

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 16 June 1995	3. REPORT TYPE AND DATES COVERED Final, 01 April 1992 - 15 May 1995	
4. TITLE AND SUBTITLE Real-Time DOA Estimation of Wideband Signals with Multidimensional Arrays via Signal Subspace Techniques			5. FUNDING NUMBERS F49620-92-J-0198	
6. AUTHOR(S) Dr. Michael D. Zoltowski				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Purdue University, W. Lafayette, IN 47907			8. PERFORMING ORGANIZATION REPORT NUMBER Purdue TR-EE-95	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research 110 Duncan Avenue, Suite B115 Bolling AFB, DC 20332			10. SPONSORING/MONITORING AGENCY REPORT NUMBER AFOSR-TR-95 0487	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Distribution Unlimited		DISTRIBUTION STATEMENT A Approved for public release Distribution Unlimited DTIC SELECTED JUL 28 1995 F		
13. ABSTRACT (Maximum 200 words) 2D Unitary ESPRIT is developed as a closed-form 2-D angle estimation algorithm for use in conjunction with a uniform rectangular array (URA). In the final stage of the algorithm, the real and imaginary parts of the i-th eigenvalue of a matrix are one-to-one related to the respective direction cosines of the i-th source relative to the two major array axes. Reduced dimension beamspace implementations of 2D Unitary ESPRIT are developed along with adaptations for other array configurations including two orthogonal linear arrays. A novel approach to angle estimation in beamspace is also developed based on the observation that beamspace noise eigenvectors may be transformed to vectors in the element-space noise subspace. The transformed noise eigenvectors are bandpass, facilitating multirate processing involving modulation to baseband, filtering, and decimation. As these operations are linear, the Root-MUSIC (ESPRIT) based algorithm merely premultiplies each beamspace noise (signal) eigenvector by a precomputed transformation matrix. Compared to previous beamspace implementations of Root-MUSIC or ESPRIT, this approach places no restrictions on the structure of the matrix beamformer. Extensions for the URA are developed based on Multidimensional multirate processing. DTIC QUALITY INSPECTED 5				
14. SUBJECT TERMS angle estimation, antenna arrays multirate processing, beamforming frequency estimation, direction finding			15. NUMBER OF PAGES 132	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet optical scanning requirements.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.

NASA - Leave blank.

NTIS - Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

REAL-TIME DIRECTION-OF-ARRIVAL ESTIMATION OF WIDEBAND SIGNALS WITH MULTIDIMENSIONAL ARRAYS VIA SIGNAL SUBSPACE TECHNIQUES

Final Technical Report
Air Force Office of Scientific Research

Grant/Contract Number: F49620-92-J-0198
Period Covered: 01 April 1992 - 15 May 1995

Principal Investigator:

Michael D. Zoltowski

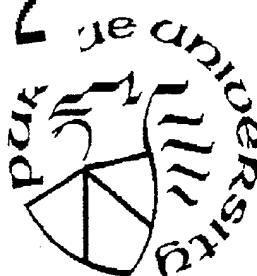
School of Electrical Engineering
1285 Electrical Engineering Building
Purdue University
West Lafayette, IN 47907 USA
e-mail: mikedz@ecn.purdue.edu
Phone: 317-494-3512
FAX: 317-494-0880

Program Director:

Jon A. Sjogren

AFOSR/NM
110 Duncan Ave., Suite B115
Bolling Air Force Base
Washington, DC 20332
jas@eng.umd.edu
Phone: 202-767-4940
FAX: 202-404-7496

19950727 012



Accession For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

Summary of Efforts

Closed-Form 2D Angle Estimation with Rectangular Arrays

UCA-ESPRIT is a recently developed closed form algorithm for use in conjunction with a uniform circular array (UCA) that provides automatically paired source azimuth and elevation angle estimates. *2D Unitary ESPRIT* is presented as an algorithm providing the same capabilities for a uniform rectangular array (URA). In the final stage of the algorithm, the real and imaginary parts of the i -th eigenvalue of a matrix are one-to-one related to the respective direction cosines of the i -th source relative to the two major array axes. *2D Unitary ESPRIT* offers a number of advantages over other recently proposed *ESPRIT* based closed-form 2D angle estimation techniques. First, except for the final eigenvalue decomposition of dimension equal to the number of sources, it is efficiently formulated in terms of real-valued computation throughout. Second, it is amenable to efficient beamspace implementations that will be presented. Third, it is applicable to array configurations that do not exhibit identical subarrays, e. g., two orthogonal linear arrays. Finally, *2D Unitary ESPRIT* easily handles sources having one member of the spatial frequency coordinate pair in common. Simulation results are presented verifying the efficacy of the method.

