \phi\right] \\
&\quad = \cos[2\pi f_j n - \phi] \qquad \text{for } 3\frac{F_s}{2} < F_j < 2F_s \\
&\qquad \vdots
\end{aligned} \tag{1}$$

For each range of the analog baseband frequency, the corresponding digital frequency f_j is between 0 and 0.5, i.e., $0 < f_j < .5$. Continuing this development, we obtain the aliasing function $g(F)$ plotted in Fig. 3 for the case of $F_s = 250$ MHz corresponding to our prototype system. With the aliasing function thus defined, the digital frequency f_j is related to analog baseband frequency F_j as $f_j = g(F_j)/F_s$. The analog aliased frequency is defined as $F_j^a = f_j F_s = g(F); F_j^a$ is the frequency one would obtain if the analog sinusoidal signal was reconstructed from its samples. An important observation in (1) is that when F_j is in a range where the slope of the aliasing function $g(F)$ is negative, the constant phase offset of the sampled sinusoid is the negative of that associated with the continuous-time sinusoid.

In order to estimate the base-band frequency of each signal despite aliasing, the base-band output of each antenna is sent through two 250-MHz sampled channels where one is delayed by τ relative to the other (prior to sampling). We here assume that the time delay τ is less than or equal to the Nyquist sampling interval for the baseband bandwidth, W, i.e., $\tau \leq 1/(2W)$. In the prototype system depicted in Fig. 1, $W = 1$ GHz and $\tau = 0.5$ ns $= 0.5 \times 10^{-9}$ s.

The sampled versions of the reference and time-delayed data sets, which are referred to as the X and Y data sets, respectively (one pair of data sets for each antenna), may be described as

$$\begin{aligned}
x_i(n) &= \sum_{j=1}^{J} \left\{ \frac{A_j}{2} e^{j\kappa_j \gamma_{j^o}} e^{j\kappa_j \gamma_j(i)} e^{j2\pi f_j n} \right. \\
&\qquad \left. + \frac{A_j}{2} e^{-j\kappa_j \gamma_{j^o}} e^{-j\kappa_j \gamma_j(i)} e^{-j2\pi f_j n} \right\} \qquad (2) \\
y_i(n) &= \sum_{j=1}^{J} \left\{ \frac{A_j}{2} e^{j\kappa_j \gamma_{j^o}} e^{j\kappa_j \gamma_j(i)} e^{-j\kappa_j 2\pi F_j \tau} e^{j2\pi f_j n} \right. \\
&\qquad \left. + \frac{A_j}{2} e^{-j\kappa_j \gamma_{j^o}} e^{-j\kappa_j \gamma_j(i)} e^{j\kappa_j 2\pi F_j \tau} e^{-j2\pi f_j n} \right\}
\end{aligned}$$

where, for the moment, we are neglecting the effects of noise. The various quantities in (2) are described below. J is the total number of signals in a particular 1-GHz base-band bandwidth. A_j is the amplitude of the jth signal, whereas γ_{jo} is the phase of jth signal at the origin of the antenna array system. $\gamma_j(i)$ is the relative phase of the jth signal arrival at the ith antenna. If the ith antenna is at the x-y coordinate pair (x_i, y_i) and the jth source is at an azimuth angle of θ_j and an elevation angle of ϕ_j

$$\begin{aligned}
\gamma_j(i) &= -\frac{2\pi}{\lambda_j}(x_i \cos\theta_j \sin\phi_j + y_i \sin\theta_j \sin\phi_j) \\
&\quad i = 1, \ldots, M
\end{aligned} \tag{3}$$

where λ_j is the wavelength of the jth signal arrival, and M is the total number of antennas comprising the array. κ_j is the slope of the aliasing function $g(F)$ at $F = F_j$ equal to either $+1$ or -1. In accordance with (1), κ_j takes into account the conjugation that occurs when F_j is in an interval where the aliasing function is downward sloping. Note that in the prototype system, the observation interval is 0.5 μs $= 0.5 \times 10^{-6}$s, yielding roughly $N = 128$ samples for each of the M antennas.

As indicated in Fig. 1, the first processing step is to compute an FFT of both the X and Y data sets at each antenna output. Ultimately, ESPRIT [2], [3] is applied to a small set of DFT values around each spectral peak in the positive frequency portion of each of the 2M spectra. We are effectively using the DFT as a narrow-band pass-band filter. This is done for two reasons. First, by isolating only positive frequencies, we are able to resolve a 180° ambiguity in azimuth angle. Second, in processing a given peak, the eigenvalue decompositions (EVD's) required are done on matrices of dimension equal to the number of DFT values that are less than the number of antennas. Separate peaks may be processed in parallel. Recall that sources well separated in base-band frequency may be aliased to very nearly the same digital frequency due to undersampling. Thus, several sources may be contributing to a given spectral peak.

The respective N-point DFT's of the X and Y data sets for the ith antenna are denoted $X_i(k)$ and $Y_i(k)$, $i = 1, \ldots, M$,

and may be expressed as

$$X_i(k) = \sum_{j=1}^{J} \left\{ \frac{A_j}{2} e^{j\kappa_j\gamma_{j_o}} e^{j\kappa_j\gamma_j(i)} \text{sinc}_{N_j}\left(f_j - \frac{k}{N}\right) + \frac{A_j}{2} e^{-j\kappa_j\gamma_{j_o}} e^{-j\kappa_j\gamma_j(i)} \text{sinc}_{N_j}\left(f_j + \frac{k}{N}\right) \right\}$$

$$Y_i(k) = \sum_{j=1}^{J} \left\{ \frac{A_j}{2} e^{j\kappa_j\gamma_{j_o}} e^{j\kappa_j\gamma_j(i)} e^{-j\kappa_j 2\pi F_j\tau} \text{sinc}_{N_j}\left(f_j - \frac{k}{N}\right) + \frac{A_j}{2} e^{-j\kappa_j\gamma_{j_o}} e^{-j\kappa_j\gamma_j(i)} e^{j\kappa_j 2\pi F_j\tau} \text{sinc}_{N_j}\left(f_j + \frac{k}{N}\right) \right\} \quad (4)$$

where N_j is the number of samples for which the jth signal is "turned on," and the periodic sinc function is defined as $\text{sinc}_N(f) = e^{-j\pi(N-1)f}\frac{\sin(N\pi f)}{\sin(\pi f)}$. Note, in contrast to convention, that we include the phase term $e^{-j\pi(N-1)f}$ in the definition of $\text{sinc}_N(f)$ for the sake of notational simplicity.

The next processing step is to locate spectral peaks. We here assume that a simple peak-picking algorithm is employed. Note that only coarse estimates of the peak locations are required for the algorithm to perform well. The respective DFT spectra for the X and Y data set for each antenna, $2M$ DFT spectra all together, should exhibit peaks at the same locations. At this point, we concentrate on a single peak in each DFT spectrum at the same location located at or near the DFT value $k = k_o$ without loss of generality. $L = 2L' + 1$ DFT values around the corresponding peak in each DFT spectrum are collected to construct the following set of $2M$ $L \times 1$ vectors:

$$\mathbf{X}_i(k_o) = [X_i(k_o - L'), \ldots, X_i(k_o), \ldots, X_i(k_o + L')]^T$$
$$\mathbf{Y}_i(k_o) = [Y_i(k_o - L'), \ldots, Y_i(k_o), \ldots, Y_i(k_o + L')]^T.$$

To give a perspective on the computational complexity, in the simulations presented in Section VI, we ran cases where $\mathbf{X}_i(k_o)$ and $\mathbf{Y}_i(k_o)$ are 4×1 and cases where $\mathbf{X}_i(k_o)$ and $\mathbf{Y}_i(k_o)$ are 5×1. The governing factor is that the number of DFT values selected around a peak should be at least one greater than the number of sources making significant contributions to that peak, which is denoted J'.

Substituting (4) into (5), the $L \times 1$ vector of X DFT values around k_o may be expressed as

$$\mathbf{X}_i(k_o) = \sum_{j=1}^{J} \left\{ \frac{A_j}{2} e^{j\kappa_j\gamma_{j_o}} e^{j\kappa_j\gamma_j(i)} \mathbf{d}(f_j) + \frac{A_j}{2} e^{-j\kappa_j\gamma_{j_o}} e^{-j\kappa_j\gamma_j(i)} \mathbf{d}(-f_j) \right\} \quad (6)$$

where $\mathbf{d}(f_j)$ is the $L \times 1$ vector

$$\mathbf{d}(f_j) = \left[\text{sinc}_{N_j}\left[f_j - \frac{k_o - L'}{N}\right], \ldots, \text{sinc}_{N_j}\left[f_j - \frac{k_o}{N}\right], \ldots, \text{sinc}_{N_j}\left[f_j - \frac{k_o + L'}{N}\right] \right]^T. \quad (7)$$

As long as the window of DFT values is not either near $k = 0$ or near $k = N/2$, the DFT acts as a narrow-band band-pass filter such that $\mathbf{d}(-f_j)$ is small enough relative to $\mathbf{d}(f_j)$ to be negligible. To simplify the development, we will neglect the contribution of $\mathbf{d}(-f_j)$. If $\mathbf{d}(-f_j)$ is not negligible, then the algorithm to be developed will indicate a source having a negative aliased frequency that potentially may be screened out.

Neglecting the negative frequency contributions, $\mathbf{X}_i(k_o) = \sum_{j=1}^{J'} \frac{A_j}{2} e^{j\kappa_j\gamma_{j_o}} e^{j\kappa_j\gamma_j(i)} \mathbf{d}(f_j)$, where $J' \leq J$ is the number of sources making a significant contribution to the spectral peak at or near the digital frequency k_o/N. This expression describes the vector of DFT values around a peak in the DFT spectrum of a single antenna. The DFT vectors from all M antennas are placed as the columns of an $L \times M$ matrix as $\mathbf{X} = [\mathbf{X}_1(k_o) \vdots \mathbf{X}_2(k_o) \vdots \cdots \vdots \mathbf{X}_M(k_o)]$. $\mathbf{X}$ may be expressed in factored form as

$$\mathbf{X} = \sum_{j=1}^{J'} A'_j e^{j\kappa_j\gamma_{j_o}} \mathbf{d}(f_j)\mathbf{a}^T(\theta_j, \phi_j, \kappa_j) \qquad (L \times M) \quad (8)$$

where $A'_j = A_j/2$ and $\mathbf{a}(\theta_j, \phi_j, \kappa_j) = [e^{j\kappa_j\gamma_j(1)}, e^{j\kappa_j\gamma_j(2)}, \ldots, e^{j\kappa_j\gamma_j(M)}]^T$ with $\gamma_j(i)$ defined by (3). $\mathbf{a}(\theta_j, \phi_j, \kappa_j)$ for $\kappa_j = 1$ is the $M \times 1$ array manifold vector for a signal incident from the (θ_j, ϕ_j) direction. The dependence on κ_j, which is the slope of the aliasing function at F_j, is introduced as a simplistic means of denoting a conjugation; it allows us to avoid breaking the sum in (8) into terms for which the array manifold is conjugated and those for which it is not conjugated.