Beamspace DOA Estimation Featuring Multirate Eigenvector Processing

A novel approach to angle of arrival estimation in beamspace has been developed. Beamspace noise eigenvectors may be transformed to vectors in the element-space noise subspace. The transformed noise eigenvectors are bandpass, facilitating multirate processing involving modulation to baseband, filtering, and decimation. As these operations are linear, a matrix transformation applied to the eigenvectors may be constructed a priori. Incorporation of the technique into either the Root-MUSIC or ESPRIT prescriptions provides a computationally efficient procedure. Compared to past efforts to adapt Root-MUSIC and ESPRIT to beamspace, this approach circumvents the need for restrictive requirements on the form of the beamforming transformation. An asymptotic theoretical performance analysis is also included to provide an alternative to computationally intensive Monte-Carlo simulations. Simulation studies show the validity of the performance predictive expressions and verify that the procedure, when incorporated into the Root-MUSIC/ESPRIT formulations, produces a direction finding technique that nearly attains the Cramer-Rao bound.

Multidimensional Multirate DOA Estimation in Beamspace

The 1D multirate approach was extended to the more general case of 2D angle estimation with a uniform rectangular array (URA) of sensors. Multidimensional multirate processing is employed to ultimately yield a small order polynomial in two variables. Again, due to the linearity of the 2D filtering and 2D decimation operations, the actual algorithm merely premultiplies each beam space noise eigenvector by a precomputed transformation matrix. To avoid the spectral search, despite the fact that the fundamental theorem of algebra does not hold in 2D, we propose taking the orthogonal complement of the resulting transformed noise eigenvectors and applying a novel version of ESPRIT facilitating closed-form 2D angle estimation. Simulations demonstrating the efficacy of the approach are presented along with theoretical performance analysis.

Real-Time Frequency And 2-D Angle Estimation With Sub-Nyquist Spatio-Temporal Sampling

An algorithm has been developed for real-time estimation of the frequency and azimuth and elevation angles of each signal incident upon 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 .5 ns, 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 unambiguous 2-D angle estimation 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 automatically paired with corresponding frequency estimates despite aliasing. Simulations are presented demonstrating the capabilities of the algorithm.

Contents

1	Closed-Form 2D Angle Estimation with Rectangular Arrays	1
1.1	Introduction	2
1.2	Real-Valued Processing with Uniform Linear Array	4
1.3	Unitary ESPRIT for Uniform Linear Array	6
1.4	DFT Beam-space ESPRIT for Uniform Linear Array	9
1.4.1	Relationship Between Unitary ESPRIT and DFT Beam-space ESPRIT	11
1.4.2	Relationship Between DFT Beam-space ESPRIT and Beam-space ESPRIT	12
1.5	2D Unitary ESPRIT for Uniform Rectangular Array	13
1.5.1	2D Unitary ESPRIT vs. ACMP	17
1.6	2D DFT Beam-space ESPRIT for Uniform Rectangular Array	18
1.6.1	Reduced Dimension Example	20
1.6.2	Comparison with UCA-ESPRIT	21
1.7	2D DFT Beam-space ESPRIT for Cross Array	23
1.8	Simulations	25
1.9	Conclusions	27
1.10	References	28
1.11	Figures	31
2	Beam-space DOA Estimation Featuring Multirate Eigenvector Processing	34
2.1	Introduction	35
2.2	Array Signal Model	37
2.3	Development of DOA Estimators Featuring Multirate Eigenvector Processing	39
2.3.1	Multirate Noise Eigenvector Processing	39
2.3.2	Incorporation of Filter Deconvolution	43
2.3.3	Root-MUSIC Incorporating Multirate Eigenvector Processing	46
2.3.4	TLS-ESPRIT Incorporating Multirate Eigenvector Processing	48
2.3.5	Location of Extraneous Roots Created by Filtering	50
2.4	Theoretical Performance Analysis	51
2.4.1	Performance Analysis of Root-MUSIC Formulation	52
2.4.2	Performance Analysis of ESPRIT Formulation	53
2.5	Computer Simulations	54
2.6	Conclusions/Remarks	57
2.7	Appendix: Asymptotic Variance of ESPRIT Formulation	58
2.8	References	59
2.9	Figures	61
3	Multidimensional Multirate DOA Estimation in Beam-space	66
3.1	Introduction	67
3.2	Array Geometry	69
3.3	Beamforming	73