Similarly, the corresponding DFT outputs from all M antennas for the Y (time-delayed) data is collected as $\mathbf{Y} = [\mathbf{Y}_1(k_o) \vdots \mathbf{Y}_2(k_o) \vdots \cdots \vdots \mathbf{Y}_M(k_o)]$. Neglecting negative frequency components, $\mathbf{Y}$ may be expressed as

$$\mathbf{Y} = \sum_{j=1}^{J'} A'_j e^{j\kappa_j\gamma_{j_o}} e^{-j\kappa_j 2\pi F_j\tau} \mathbf{d}(f_j)\mathbf{a}^T(\theta_j, \phi_j, \kappa_j) \qquad (L \times M). \quad (9)$$

Equations (8) and (9) represent the pure signal component of the spatio-temporal data model assumed throughout. Again, κ_j is the slope of the aliasing function $g(F)$ in Fig. 3 at $F = F_j$ equal to either $+1$ or -1. κ_j is a notational tool that takes into account the conjugation that occurs when F_j is in an interval where the aliasing function is downward sloping.

III. ESPRIT-Based Frequency Estimation with Temporal Undersampling

Given the data model described by (8) and (9), the applicability of ESPRIT [2], [3] is evident.

$$\mathbf{Y} = \mu\mathbf{X} = \sum_{j=1}^{J'} A'_j e^{j\kappa_j\gamma_{j_o}} \{e^{-j\kappa_j 2\pi F_j\tau} - \mu\} \mathbf{d}(f_j)\mathbf{a}^T(\theta_j, \phi_j, \kappa_j). \quad (10)$$

The critical observation for estimating F_j is that when $\mu = e^{-j\kappa_l 2\pi F_l\tau}$, the rank of $\mathbf{Y} - \mu\mathbf{X}$ drops from J' to $J' - 1$ since the lth term drops out of the sum. Thus, $\mu_j = e^{-j\kappa_j 2\pi F_j\tau}, j =$

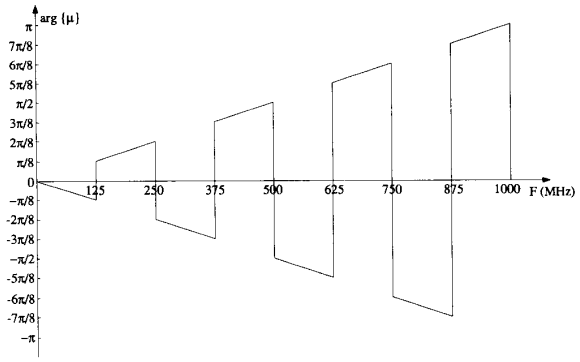


Fig. 4. Phase of ESPRIT eigenvalue as a function of analog base-band frequency with 250-MHz sampling rate.

$1, \ldots, J'$ are J' distinct generalized eigenvalues of the $L \times M$ rectangular matrix pencil $\{\mathbf{Y}, \mathbf{X}\}$.

The argument of the ESPRIT eigenvalue $\arg\{\mu_j\} = -\kappa_j 2\pi F_j \tau$ is plotted as a function of the base-band frequency F_j for $F_s = 250$ MHz in Fig. 4. Recall that κ_j is the slope of the aliasing function at $F = F_j$. Note that certain ranges of phase within $(-\pi, \pi)$ are not permissible as the argument of μ_j. In fact, only half of the 2π interval $(-\pi, \pi)$ is permissible. For example, under ideal noiseless conditions, no value of phase in the interval $(-\pi/8, -2\pi/8)$ is permissible as the argument of $\mu_j = e^{-j\kappa_j 2\pi F_j \tau}$.

The PRO-ESPRIT [3] variant of ESPRIT is here employed as a "fast" implementaton of ESPRIT for estimating the phase factors $-\kappa_j 2\pi F_j \tau, j = 1, \ldots, J'$. PRO-ESPRIT operates on the $L \times L$ autocorrelation and cross-correlation matrices $\mathbf{R}_{xx} = \frac{1}{M}\sum_{i=1}^{M} \mathbf{X}_i(k_o)\mathbf{X}_i^H(k_o) = \frac{1}{M}\mathbf{X}\mathbf{X}^H$ and $\mathbf{R}_{yx} = \frac{1}{M}\sum_{i=1}^{M} \mathbf{Y}_i(k_o)\mathbf{X}_i^H(k_o) = \frac{1}{M}\mathbf{Y}\mathbf{X}^H$. Note that the number of DFT values selected around the peak at k_o (L) may be as small as two if only a single source is contributing to the peak. The algorithm is first summarized and then briefly justified.

First, compute an EVD of $\mathbf{R}_{XX}$:$\mathbf{R}_{XX}\mathbf{u}_i = \lambda_i \mathbf{u}_i, i = 1, \ldots, L$, where the eigenvalues are indexed in order of decreasing magnitude. The number of complex sinusoids with aliased frequency components in the vicinity of k_o, $J' \leq J$ may be determined from a number of techniques including statistical tests that examine the eigenvalues such as AIC or MDL. With the $J' < L$ largest (signal) eigenvalues and corresponding signal eigenvectors, construct the $J' \times J'$ diagonal matrix Σ_S and the $L \times J'$ matrix $\boldsymbol{U}_S$ as

$$\Sigma_S = \operatorname{diag}\{(\lambda_1 - \lambda_{\min})^{1/2}, (\lambda_2 - \lambda_{\min})^{1/2}, \ldots, (\lambda_{J'} - \lambda_{\min})^{1/2}\} \tag{11}$$

$$\mathbf{U}_S = [\mathbf{u}_1 \vdots \mathbf{u}_2 \vdots \cdots \vdots \mathbf{u}_{J'}]. \tag{12}$$

The smallest eigenvalue $\lambda_{\min}$ is asymptotically equal to the noise power that affects the diagonal elements of the autocorrelation matrix $\mathbf{R}_{xx}$. Note that for a given antenna output, even if the noise is not white, i.e., the noise spectral density is not flat over the entire 1-GHz bandwidth, it can be shown that the noise contaminating a small set of successive DFT values is approximately i.i.d. The final major step is an EVD of

$$\mathbf{\Psi} = \Sigma_S^{-1}\mathbf{U}_S^H \mathbf{R}_{yx} \mathbf{U}_S \Sigma_S^{-1} \qquad (J' \times J'). \tag{13}$$

The eigenvalues of $\mathbf{\Psi}$ are estimates of $\mu_j = e^{-j\kappa_j 2\pi F_j \tau}, j = 1, \ldots, J'$.

Proof: Let $\mathbf{X} = \mathbf{U}_S \Sigma_S \mathbf{V}_S^H$ be the SVD of $\mathbf{X}$, including only the J' nonzero values and corresponding left and right singular vectors; $\mathbf{U}_S$ is $L \times J'$, Σ_S is $J' \times J'$, and $\mathbf{V}_S$ is $M \times J'$. It follows from (8) and (9) that $\text{range}\{\mathbf{U}_S\} = \text{range}\{\mathbf{Y}\} = \text{span}\{\mathbf{d}(f_1), \ldots, \mathbf{d}(f_{J'})\}$ and $\text{range}\{\mathbf{V}_S\} = \text{range}\{\mathbf{Y}^H\} = \text{span}\{\mathbf{a}^*(\theta_1, \phi_1, \kappa_1), \ldots, \mathbf{a}^*(\theta_{J'}, \phi_{J'}, \kappa_{J'})\}$ such that $\mathbf{Y} = \mathbf{U}_S\mathbf{U}_S^H\mathbf{Y}\mathbf{V}_S\mathbf{V}_S^H$, where $\mathbf{U}_S\mathbf{U}_S^H$ and $\mathbf{V}_S\mathbf{V}_S^H$ are projection operators. Thus

$$\begin{aligned}\mathbf{Y} - \mu\mathbf{X} &= \mathbf{U}_S\mathbf{U}_S^H\mathbf{Y}\,\mathbf{V}_S^H - \mu\mathbf{U}_S\Sigma_S\mathbf{V}_S^H \\ &= \mathbf{U}_S\Sigma_S\{\Sigma_S^{-1}\mathbf{U}_S^H\mathbf{Y}\mathbf{V}_S - \mu\mathbf{I}_{J'}\}\mathbf{V}_S^H \\ &= \mathbf{U}_S\Sigma_S\{\Sigma_S^{-1}\mathbf{U}_S^H\mathbf{Y}\mathbf{V}_S\Sigma_S\mathbf{U}_S^H\mathbf{U}_S\Sigma_S^{-1} - \mu\mathbf{I}_{J'}\}\mathbf{V}_S^H \\ &= \mathbf{U}_S\Sigma_S\{\Sigma_S^{-1}\mathbf{U}_S^H\mathbf{Y}\mathbf{X}^H\mathbf{U}_S\Sigma_S^{-1} - \mu\mathbf{I}_{J'}\}\mathbf{V}_S^H \end{aligned} \tag{14}$$

where we have used the fact that $\Sigma_S\mathbf{U}_S^H\mathbf{U}_S\Sigma_S^{-1} = I_j$. Thus, the J' nontrivial generalized eigenvalues of the $L \times M$ matrix pencil $\{\mathbf{Y}, \mathbf{X}\}$ may be computed as the eigenvalues of the $J' \times J'$ matrix $\Sigma_S^{-1}\mathbf{U}_S^H\mathbf{Y}\mathbf{X}^H\mathbf{U}_S\Sigma_S^{-1}$. The proof is completed by recognizing that $\mathbf{X}\mathbf{X}^H = \mathbf{U}_S\Sigma_S^2\mathbf{U}_S^H$.

In the prototype system $\tau = 0.5 \times 10^{-9}$ s such that F_j may be estimated from the phase of the jth ESPRIT eigenvalue according to $F_j = |\arg\{\mu_j\}/2\pi|(2 \times 10^9)$ Hz, $j = 1, \ldots, J'$, where $\arg\{z\}$ is the phase angle of the complex number z. Any error in $\arg\{\mu_j\}$ due to noise is grossly magnified due to the multiplication by 10^9, i.e., multiplication by 1 GHz. Simulations presented in Section VI reveal that the variance of the base-band frequency estimates obtained from ESPRIT in this manner are on the order of 10 MHz, whereas the CRB on the variance of any unbiased estimator of frequency is on the order of 10 KHz. This extreme differential motivates us to see if we can obtain a performance closer to the CRB without incurring too much additional computation.

The above approach is referred to as the *Direct ESPRIT* approach. An alternative approach is referred to as *Indirect ESPRIT*. The steps in Indirect ESPRIT are as follows:

i) Estimate the digital frequency f_j.
ii) Convert f_j to the aliased analog frequency via $F_j^a = f_j F_s$.
iii) Translate F_j^a up to the proper aliasing zone using the phase of the ESPRIT eigenvalue μ_j, in conjunction with Fig. 4, where F_j^a is either added to or subtracted from an integer multiple of the sampling rate to estimate the actual base-band frequency.

Two computationally efficient, high-resolution algorithms for estimating the aliased frequencies using DFT values as input are Beamspace Root-MUSIC [4] and Beamspace ESPRIT [5]. Recall that high-resolution capability is necessary since sources well separated in analog frequency may be very closely spaced in digital frequency due to aliasing. However, despite their relative computational efficiency, implementing

either of these two algorithms represents a substantial increase in computational complexity.

More important, though, is the data association problem wherein the aliased frequency estimates must be paired with the correct ESPRIT eigenvalue so that it is translated to the proper alias zone. If the aliased frequencies are estimated independently of the ESPRIT eigenvalues, this pairing problem is very difficult and insurmountable when sources are closely spaced in frequency after aliasing. Fortuitously, eigenvector information provided by PRO-ESPRIT facilitates automatic pairing of the aliased frequency estimates with the corresponding ESPRIT eigenvalues. In addition, the eigenvector information generated by PRO-ESPRIT provides a means for isolating the individual contribution of each source despite aliasing. This facilitates simple estimation of the aliased frequency associated with each source. It may be done on an individual basis, assuming a single source leading to a simple closed-form formula as shown shortly.

The jth $M \times 1$ right generalized eigenvector r_j of the $L \times M$ rectangular matrix pencil $\{\mathbf{Y}, \mathbf{X}\}$ is that vector satisfying $\{\mathbf{Y} - \mu_j \mathbf{X}\}\mathbf{r}_j = \mathbf{0}$. Substituting the noiseless (ideal) forms of the X and Y data matrices:

$$\left\{\sum_{j=1}^{J'} A'_j e^{j\kappa_j \gamma_{j_o}} \{e^{-j\kappa_j 2\pi F_j \tau} - \mu\} \mathbf{d}(f_j)\mathbf{a}^T(\theta_j, \phi_j, \kappa_j)\right\}\mathbf{r}_j = \mathbf{0}. \tag{15}$$

When $\mu = e^{-j\kappa_l 2\pi F_l \tau}$, the lth term $\mathbf{d}(f_l)\mathbf{a}^T(\theta_l, \phi_l, \kappa_l)$ drops out of sum such that $\mathbf{a}^T(\theta_l, \phi_l, \kappa_l)\mathbf{r}_j = 0$ for $l = 1, \ldots, J', l \neq j$. Hence, $\mathbf{r}_j$ can be used to extract $\mathbf{d}(f_j)$ to within a scalar multiple:

$$\mathbf{X}\mathbf{r}_j \propto \mathbf{d}(f_j) \qquad (\mathbf{Y}\mathbf{r}_j \propto \mathbf{d}(f_j)).$$

A key point is that the estimate of $\mathbf{d}(f_j)$ obtained in this manner is automatically paired with the ESPRIT eigenvalue that is an estimate of $\mu_j = e^{-j\kappa_j 2\pi F_j \tau}$ since $\mathbf{r}_j$ is the right generalized eigenvector associated with μ_j. Thus, a frequency estimation algorithm that assumes a single source may be applied to $\hat{\mathbf{d}}(f_j)$ to estimate f_j.

Note that we only desire $\mathbf{r}_j$ in order to compute $\mathbf{X}\mathbf{r}_j$ as our estimate of $\mathbf{d}(f_j)$ to within a scalar multiple. We can bypass the computation of $\mathbf{r}_j$ and construct $\mathbf{X}\mathbf{r}_j$ directly from the $\mathbf{J}' \times 1$ right eigenvectors of $\mathbf{\Psi}$, which is defined by (13), satisfying $\mathbf{\Psi}\beta_j = \mu_j \beta_j, \quad j = 1, \ldots, J'$. From (14) and (15), it follows that

$$\hat{\mathbf{d}}(f_j) \quad \propto \quad \mathbf{X}\mathbf{r}_j = \mathbf{U}_S \Sigma_S \beta_j \qquad j = 1, \ldots, J' \tag{16}$$

where Σ_S and $\mathbf{U}_S$ are constructed from the J' largest eigenvalues and corresponding eigenvectors of $\mathbf{R}_{xx}$ according to (11) and (12), respectively.

Next, we apply Beamspace ESPRIT [5] to $\hat{\mathbf{d}}(f_j)$ to estimate f_j. After much algebraic manipulation, the single source assumption leads to the following simple formula for estimating f_j:

$$\hat{f}_j = \frac{1}{2\pi}\arg\left\{\hat{\mathbf{d}}^H(f_j)\Delta^* \mathbf{P}_1^{\perp} \hat{\mathbf{d}}(f_j)\right\} \quad j = 1, \ldots, J' \tag{17}$$

where $\hat{\mathbf{d}}(f_j)$ is computed as in (16), and $\mathbf{P}_1^{\perp}$ and Δ are each $L \times L$ matrices defined as

$$\mathbf{P}_1^{\perp} = \mathbf{I}_L - \frac{1}{L}\mathbf{1}\mathbf{1}^T$$
$$\Delta = \operatorname{diag}\left\{e^{-j2\pi \frac{k_o - L'}{N}}, \ldots, e^{-j2\pi \frac{k_o}{N}}, \ldots, e^{-j2\pi \frac{k_o + L'}{N}}\right\} \tag{18}$$

where $\mathbf{1}$ is an $L \times 1$ vector composed of all ones. The aliased analog frequency is then estimated as $\hat{F}_j^a = F_s \hat{f}_j$, where $F_s = 250$ MHz in our prototype system. Rather than develop the formula in (17) as a simplification of the general Beamspace ESPRIT algorithm presented in [5], due to space limitations, we here simply present a proof that it works when $\hat{\mathbf{d}}(f_j) = \mathbf{d}(f_j)$.

Proof: First, we need to define some quantities. Let $\mathbf{W}$ denote an $L \times N$ matrix whose rows are L successive rows of the $N \times N$ DFT matrix associated with the DFT indices $k_o - L', \ldots, k_o, \ldots, k_o + L'$. Let $\mathbf{W}_1$ and $\mathbf{W}_2$ be composed of the first and last $N-1$ columns of $\mathbf{W}$, respectively. $\mathbf{W}_1$ and $\mathbf{W}_2$ are each $Lx(N-1)$ and related as $\mathbf{W}_2 = \Delta\mathbf{W}_1$. Finally, let $\mathbf{w}_N$ denote the last column of $\mathbf{W}$; $\mathbf{w}_N = \operatorname{diag}(\Delta^{N-1})$, where diag $(\bullet)$ converts the $L \times L$ diagonal matrix Δ^{N-1} to an $L \times 1$ column vector. Note that the first column of $\mathbf{W}$ is $\mathbf{1}$ such that $\mathbf{P}_1^{\perp}\mathbf{W} = [\mathbf{0}_L \ \vdots \ \mathbf{P}_1^{\perp}\mathbf{W}_2] = [\mathbf{0}_L \ \vdots \ \mathbf{P}_1^{\perp}\Delta\mathbf{W}_1]$, where $\mathbf{0}_L$ is an $L \times 1$ vector composed of all zeroes.

Next, define $\mathbf{v}(f)$ as the $N \times 1$ Vandermonde vector $\mathbf{v}(f) = [1, e^{j2\pi f}, e^{j4\pi f}, \ldots, e^{j(N-1)2\pi f}]^T$. Let $\mathbf{v}_1(f)$ and $\mathbf{v}_2(f)$ be composed of the first and last $N-1$ elements of $\mathbf{v}(f)$, respectively. $\mathbf{v}_1(f)$ and $\mathbf{v}_2(f)$ are each $(N-1) \times 1$ and related as $\mathbf{v}_2(f) = e^{j2\pi f}\mathbf{v}_1(f)$. With these definitions and relationships, it follows that

$$\begin{aligned}
&\mathbf{d}^H(f_j)\mathbf{\Delta}^*\mathbf{P}_1^{\perp}\mathbf{d}(f_j) \\
&\quad = \mathbf{v}^H(f_j)\mathbf{W}^H\mathbf{\Delta}^*\mathbf{P}_1^{\perp}\mathbf{W}\mathbf{v}(f_j) \\
&\quad = \mathbf{v}^H(f_j)\mathbf{W}^H\mathbf{\Delta}^*\mathbf{P}_1^{\perp}\mathbf{\Delta}\mathbf{W}_1\mathbf{v}_1(f_j)e^{j2\pi f_j} \\
&\quad = \left\{\mathbf{v}_1^H(f_j)\mathbf{W}_1^H + e^{-j(N-1)2\pi f_j}\mathbf{w}_N^H\right\} \\
&\qquad \times \mathbf{\Delta}^*\mathbf{P}_1^{\perp}\mathbf{\Delta}\mathbf{W}_1\mathbf{v}_1(f_j)e^{j2\pi f_j} \\
&\quad = \mathbf{v}_1^H(f_j)\mathbf{W}_1^H\mathbf{\Delta}^*\mathbf{P}_1^{\perp}\mathbf{\Delta}\mathbf{W}_1\mathbf{v}_1(f_j)e^{j2\pi f_j} \\
&\qquad + \mathbf{1}^T\mathbf{P}_1^{\perp}\mathbf{\Delta}\mathbf{W}_1\mathbf{v}_1(f_j)e^{-j(N-2)2\pi f_j} \\
&\quad = \{\mathbf{v}_1^H(f_j)\mathbf{W}_1^H\mathbf{\Delta}^*\mathbf{P}_1^{\perp}\mathbf{\Delta}\mathbf{W}_1\mathbf{v}_1(f_j)\}e^{j2\pi f_j}
\end{aligned}$$

where we have used the fact that $\mathbf{w}_N^H\mathbf{\Delta}^* = \operatorname{diag}^H(\Delta^{N-1})\Delta^* = \operatorname{diag}^H(\Delta^{N-1}\Delta) = \operatorname{diag}^H(1_L) = 1^T$, where $\operatorname{diag}^H(\mathbf{D})$ is intended to mean convert the diagonal matrix $\mathbf{D}$ to a column vector and conjugate transpose (in that order).

Since $\mathbf{v}_1^H(f_j)\mathbf{W}_1^H\mathbf{\Delta}^*\mathbf{P}_1^{\perp}\mathbf{\Delta}\mathbf{W}_1\mathbf{v}_1(f_j)$ is real valued, it follows that $\arg\{\mathbf{d}^H(f_j)\Delta^*\mathbf{P}_1^{\perp}\mathbf{d}(f_j)\} = 2\pi f_j$. □

Comparing the Direct and Indirect ESPRIT methods, in the former, the phase of the jth eigenvalue of $\mathbf{\Psi}$ is multiplied by $\pi \times 10^9$, whereas in the latter, the phase of $\hat{\mathbf{d}}^H(f_j)\Delta^*\mathbf{P}_1^{\perp}\hat{\mathbf{d}}(f_j)$ is multiplied by $250 \times 10^6/2\pi$. The multiplicative factor in the latter is three orders of magnitude lower than that in the former. This is a heuristic explanation as to why the performance of

the Indirect ESPRIT method comes much closer to achieving the CRB than the Direct ESPRIT method.

The formula for translating F_j^a up to the proper aliasing zone is dictated by Fig. 4, wherein the phase of the ESPRIT eigenvalue $\mu = e^{-j\kappa 2\pi F\tau}$, where $\tau = 0.5 \times 10^{-9}$ s, is plotted as a function of the analog baseband frequency $0 \leq F \leq 1$ GHz. Within the interval $(-\pi, \pi)$ are eight disjoint permissible intervals, each having a width of $\pi/8$ and a one-to-one correspondence with each of the eight aliasing zones depicted in Fig. 3. If the phase of the ESPRIT eigenvalue lies within one of these permissible intervals, F_j^a is translated to the corresponding aliasing zone accordingly, where it is either added to or subtracted from the appropriate integer multiple of 250 MHz. If, due to noise, the phase of the ESPRIT eigenvalue lies within one of the impermissible regions, it is projected into the nearest permissible region. This decomposes the interval $(-\pi, \pi)$ into eight distinct intervals, each having a width of $\pi/4$ and having a one-to-one correspondence with each of the eight possible aliasing zones plotted in Fig. 3. The baseband frequency of the jth source F_j is ultimately determined from the aliased frequency estimate $\hat{F}_j^a$ according to

$$\begin{aligned} \hat{F}_j &= \left| \hat{F}_j^a - 250 \times 10^6 \text{round}\left\{ \frac{\arg\{\mu_j\} + \pi/16}{\pi/4} \right\} \right| \text{ Hz} \\ &\quad \text{for } -15\pi/16 < \arg\{\mu\} < \pi \\ \hat{F}_j &= 1 \times 10^9 - \hat{F}_j^a \text{ Hz} \\ &\quad \text{for } -\pi < \arg\{\mu_j\} < -15\pi/16 \end{aligned} \tag{19}$$

where round $[x]$ is the nearest integer to x, as defined previously.

As an example, if $\arg\{\mu_j\}$ is either in the impermissible region $\pi/16 < \arg\{\mu_j\} < \pi/8$, the permissible region $\pi/8 < \arg\{\mu_j\} < 2\pi/8$, or the impermissible region $2\pi/8 < \arg\{\mu_j\} < 5\pi/16$, $\hat{F}_j^a$ is subtracted from 250 MHz to obtain $\hat{F}_j$. Simulations presented in Section VI reveal (19) to be a very robust formula for translating $\hat{F}_j^a$ to the proper aliasing zone. Note that if we are off by one in selecting the correct aliasing zone, a very large error may be incurred. Two adjacent aliasing zones differ in that in one F_j^a is added to nF_s, whereas in the other, it is subtracted from $(n+1)F_s$.

IV. 2-D Angle Estimation with Spatial Undersampling via PRO-ESPRIT and Integer Search

A. Estimation of the Array Manifold for Each Source

In Section III, we saw that use of the right generalized eigenvectors of the $L \times M$ matrix pencil $\{\mathbf{Y}, \mathbf{X}\}$ facilitates a simple procedure for estimating the aliased frequency of a source that was automatically paired with an ESPRIT eigenvalue, thereby, in turn, facilitating simple translation up to the proper aliasing zone. The left generalized eigenvectors of the $L \times M$ matrix pencil $\{\mathbf{Y}, \mathbf{X}\}$ play a similar role in the problem of estimating the azimuth and elevation angle of each source contributing to a given peak in the DFT spectrum. Specifically, the jth left generalized eigenvector of $\{\mathbf{Y}, \mathbf{X}\}$ is used to extract from the X and Y data an estimate of the array manifold for the jth source, which is denoted $\mathbf{a}(\theta_j, \phi_j, \kappa_j)$. Recall the inclusion of κ_j in the definition of the array manifold is a notational tool to reflect the fact that the array manifold is conjugated when the base-band frequency is located on a downward sloping portion of the aliasing function.

The jth $L \times 1$ left generalized eigenvector $\mathbf{l}_j$ of the $L \times M$ rectangular matrix pencil $\{\mathbf{Y}, \mathbf{X}\}$ is that vector satisfying $\mathbf{l}_j^H\{\mathbf{Y} - \mu_j\mathbf{X}\} = \mathbf{0}$. Substituting the noiseless (ideal) forms of the X and Y data matrices, we have

$$\mathbf{l}_j^H \left\{ \sum_{j=1}^{J} A_j' e^{j\kappa_j \gamma_{j^o}} \left\{ e^{-j\kappa_j 2\pi F_j \tau} - \mu \right\} \mathbf{d}(f_j) \mathbf{a}^T(\theta_j, \phi_j, \kappa_j) \right\} = \mathbf{0}. \tag{20}$$

When $\mu = e^{-j\kappa_l 2\pi F_l \tau}$, the lth term $\mathbf{d}(f_l)\mathbf{a}^T(\theta_l, \phi_l, \kappa_l)$ drops out of the sum such that $\mathbf{l}_j^H \mathbf{d}(f_l) = \mathbf{0}$ for $l = 1, \ldots, J', l \neq j$. Hence, $\mathbf{l}_j$ can be used to extract $\mathbf{a}(\theta_j, \phi_j, \kappa_j)$ to within a scalar multiple:

$$\begin{aligned} \mathbf{l}_j^H \mathbf{X} &\propto \mathbf{a}^T(\theta_j, \phi_j, \kappa_j) \Rightarrow \mathbf{X}^T \mathbf{l}_j^* &&\propto \mathbf{a}(\theta_j, \phi_j, \kappa_j) \\ \mathbf{l}_j^H \mathbf{Y} &\propto \mathbf{a}^T(\theta_j, \phi_j, \kappa_j) \Rightarrow \mathbf{Y}^T \mathbf{l}_j^* &&\propto \mathbf{a}(\theta_j, \phi_j, \kappa_j). \end{aligned} \tag{21}$$

Thus, applying the jth left generalized eigenvector allows us to extract an estimate of the array manifold for the jth source, which, in turn, may be operated on to estimate the azimuth and elevation angles of the jth source. The latter problem is greatly simplified, specifically in cases where sources are very closely spaced in digital frequency due to aliasing, due to the ability to isolate a single source contribution. In addition, since $\mathbf{l}_j$ is associated with the ESPRIT eigenvalue $\mu_j = e^{-jk_j 2\pi F_j \tau}$, the azimuth and elevation angle estimates obtained by processing the estimate of $\mathbf{a}(\theta_j, \phi_j, \kappa_j)$ are automatically paired with the estimate of F_j obtained via the algorithm developed in Section III. Knowledge of F_j is tantamount to knowledge of the proper aliasing zone. This allows us to determine the value of κ_j, enabling us to resolve a 180° ambiguity in the azimuth angle estimate (flipping the sign of κ_j introduces a 180° change in azimuth angle).

Similar to the case with the right generalized eigenvectors, the jth $L \times 1$ left generalized eigenvector $\mathbf{l}_j$ of the $L \times M$ rectangular matrix pencil $\{\mathbf{Y}, \mathbf{X}\}$ may be efficiently computed from the $J' \times 1$ jth left eigenvector α_j of $\mathbf{\Psi}$ in (13), satisfying $\alpha_j^H \mathbf{\Psi} = \mu_j \alpha_j^H, j = 1, \ldots, J'$. From (14) and (20), it follows that

$$\mathbf{l}_j = \mathbf{U}_S \Sigma_S^{-1} \alpha_j \quad j = 1, \ldots, J'. \tag{22}$$

Recall that J' is the number of sources making a nonnegligible contribution to a particular DFT spectral peak, which may be as small as one if sources are well separated in digital frequency.

In general, the problem of 2-D angle estimation is considerably more computationally complex than the problem of 1-D angle estimation. Fortuitously, the isolation of single-source components via PROESPRIT facilitates separable 2-D angle estimation given an appropriate array geometry. For example, consider a 2-D array consisting of two orthogonal linear arrays, e.g., an L-shaped array. Since we have isolated a single source component, we can determine the direction

cosine of a source relative to each axis independently. Each leg may be processed independently applying an appropriate 1-D angle-estimation algorithm. The x and y direction cosines are automatically paired with each other as well as with the corresponding frequency estimate. Simple trigonometry may be invoked to convert the x and y direction cosines into azimuth and elevation angle estimates.

4.2 Prescription for Nonuniform Element Spacing Facilitating Nonambiguous Angle Estimation

In accordance with the discussion in Section I, to achieve a high degree of resolution power and estimator accuracy and yet avoid mutual coupling, the elements of each leg of the L-shaped array are spaced nonuniformly with interelement spacings much greater than a half wavelength. The design problem is twofold: i) development of a prescription for "good" interelement spacings for unambiguous angle estimation relative to each array axis and ii) development of a computationally simple algorithm for processing the estimate of the array manifold provided by PRO-ESPRIT to estimate the direction cosine of a source with respect to each axis. Here, we assume a small number of antenna elements due to cost and complexity of hardware considerations and space limitations on the antenna platform attached to the aircraft.

The L-shaped array geometry employed in the simulations presented in Section VI is depicted in Fig. 2. The corresponding array manifold is shown in (23), which appears at the bottom of the page, where u_j and v_j are the direction cosines of the jth source relative to the x and y axes, respectively, and λ_j is the wavelength of the jth source related to the RF frequency as $c = \lambda_j F_j^{rf}$, where c is the speed of light: $\boldsymbol{a}(\theta_j, \phi_j, \kappa_j)$ may be estimated to within an unknown scalar multiple as

$$\begin{aligned} \boldsymbol{z}_{x_j} &= \boldsymbol{X}^T \boldsymbol{l}_j^* = \sigma_{x_j} \boldsymbol{a}(\theta_j, \phi_j, \kappa_j) \\ \boldsymbol{z}_{y_j} &= \boldsymbol{Y}^T \boldsymbol{l}_j^* = \sigma_{y_j} \boldsymbol{a}(\theta_j, \phi_j, \kappa_j) \end{aligned} \tag{24}$$

The processing of the first five elements of the estimate of $\boldsymbol{a}(\theta_j, \phi_j, \kappa_j)$, which is mathematically represented as

$$\begin{aligned} \boldsymbol{z}_{x_j}^v &= \left[\boldsymbol{I}_5 \vdots \boldsymbol{O}_{5x4}\right] \boldsymbol{z}_{x_j} = \sigma_{x_j} \boldsymbol{a}(v_j, \kappa_j) \\ \boldsymbol{z}_{y_j}^v &= \left[\boldsymbol{I}_5 \vdots \boldsymbol{O}_{5x4}\right] \boldsymbol{z}_{y_j} = \sigma_{y_j} \boldsymbol{a}(v_j, \kappa_j) \end{aligned} \tag{25}$$

corresponds to processing the leg parallel to the y axis and yields a measurement of $v_j = \sin\theta_j \sin\phi_j$ $(-1 \le v_j \le 1)$. Here, $\boldsymbol{a}(v_j, \kappa_j) = \left[\boldsymbol{I}_5 \vdots \boldsymbol{O}_{5x4}\right] \boldsymbol{a}(\theta_j, \phi_j, \kappa_j)$. Likewise, the processing of the last five elements of $\boldsymbol{a}(\theta_j, \phi_j, \kappa_j)$, which is mathematically represented as

$$\begin{aligned} \boldsymbol{z}_{x_j}^u &= \left[\boldsymbol{O}_{5x4} \vdots \boldsymbol{I}_5\right] \boldsymbol{z}_{x_j} = \sigma_{x_j} \boldsymbol{a}(u_j, \kappa_j) \\ \boldsymbol{z}_{y_j}^u &= \left[\boldsymbol{O}_{5x4} \vdots \boldsymbol{I}_5\right] \boldsymbol{z}_{y_j} = \sigma_{y_j} \boldsymbol{a}(u_j, \kappa_j) \end{aligned} \tag{26}$$

corresponds to processing the leg parallel to the x axis and yields a measurement of $u_j = \cos\theta_j \sin\phi_j$. Here, $\boldsymbol{a}(u_j, \kappa_j) = \left[\boldsymbol{O}_{5x4} \vdots \boldsymbol{I}_5\right] \boldsymbol{a}(\theta_j, \phi_j, \kappa_j)$.

In the development to follow, we concentrate primarily on the estimation of v_j. The procedure for estimating u_j is nearly identical. We begin the development by determining restrictions on d_1 and d_2 so that the direction cosine of a source relative to the y axis may be unambiguously determined from the interelement phase measurements provided by elements 1, 2, and 3 in Fig. 4. We then proceed to incorporate the other interelement phase measurements available to us. In the development, the true direction cosine of the jth source v_j is denoted as v_{jo} in order to not confuse j with other integer indices to be introduced.

Referring to the vertical leg of the L-array pictured in Fig. 2, the element pair 1-2 provides a measurement of $e^{j\frac{2\pi}{\lambda_j} d_1 v_{jo}} = e^{j\frac{2\pi}{\lambda_j} d_1 \left(v_{jo} + \tilde{k}\frac{\lambda_j}{d_1}\right)}$, where $\tilde{k}$ is an integer. Since $-1 \le v_{jo} \le 1$, the ambiguous angles resulting from the 1-2 measurement are ($\tilde{k}$, integer)

$$\begin{aligned} v_{\tilde{k}} &= v_{jo} + \tilde{k}\frac{\lambda_j}{d_1} \\ \tilde{k} &\in \left\{ \text{ceiling}\left[-\frac{d_1}{\lambda_j}(1 + v_{jo})\right], \text{floor}\left[\frac{d_1}{\lambda_j}(1 - v_{jo})\right] \right\} \end{aligned} \tag{27}$$

where ceiling $[x]$ is the integer closest to but greater than x, and floor $[x]$ is the integer closest to but less than x. It is thus apparent that if $d_1 \gg \lambda_j/2$, ambiguities arise. The ambiguity can only be resolved with another spatial sample. Referring to the vertical leg of the L-array pictured in Fig, 2, the ambiguous angles associated with the 1-3 measurement ($\tilde{l}$, integer) are

$$\begin{aligned} v_{\tilde{l}} &= v_{jo} + \tilde{l}\frac{\lambda_j}{d_2} \\ \tilde{l} &\in \left\{ \text{ceiling}\left[-\frac{d_2}{\lambda_j}(1 + v_{jo})\right], \text{floor}\left[\frac{d_2}{\lambda_j}(1 - v_{jo})\right] \right\}. \end{aligned} \tag{28}$$

The objective is to choose d_1 and d_2 so that alignment, i.e., $v_{\tilde{k}} = v_{\tilde{l}}$, only occurs for $\tilde{k} = \tilde{l} = 0$.

Equating the expressions for $v_{\tilde{k}}$ and $v_{\tilde{l}}$

$$v_{jo} + \tilde{k}\frac{\lambda_j}{d_1} = v_{jo} + \tilde{l}\frac{\lambda_j}{d_2} \rightarrow \frac{d_2}{d_1} = \frac{\tilde{l}}{\tilde{k}}. \tag{29}$$

This indicates that ambiguities may possibly arise if $\frac{d_2}{d_1}$ is rational. Express $\frac{d_2}{d_1}$ as $\frac{d_2}{d_1} = \frac{l'}{k'}$, where l' and k' are relatively

$$\boldsymbol{a}(\theta_j, \phi_j, \kappa_j) = \left[e^{j\kappa_j 2\pi \frac{2d_2}{\lambda_j} v_j}, e^{j\kappa_j 2\pi \frac{2d_2 - d_1}{\lambda_j} v_j}, e^{j\kappa_j 2\pi \frac{d_2}{\lambda_j} v_j}, e^{j\kappa_j 2\pi \frac{d_1}{\lambda_j} v_j}, 1, e^{j\kappa_j 2\pi \frac{d_1}{\lambda_j} u_j}, e^{j\kappa_j 2\pi \frac{d_2}{\lambda_j} u_j}, e^{j\kappa_j 2\pi \frac{2d_2 - d_1}{\lambda_j} u_j}, e^{j\kappa_j 2\pi \frac{2d_2}{\lambda_j} u_j} \right]^T \tag{23}$$

prime, i.e., have no common factors other than unity. The set of ambiguous angles is then $v_n = v_{jo} + nk'\frac{\lambda_j}{d_1} = v_{jo} + nl'\frac{\lambda_j}{d_2}$ for any n for which $-1 < v_n < 1$. Consider the case of $n = 1$. If we make sure that either $v_{jo} + k'\frac{\lambda_j}{d_1}$ **or** $v_{jo} + l'\frac{\lambda_j}{d_2}$ lies outside the visible region, i.e., is either less than -1 or greater than $+1$, then there is no ambiguity, that is, within the visible region corresponding to direction cosines with absolute value less than 1, there is only alignment at $n = 0$ or $\tilde{k} = \tilde{l} = 0$.

Part of the design procedure then is to select d_1 and d_2 such that the relatively prime factors l' and k' comply with *one* of the following conditions: *Either*

$$k' \notin \left\{\text{ceiling}\left[-\frac{d_1}{\lambda_j}(1+v_{jo})\right], \text{floor}\left[\frac{d_1}{\lambda_j}(1-v_{jo})\right]\right\} \text{ or}$$
$$l' \notin \left\{\text{ceiling}\left[-\frac{d_2}{\lambda_j}(1+v_{jo})\right], \text{floor}\left[\frac{d_2}{\lambda_j}(1-v_{jo})\right]\right\}.$$

These conditions depend on the direction cosine of the source. To remove the data dependence, we overspecify and let $v_{j^o} = 1$ for the lower bound limit and $v_{j^o} = -1$ for the upper bound limit. The goal then is to select d_1 and d_2 such that the relatively prime factors l' and k' comply with *one* of the following conditions:

$$k' \notin \left\{-\text{floor}\left[\frac{2d_1}{\lambda_j}\right], \text{floor}\left[\frac{2d_1}{\lambda_j}\right]\right\} \text{ or}$$
$$l' \notin \left\{-\text{floor}\left[\frac{2d_2}{\lambda_j}\right], \text{floor}\left[\frac{2d_2}{\lambda_j}\right]\right\}. \quad (30)$$

With $\frac{d_2}{d_1} = \frac{l'}{k'}$, where k' and l' are relatively prime, if either of the conditions above are satisfied, then within the visible region, the ambiguities only align at the true source direction cosine v_{j^o}.

Note that satisfying the condition above at 18 GHz guarantees that ambiguities may be resolved at lower frequencies since $|k'_{\min}| = |k'_{\max}| = \text{floor}[2d_1/\lambda_j]$ decreases with decreasing frequency (increasing λ_j) as does $|l'_{\min}| = |l'_{\max}| = \text{floor}[2d_2/\lambda_j]$.

As an illustrative example, for the simulations, we chose $d_1 = 2.3$ in and $d_2 = 5.3$ in. Consider the upper limit of the 2–18 GHz spectrum (18 GHz) for which the wavelength is $\lambda = 2/3$ in. (30) dictates that at 18 GHz, $k' \notin \{-6, 6\}$, and $\ell' \notin \{-14, 14\}$. Expressing d_2/d_1 as the ratio of two relatively prime numbers as $\frac{d_2}{d_1} = \frac{5.3}{2.3} = \frac{53}{23} = \frac{l'}{k'}$, we see that $k' = 23 \notin \{-6, 6\}$ and $l' = 53 \notin \{-14, 14\}$ so that both conditions in (30) are satisfied, and the direction cosine may be uniquely determined over the entire 2–18 GHz spectrum.

C. Integer Search Algorithm for Direction Cosine Estimation

We have shown that through judicious selection of the interelement spacings, it is theoretically possible to uniquely determine the true direction cosine. We now develop an algorithm to do just that. With respect to Fig 2, element pairs 1–2 and 4–5 provide two measurements of $\psi_1 = \arg\{e^{j2\pi\frac{d_1}{\lambda_j}v_{j^o}}\}$. The candidate estimates of v_{j^o} in the "visible" region $-1 \le v \le 1$ are

$$v_k^{(1)} = \frac{\lambda_j}{2\pi d_1}\psi_1 + k\frac{\lambda_j}{d_1}$$
$$k \in \left\{\text{ceiling}\left[-\frac{d_1}{\lambda_j} - \frac{\psi}{2\pi}\right], \text{floor}\left[\frac{d_1}{\lambda_j} - \frac{\psi_1}{2\pi}\right]\right\}. \quad (31)$$

Let k^* be that for which $v_{\text{jo}} = \frac{\lambda_j}{2\pi d_1}\psi_1 + k^*\frac{\lambda_j}{d_1}$. We will determine k^* by stepping through the integers in the range of k in (31), evaluating a metric for each corresponding $v_k^{(1)}$, and selecting that value for which the metric is minimum. An appropriate metric is developed below. Note that since d_1 is the smallest interelement spacing represented in the array, the number of ambiguous angles associated with the corresponding phase measurement ψ_1 is least. This is in line with the overridding goal of keeping the computational load as small as possible.

Element pairs 1–3 and 3–5 provide two measurements of $\psi_2 = \arg\{e^{j2\pi\frac{d_2}{\lambda_j}v_{\text{jo}}}\}$. The candidate estimates of v_{jo} in the "visible" region are

$$v_l^{(2)} = \frac{\lambda_j}{2\pi d_2}\psi_2 + l\frac{\lambda_j}{d_2}$$
$$l \in \left\{\text{ceiling}\left[-\frac{d_2}{\lambda_j} - \frac{\psi_2}{2\pi}\right], \text{floor}\left[\frac{d_2}{\lambda_j} - \frac{\psi_2}{2\pi}\right]\right\}. \quad (32)$$

Let l^* denote that value of l such that $v_{l^*}^{(2)} = v_{\text{jo}}$. Equating the expressions for $v_k^{(1)}$ and $v_l^{(2)}$ in (31) and (32), respectively, yields $\frac{\lambda_j}{2\pi d_1}\psi_1 + k\frac{\lambda_j}{d_1} = \frac{\lambda_j}{2\pi d_2}\psi_2 + l\frac{\lambda_j}{d_2}$. Selecting d_1 and d_2 in accordance with the prescription developed previously, $v_k^{(1)} = v_l^{(2)}$ only when $k = k^*$ and $l = l^*$. Solving for l yields

$$l = k\frac{d_2}{d_1} + \frac{d_2}{d_1} - \frac{\psi_1}{2\pi} - \frac{\psi_2}{2\pi}. \quad (33)$$

It follows that in stepping through the range of feasible integers k, (33) yields an integer value of l only when $k = k^*$ for which $l = l^*$.

An algorithm for determining k^* is then as follows. For each integer k in $\{\text{ceiling}\,[-\frac{d_1}{\lambda_j} - \frac{\psi_1}{2\pi}], \text{floor}\,[\frac{d_1}{\lambda_j} - \frac{\psi_1}{2\pi}]\}$, compute the corresponding l according to (33). Select k^* as that for which $|\ell - \text{round}[l]|$ is minimum, where round $[l]$ is the integer closest to l. Although this is a rather *ad hoc* technique, it is computationally simple, and simulations reveal that it performs very well with respect to resolving the ambiguity.

Thus far, we have only made use of the relative phase measurements associated with the interelement spacings d_1 and d_2. Element pairs 2–3 and 3–4 provide two measurements of $\psi_3 = \arg\{e^{j2\pi\frac{(d_2-d_1)}{\lambda_j}v_{\text{jo}}}\}$. Equating $v_m^{(3)} = \frac{\lambda_j}{2\pi(d_2-d_1)}\psi_3 + m\frac{\lambda_j}{d_2-d_1}$ with the expression for $v_k^{(1)}$ and solving for m yields

$$m = k\frac{d_2-d_1}{d_1} + \frac{d_2-d_1}{d_1}\frac{\psi_1}{2\pi} - \frac{\psi_3}{2\pi}. \quad (34)$$

Relative to the prototype array in Fig. 2, $\frac{d_2-d_1}{d_1} = \frac{3.0}{2.3} = \frac{30}{23} = \frac{m'}{k'}$. Equation (30) dictates that at 18 GHz, $k' \notin \{-6, 6\}$ and $m' \notin \{-8, 8\}$. Since $k' = 23$ and $m' = 30$, the conditions are satisfied so that (34) only yields an integer when $k = k^*$.

Similarly, element pairs 1–4 and 2–5 provide two measurements of $\psi_4 = \arg\{e^{j2\pi\frac{(2d_2-d_1)}{\lambda_j}v_{j^o}}\}$. Equating $v_n^{(4)} = \frac{\lambda_j}{2\pi(2d_2-d_1)}\psi_4 + n\frac{\lambda_j}{2d_2-d_1}$ with the expression for $v_k^{(1)}$ and

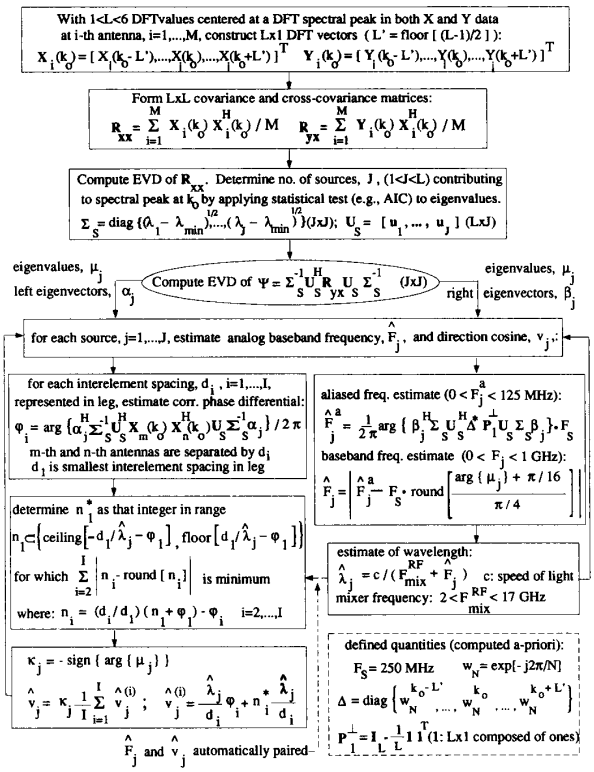


Fig. 5. Flowchart of frequency and 2-D angle estimation algorithm.

solving for n yields

$$n = k\frac{2d_2 - d_1}{d_1} + \frac{2d_2 - d_1}{d_1}\frac{\psi_1}{2\pi} - \frac{\psi_4}{2\pi}. \tag{35}$$

Relative to the prototype array in Fig. 2, $\frac{2d_2-d_1}{d_1} = \frac{8.3}{2.3} = \frac{83}{23} = \frac{n'}{k'}$. Equation (30) dictates that at 18 GHz, $k' \notin \{-6, 6\}$, and $n' \notin \{-22, 22\}$. Since $k' = 23$ and $n' = 83$, the conditions are satisfied so that (35) only yields an integer when $k = k^*$.

A refined algorithm for determining k^* is as follows. For each integer k in {ceiling $[-\frac{d_1}{\lambda_j} - \frac{\psi_1}{2\pi}]$, floor $[\frac{d_1}{\lambda_j} - \frac{\psi_1}{2\pi}]$}, compute the corresponding value of l, m, and n according to (33)–(35), respectively. Select k^* as that for which $|l\text{–round}[l]| + |m\text{–round}[m]| + |n\text{–round}[n]|$ is minimum. Once k^* is determined, compute l^*, m^*, and n^* by substituting k^* into (33)–(35), respectively. Compute the corresponding estimates of the direction cosine according to $v_{k^*}^{(1)} = \frac{\lambda_j}{2\pi d_1}\psi_1 + k^*\frac{\lambda_j}{d_1}$, $v_{l^*}^{(2)} = \frac{\lambda_j}{2\pi d_2}\psi_2 + l^*\frac{\lambda_j}{d_2}$, $v_{m^*}^{(3)} = \frac{\lambda_j}{2\pi(d_2-d_1)}\psi_3 + m^*\frac{\lambda_j}{d_2-d_1}$, and $v_{n^*}^{(4)} = \frac{\lambda_j}{2\pi(2d_2-d_1)}\psi_4 + n^*\frac{\lambda_j}{2d_2-d_1}$. The direction cosine relative to the vertical axis is estimated as a weighted sum of these estimates. Each direction cosine estimate is weighted by the corresponding interelement distance as the accuracy of the estimate increases with increasing distance, provided one can resolve the ambiguity.

A similar procedure may be used to estimate the direction cosine relative to the horizontal axis. A flowchart of the overall algorithm, including frequency estimation, is depicted in Fig. 5. The computational simplicity is evident. Note, that due to space limitations, the processing of the left eigenvectors indicated in the flowchart is only relative to a single leg and needs to be repeated for each leg.

TABLE I
SIGNAL PARAMETERS FOR SIMULATION EXAMPLE 1

j	$F_j(RF)$ (GHz)	F_j (MHz)	F_j^a (MHz)	$k_j = 128F_j^a/f_s$	θ_j (deg.)	ϕ_j (deg.)
1	(2/17).227	227	23	11.8	20	40
2	(2/17).275	275	25	12.8	50	30

V. SIMULATION EXAMPLES

The performance of the frequency and 2-D angle estimation algorithm summarized in the flowchart in Fig. 5 was examined in two simulation examples. Example 1 involves two sources very closely spaced in frequency after sampling due to aliasing. Example 2 represents a very stressful signal environment involving four sources that were very closely spaced in frequency after sampling. In both cases, simulations were conducted at the lower and upper ends of the 2–18 GHz spectrum. This was done to show that the algorithm works properly over a very wide bandwidth using the same physical array: the $M = 9$ element L-shaped array with geometry depicted in Fig. 2. Note that at 18 GHz, the wavelength is roughly 0.67 in. such that the smallest interelement spacing in the L-array ($d_1 = 2.3$ in) is roughly seven times a half wavelength. In general, both d_1 and d_2 are several times greater than a half wavelength at all frequencies in the band 2–18 GHz.

The simulation parameters indicated in Figs. 1 and 2 were common to all simulation runs. In all cases, the signal scenario was composed of equipowered RF-pulsed signals (monochromatic planewaves) that were "turned on" during the entire 0.5-μs interval in which 128 samples were collected. White Gaussian noise was added to the raw data samples output from each channel of each antenna, in accordance with the raw data model described in (36) and (37) of Appendix A, prior to computing the 128-point DFT. Finally, the CRB for a particular set of simulation parameters was computed according to expressions developed in Appendix A.

Example 1: The parameters describing the two signal arrivals are listed in Table I. In the one set of simulations, the signals were in the 2–3 GHz band, and the mixing frequency was 2 GHz, whereas in the other, the signals were in the 17–18 GHz band, and the mixing frequency was 17 GHz. A typical DFT spectrum representative of any of the 18 sampled channels (two channels for each of $M = 9$ antennas) for either signal band (2–3 GHz or 17–18 GHz) is plotted in Fig. 8. Due to their relative proximity, the two signal arrivals give rise to a single peak in the positive frequency portion of the spectrum. The frequency and 2-D angle estimation algorithm was applied to the DFT values in the range 11–14. In each run, the major computations were a 4×4 EVD followed by a 2×2 EVD. Sample statistics computed from 250 independent runs for each of a number of different SNR's are plotted in Figs. 6, 7, 9, and 10.

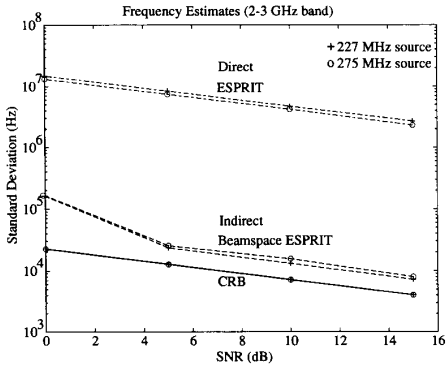


Fig. 6. Frequency estimation performance for Example 1 with signals in the 2–3 GHz band.

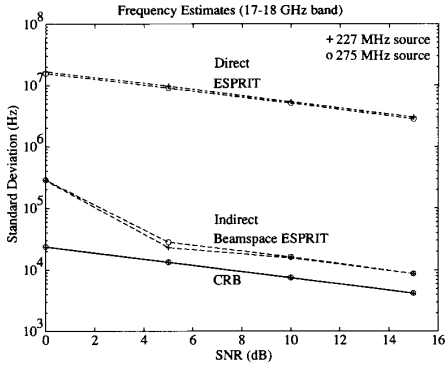


Fig. 9. Frequency estimationperformance for Example 1 with signals in 17–18 GHz band.

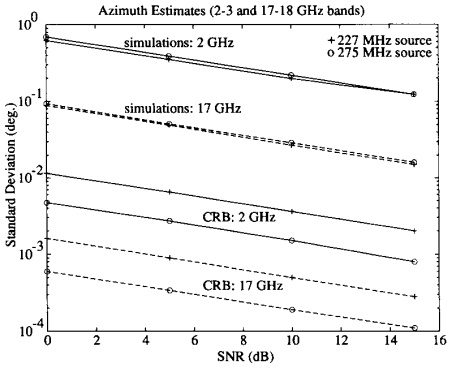


Fig. 7. Azimuth estimation performance for Example 1.

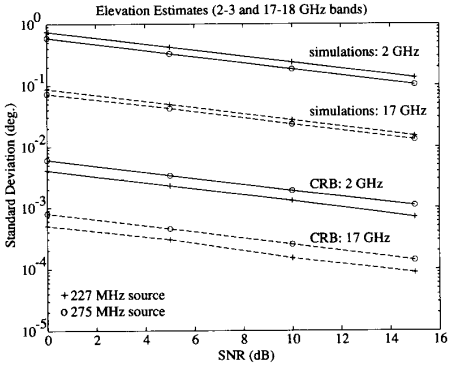


Fig. 10. Elevation estimation performance for Example 1.

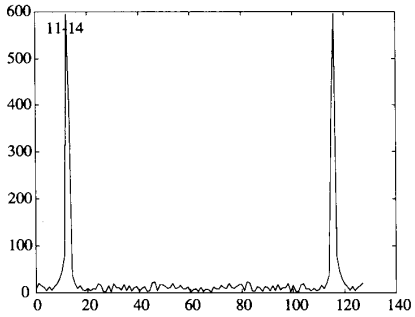


Fig. 8. Sample DFT spectrum of X data for Example 1.

Figs. 6 and 9 reveal the high variance of the Direct ESPRIT frequency estimates, which are three orders of magnitude greater than the CRB, in accordance with the discussion in Section III. The sample standard deviations of the Indirect Beamspace ESPRIT frequency estimates are very close to the CRB, particularly for SNR's greater than 4 dB. An important point to note is that despite how closely spaced the two sources are in frequency after aliasing, in all cases, i.e., for each source, for each SNR tested, and for each of 250 independent runs, the aliased frequency estimate obtained from Beamspace ESPRIT was translated to the proper aliasing zone. This demonstrates the robustness of the translation formula in (19). Note that the biases of the frequency estimates were always less than or equal to 1 MHz, which is negligible relative to the actual RF frequencies that are in the 2–18 GHz band.

Relative to the appropriate CRB, the performance of the angle estimation subroutine is not nearly as good as that of the frequency estimation subroutine. The sample standard deviations of the angle estimates obtained from the integer search algorithm are roughly two orders of magnitude greater than the CRB. This is true for both azimuth and elevation angle estimation as evidenced in Figs. 7 and 10, respectively, and for both ends of the 2–18 GHz spectrum. Better performance may be achieved by using the angle estimates from the integer search algorithm as starting points for localized Newton searches of a 1-D or 2-D MUSIC spectrum, or for initializing the expectation maximization algorithm, for example. However, imperfections in the hardware implementation of the algorithm may preclude achieving the CRB, which, for the case where the signals are in the 17–18 GHz band, is roughly 0.001°. It may be very difficult to achieve this kind of accuracy in practice even if it is achieved in simulation. Note that although the sample variances of the angle estimates were large relative to the CRB, the sample biases were very small. Although not plotted, the sample biases obtained in the 2–3 GHz range were less than 0.1° in all cases, even at 0 dB SNR, whereas the sample biases obtained in the 17–18 GHz range were less than 0.01° in all cases, even at 0 dB SNR.

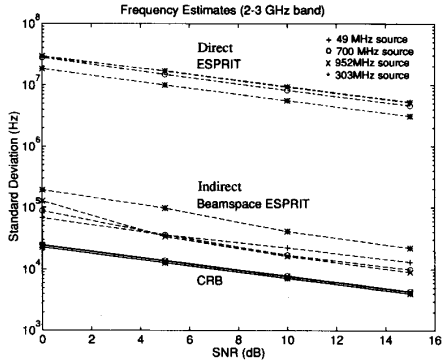


Fig. 11. Frequency estimation performance for Example 2 with signals in the 2–3 GHz band.

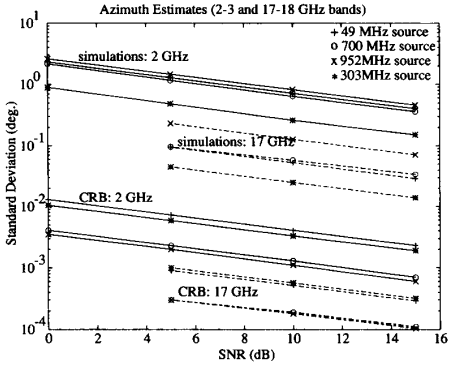


Fig. 12. Azimuth estimation performance for Example 2.

TABLE II
SIGNAL PARAMETERS FOR SIMULATION EXAMPLE 2

j	$F_j(RF)$ (GHz)	F_j (MHz)	F_j^a (MHz)	$k_j = 128F_j^a/f_s$	θ_j (deg.)	ϕ_j (deg.)
1	(2/17).952	952	48	24.6	120	15
2	(2/17).049	49	49	25.1	20	40
3	(2/17).700	700	50	25.6	50	30
4	(2/17).303	303	53	27.1	200	45

Again, for signals in the 17–18 GHz band, the smallest interelement spacing in the L array employed is roughly seven times greater than a half wavelength. For a given source in a given run and for a given leg of the array, the integer search algorithm had to choose which of roughly seven possible angles is the correct one. For all SNR's tested, the algorithm chose an angle in the vicinity of the actual angle in all 250 independent runs, despite how closely spaced the two sources were in frequency after aliasing.

Note that whereas the performance of the frequency-estimation phase of the algorithm did not vary significantly from one end of the 2–18 GHz spectrum to the other, the performance of the angle estimation phase of the algorithm did. The sample standard deviations of the angle estimates obtained in the 17–18 GHz range are roughly an order of magnitude smaller than those for the corresponding sources in the 2–3 GHz range. This is to be expected since the aperture length in terms of wavelengths at 18 GHz is roughly an order of magnitude greater than that at 2 GHz.

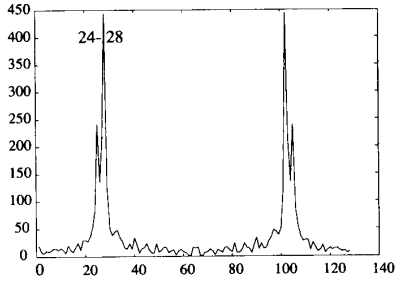


Fig. 13. Sample DFT spectrum of X data for Example 2.

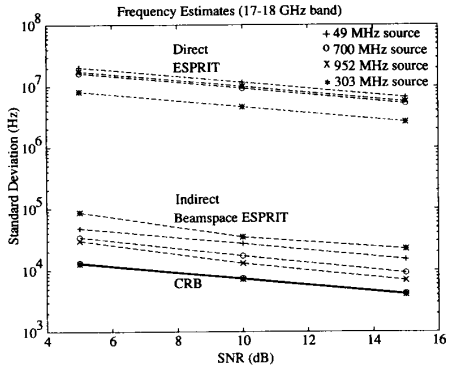


Fig. 14. Frequency estimation performance for Example 2 with signals in the 17–18 GHz band.

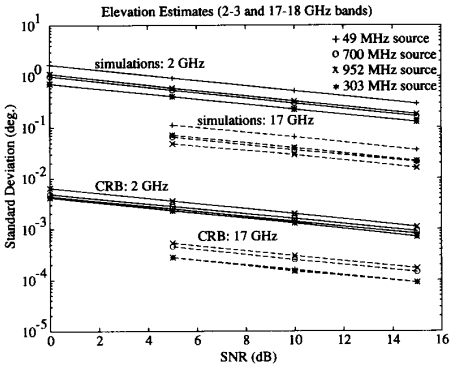


Fig. 15. Elevation estimation performance for Example 2.

Example 2: This simulation example is presented to demonstrate the power of the algorithm in light of the stressful nature of the signal scenario. The parameters describing each of the four signal arrivals simulated are listed in Table II. A typical DFT spectrum is plotted in Fig. 13. The four signal arrivals give rise to a single split peak in the positive frequency portion of the spectrum. The frequency and 2-D angle estimation algorithm was applied to the DFT values in the range 24–28. In each run, the major computations are a 5×5 EVD followed by a 4×4 EVD. More sample statistics computed from 250 independent runs for each of a number of different SNR's are plotted in Figs. 11, 12, 14, and 15.

Despite the fact that the four sources were all aliased to within a 4-MHz range, performance similar to that obtained in the much less stressful signal scenario of Example 1 was

achieved. Relative to the 17–18 GHz simulation, for a given source in a given run and for a given leg of the array, the integer search algorithm had to choose which of roughly seven possible angles is the correct one. For SNR's greater than or equal to 5 dB, the algorithm chose an angle in the vicinity of the actual angle in all 250 independent runs. At 0 dB, an erroneous angle was selected roughly 10% of the time. This yielded a very large sample variance not plotted in either Figs. 12, 14, or 15. Bearing in mind the stressful nature of the signal environment, four sources aliased to within a 4-MHz range, this is actually a remarkable performance.

VI. Final Comments

The frequency and 2-D angle estimation algorithm developed within and summarized in Fig. 5 is not able to handle sources that are aliased to exactly the same frequency. Examining Fig. 3, this will occur if i) two sources are separated in frequency by nF_s or ii) if one source is at $nF_s - \Delta F$, while another is at $nF_s + \Delta F$, where n is an integer. The failure of the algorithm in this case is due to a rank deficiency in the X and Y data matrices similar to the coherent signal problem encountered in array signal processing [9]. At the cost of a modest increase in computation, this deficiency may be overcome by working with spatial covariance matrices, as opposed to frequency domain covariance matrices, and performing a single forward-backward average when processing each leg of the L-array independently. The single forward-backward average is facilitated by the symmetric placing of elements along an axis. A more general measure would be to incorporate an additional sampled channel at a different rate, e.g., 225 MHz. This is the subject of ongoing investigation.

Appendix A
Computation of Cramer–Rao Lower Bound for Frequency and 2-D Angle Estimation

The data model used for calculating the CRB is the raw data output from the reference and time-delayed channels of each of M antenas. By raw data, we mean that which is prior to any processing, including the FFT (or DFT). Let $\mathbf{x}(n)$ denote the $M \times 1$ vector, where the ith component is the raw data output from the reference channel of the ith antenna $i = 1, \ldots, M$, at the nth sampling instant $n = 0, 1, \ldots, N-1$. Let $\mathbf{y}(n)$ be defined similarly relative to the time-delayed channel at each antenna. From the initial development in Section II, it follows that $\mathbf{x}(n)$ and $\mathbf{y}(n)$ may be expressed as

$$\mathbf{x}(n) = \mathrm{Re}\{\mathbf{A}\boldsymbol{\Omega}^n \boldsymbol{c}\} + \mathbf{n}_x(n) \quad n = 0, 1, \ldots, N-1 \tag{36}$$

$$\mathbf{y}(n) = \mathrm{Re}\{\mathbf{A}\boldsymbol{\Omega}^n \boldsymbol{\Phi} \boldsymbol{c}\} + \mathbf{n}_y(n) \quad n = 0, 1, \ldots, N-1. \tag{37}$$

The various quantities in (36) and (37) are defined below. $\boldsymbol{A}$ is the $M \times J$ DOA matrix

$$\mathbf{A} = [\mathbf{a}(\theta_1, \phi_1) \vdots \mathbf{a}(\theta_2, \phi_2) \vdots \cdots \vdots \mathbf{a}(\theta_J, \phi_J)] \tag{38}$$

where $\mathbf{a}(\theta_j, \phi_j)$ is defined by (23) with $\kappa_j = 1$. $\mathbf{c}$ is the $J \times 1$ vector

$$\mathbf{c} = [c_1, c_2, \ldots, c_J]^T = \bar{\mathbf{c}} + j\tilde{\mathbf{c}} \tag{39}$$

where $c_j = A_j e^{j\gamma_j \circ}$ is the complex amplitude of the jth source at time $n = 0$ at the reference element. $\boldsymbol{\Omega}$ is the $J \times J$ diagonal matrix

$$\boldsymbol{\Omega} = \mathrm{diag}\{e^{j\omega_1 T}, e^{j\omega_2 T}, \ldots, e^{j\omega_J T}\} \tag{40}$$

where $\omega_j = 2\pi F_j$ with F_j denoting the baseband analog frequency, and T is the sampling interval equal to the reciprocal of the sampling rate F_s. $\boldsymbol{\Phi}$ is the $J \times J$ diagonal matrix

$$\boldsymbol{\Phi} = \mathrm{diag}\{e^{j\omega_1 \tau}, e^{j\omega_2 \tau}, \ldots, e^{j\omega_J \tau}\} \tag{41}$$

where τ is the time delay equal to 0.5 ns $= 0.5 \times 10^{-9}$ s in our prototype system. $\mathbf{n}_x(n)$ and $\mathbf{n}_y(n), n = 0, 1, \ldots, N-1$, are i.i.d. multivariate Gaussian noise vectors $\mathbf{n}_x(n) \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_M)$ and $\mathbf{n}_y(n) \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_M)$.

Given the Gaussian assumption on the respective distributions of $\mathbf{n}_x(n)$ and $\mathbf{n}_y(n)$, it follows that $\mathbf{x}(n) \sim \mathcal{N}(\mathrm{Re}\{\mathbf{A}\boldsymbol{\Omega}^n\mathbf{c}\}, \sigma_n^2\mathbf{I}_M)$ and $\mathbf{y}(n) \sim \mathcal{N}(\mathrm{Re}\{\mathbf{A}\boldsymbol{\Omega}^n\boldsymbol{\Phi}\mathbf{c}\}, \sigma_n^2\mathbf{I}_M)$. The log-likelihood function is

$$\begin{aligned} \ln L(\omega, \theta, \phi, \bar{\mathbf{c}}, \tilde{\mathbf{c}}, \sigma_n^2) = \text{constant} &- NM \ln \sigma_n^2 \\ &- \frac{1}{2\sigma_n^2} \sum_{n=0}^{N-1} \|\mathbf{x}(n) - \mathrm{Re}\{\mathbf{A}\boldsymbol{\Omega}^n\mathbf{c}\}\|^2 \\ &- \frac{1}{2\sigma_n^2} \sum_{n=0}^{N-1} \|\mathbf{y}(n) - \mathrm{Re}\{\mathbf{A}\boldsymbol{\Omega}^n\boldsymbol{\Phi}\mathbf{c}\}\|^2. \end{aligned} \tag{42}$$

Let α denote the set of parameters on which the log-likelihood function depends. α contains $5J + 1$ parameters, which we group as follows: $\omega = [\omega_1, \omega_2, \ldots, \omega_J]^T$, $\theta = [\theta_1, \theta_2, \ldots, \theta_J]^T$, $\phi = [\phi_1, \phi_2, \ldots, \phi_J]^T$, $\bar{\mathbf{c}} = [A_1 \cos\gamma_{1o}, \ldots, A_J \cos\gamma_{Jo}]$, $\tilde{\mathbf{c}} = [A_1 \sin\gamma_{1o}, \ldots, A_J \sin\gamma_{Jo}]$, and σ_n^2 is the unknown noise power. Recall that J is the total number of sources.

With the $(5J + 1) \times (5J + 1)$ Fisher information matrix defined as $\mathbf{J} = \mathbf{I}(\alpha) = E\{\nabla_\alpha(\ln L)\nabla_\alpha^T(\ln L)\}$, the CRB on any unbiased estimator of the ith parameter α_i is $[\mathbf{J}^{-1}]_{ii}$, i.e., the ith diagonal element of the inverse of $\mathbf{J}$. Taking into account symmetry, the Fisher information matrix may be built up from the (1, 1) element $E[(\frac{\partial}{\partial\sigma_n^2}(\ln L))^2] = \frac{MN}{\sigma_n^4}$, the five $1 \times J$ blocks, $E[\frac{\partial}{\partial\sigma_n^2}(\ln L)\nabla_{\bar{c}}^T(\ln L)]$, $E[\frac{\partial}{\partial\sigma_n^2}(\ln L)\nabla_{\tilde{c}}^T(\ln L)]$, $E[\frac{\partial}{\partial\sigma_n^2}(\ln L)\nabla_\theta^T(\ln L)]$, $E[\frac{\partial}{\partial\sigma_n^2}(\ln L)\nabla_\phi^T(\ln L)]$, and $E[\frac{\partial}{\partial\sigma_n^2}(\ln L)\nabla_\omega^T(\ln L)]$, all of which are equal to 0, and the $J \times J$ blocks $E[\nabla_{\bar{c}}\nabla_{\bar{c}}^T]$, $E[\nabla_{\bar{c}}\nabla_{\tilde{c}}^T]$, $E[\nabla_{\bar{c}}\nabla_\theta^T]$, $E[\nabla_{\bar{c}}\nabla_\phi^T]$, $E[\nabla_{\bar{c}}\nabla_\omega^T]$, $E[\nabla_{\tilde{c}}\nabla_{\bar{c}}^T]$, $E[\nabla_{\tilde{c}}\nabla_{\tilde{c}}^T]$, $E[\nabla_{\tilde{c}}\nabla_\theta^T]$, $E[\nabla_{\tilde{c}}\nabla_\phi^T]$, $E[\nabla_{\tilde{c}}\nabla_\omega^T]$, $E[\nabla_\theta\nabla_\theta^T]$, $E[\nabla_\theta\nabla_\phi^T]$, $E[\nabla_\theta\nabla_\omega^T]$, $E[\nabla_\omega\nabla_\omega^T]$, $E[\nabla_\phi\nabla_\omega^T]$, and $E[\nabla_\phi\nabla_\omega^T]$, where it is understood that the function on which the gradient is operating in each case is the log-likelihood function in (42). The derivation of each block is straightforward. Due to space limitations, it is not feasible to present an expression for each of these 15 $J \times J$ blocks. As an example, though

$$\begin{aligned} E[\nabla_\theta\nabla_\phi^T] = &\frac{1}{\sigma_n^2} \sum_{n=0}^{N-1} \mathrm{Re}\{\boldsymbol{\Omega}^{*n}\mathbf{C}^*\mathbf{A}_\theta^H\}\mathrm{Re}\{\mathbf{A}_\phi\boldsymbol{\Omega}^n\mathbf{C}\} \\ &+ \frac{1}{\sigma_n^2} \sum_{n=0}^{N-1} \mathrm{Re}\{\boldsymbol{\Omega}^{*n}\mathbf{C}^*\boldsymbol{\Phi}^*\mathbf{A}_\theta^H\}\mathrm{Re}\{\mathbf{A}_\phi\boldsymbol{\Phi}\boldsymbol{\Omega}^n\mathbf{C}\} \end{aligned}$$

(43)

where $\mathbf{C}, \mathbf{A}_\theta$, and $\mathbf{A}_\theta$ are defined as follows:

$$\mathbf{C} = \text{diag}\{c_1, c_2, \ldots, c_J\} \tag{44}$$

$$\mathbf{A}_\theta = \left[\frac{\partial}{\partial\theta}\mathbf{a}(\theta,\phi)\Big|_{(\theta,\phi)=(\theta_j,\phi_j)}, j = 1, \ldots, J\right]$$

$$\mathbf{A}_\phi = \left[\frac{\partial}{\partial\phi}\mathbf{a}(\theta,\phi)\Big|_{(\theta,\phi)=(\theta_j,\phi_j)}, j = 1, \ldots, J\right]. \tag{45}$$

REFERENCES

[1] R. B. Sanderson, J. B. Y. Tsui and N. Freese, "Reduction of Ambiguities Through Phase Relations," *IEEE Trans. Aerospace and Electronic Systems*, Oct. 1992, pp. 950–956.

[2] R. Roy and T. Kailath, "ESPRIT–Estimation of Signal Parameters via Rotational Invariance Techniques," *IEEE Trans. Acoust., Speech, and Signal Process (ASSP)*, vol. 37, pp. 984–995, July 1989.

[3] M. D. Zoltowski and D. Stavrinides, "Sensor Array Signal Processing via a Procrusters Rotations Based Eigenanalysis of the ESPRIT Data Pencil," *IEEE Trans. ASSP*, vol. 37, June 1989, pp. 832–861.

[4] M. D. Zoltowski, G. M. Kautz and S. D. Silverstein, "Beamspace Root-MUSIC," *IEEE Trans. on Signal Processing*, vol. 41, Jan. 1993, pp. 344–364.

[5] G. Xu, S. D. Silverstein, R. Roy and T. Kailath, "Parallel Implementation and Performance Analysis of Beamspace ESPRIT," *Proc. 1991 IEEE ICASSP*, Apr. 1991, pp. 1497–1500.

[6] D. King, R. Packard and R. K. Thomas, "Unequally-Spaced Broad-Band Antenna Arrays," *IRE Trans. Antennas Propagat.*, vol. 8, Sept. 1960, pp. 498-500.

[7] A. Ishimaru, "Theory of Unequally-Spaced Arrays," *IRE Trans. Antennas Propagat.*, vol. 8, Nov. 1962, pp. 691–702.

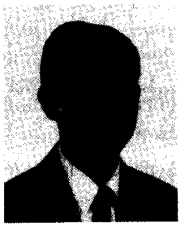
[8] M. Skolnik, G. Nemhauser and J. Sherman, "Dynamic Programming Applied to Unequally Spaced Arrays," *IRE Trans. Antennas Propagat.*, vol. 12, Jan. 1964, pp. 35–43.

[9] F. Haber and M. D. Zoltowski, "Spatial Spectrum Estimation in a Coherent Signal Environment Using an Array in Motion," *IEEE Trans. Antennas Propagat.*, vol. 34, Mar. 1986, pp. 301–310.

Michael D. Zoltowski (S'79–M'86) was born in Philadelphia, PA on August 12, 1960. He received both the B.S. and M.S. degrees in electrical engineering with highest honors from Drexel University in 1983 and the Ph.D. degree in systems engineering from the University of Pennsylvania in 1986.

From 1982 to 1986, he was an Office of Naval Research Graduate Fellow. In conjunction with this fellowship, he held a visiting research position at the Naval Research Laboratory in Washington, DC, during Summer 1986. In Fall 1986, he joined the faculty of Purdue University, where he currently holds the position of Associate Professor of Electrical Engineering. In this capacity, he was the recipient of the IEEE Outstanding Branch Counselor/Advisor Award for 1989–1990 and the Ruth and Joel Spira Outstanding Teacher Award for 1990–1991. During 1987, he was a Summer Faculty Research Fellow at the Naval Ocean Systems Center in San Diego, CA.

Dr. Zoltowski was the recipient of the IEEE Signal Processing Society's 1991 Paper Award (Statistical Signal and Array Processing Technical Area). He is a contributing author to *Adaptive Radar Detection and Estimation* (Wiley, 1991) and *Advances in Spectrum Analysis and Array Processing, Vol. III* (Prentice-Hall, 1994). He has served as a consultant to the General Electric Company. His present research interests include sensor array signal processing for mobile communications, radar, and noncooperatiuve electronic communications, adaptive beamforming, and higher order spectral analysis.



Cherian P. Mathews (M'93) was born in Madras, India, on October 26, 1966. He received the B.E. degree in electrical and electronics engineering from Anna University, Madras, in 1987, and the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1989 and 1993, respectively.

He was awarded a Purdue University dissertation fellowship in 1993. He is currently a Visiting Assistant Professor at Purdue University. He is a contributing author to *Advances in Spectrum Analysis and Array Processing, Vol. III* (Prentice-Hall, 1994). His current research interests are in sensor array signal processing for radar and mobile communications.