3.4	Eigenanalysis	74
3.5	Multirate Processing of Beam-space Noise Eigenvectors	75
3.6	TLS-ESPRIT	78
3.7	Bandlimiting the Response	82
3.8	Further Reductions in Complexity	84
3.8.1	Real Covariance Processing	84
3.8.2	Orthogonal Complement	85
3.9	Algorithm Summary	86
3.10	Performance Analysis	86
3.11	Computer Simulations	88
3.12	Conclusions	91
3.13	Appendix: Characterizing the Asymptotic Error	92
3.14	References	96
4	Real-Time Frequency And 2-D Angle Estimation With Sub-Nyquist Spatio-Temporal Sampling	98
4.1	Introduction	99
4.2	Spatio-Temporal Sampling and Data Model	101
4.3	ESPRIT Based Frequency Estimation With Temporal Undersampling	105
4.4	2-D Angle Estimation With Spatial Undersampling Via PRO-ESPRIT and Integer Search Formulation	110
4.4.1	Estimation of the Array Manifold for Each Source	110
4.4.2	Prescription for Nonuniform Element Spacing Facilitating Nonambiguous Angle Estimation	111
4.4.3	Integer Search Algorithm for Direction Cosine Estimation	114
4.5	Simulation Examples	116
4.6	Final Comments	118
4.7	References	119
4.8	Computation of Cramer Rao Lower Bound for Frequency and 2D Angle Estimation	119
4.9	Figures	122

1 Closed-Form 2D Angle Estimation with Rectangular Arrays

UCA-ESPRIT is a recently developed closed form algorithm for use in conjunction with a uniform circular array (UCA) that provides automatically paired source azimuth and elevation angle estimates. *2D Unitary ESPRIT* is presented as an algorithm providing the same capabilities for a uniform rectangular array (URA). In the final stage of the algorithm, the real and imaginary parts of the i -th eigenvalue of a matrix are one-to-one related to the respective direction cosines of the i -th source relative to the two major array axes. *2D Unitary ESPRIT* offers a number of advantages over other recently proposed *ESPRIT* based closed-form 2D angle estimation techniques. First, except for the final eigenvalue decomposition of dimension equal to the number of sources, it is efficiently formulated in terms of real-valued computation throughout. Second, it is amenable to efficient beamspace implementations that will be presented. Third, it is applicable to array configurations that do not exhibit identical subarrays, e. g., two orthogonal linear arrays. Finally, *2D Unitary ESPRIT* easily handles sources having one member of the spatial frequency coordinate pair in common. Simulation results are presented verifying the efficacy of the method.

1.1 Introduction

1.2 Real-Valued Processing with Uniform Linear Array

1.3 Unitary ESPRIT for Uniform Linear Array

1.4 DFT Beamspace ESPRIT for Uniform Linear Array

1.4.1 Relationship Between Unitary ESPRIT and DFT Beamspace ESPRIT

1.4.2 Relationship Between DFT Beamspace ESPRIT and Beamspace ESPRIT

1.5 2D Unitary ESPRIT for Uniform Rectangular Array

1.5.1 2D Unitary ESPRIT vs. ACMP

1.6 2D DFT Beamspace ESPRIT for Uniform Rectangular Array

1.6.1 Reduced Dimension Example

1.6.2 Comparison with UCA-ESPRIT

1.7 2D DFT Beamspace ESPRIT for Cross Array

1.8 Simulations

1.9 Conclusions

1.10 References

1.11 Figures

1 Introduction

For 1D arrays, if the elements are uniformly-spaced, *Root-MUSIC* and *ESPRIT*¹ [1] avert a spectral search in determining the direction of arrival (DOA) of each incident signal. Instead, the DOA of each signal is determined from the roots of a polynomial. For either *Root-MUSIC* or *ESPRIT*², the roots of interest ideally lie on the unit circle and are related one-to-one with each source as shown in Figure 1.

For 2D (planar) arrays, the fact that the fundamental theorem of algebra does not hold in two dimensions typically precludes a rooting type of formulation. Even for the highly regular uniform rectangular array (URA), *2D MUSIC* requires a spectral search of a multimodal two-dimensional surface, while both *Multiple Invariance ESPRIT* [2, 3] and Clark & Scharf's *2D IQML* [4] algorithm involve nonlinear optimization. Now, it should be pointed out that a URA lends itself to separable processing allowing one to decompose the 2D problem into two 1D problems. That is, one can estimate the DOA's with respect to one array axis via one set of calculations involving a *MUSIC* or *ESPRIT* based polynomial formulation, and also do the same with respect to another array axis. Coupling information may be employed to subsequently pair the respective members of the two sets of 1D angle estimates [5].

In the Algebraically Coupled Matrix Pencil (*ACMP*) method of van der Veen *et al*³ [6], eigenvector information is employed to pair the respective members of the two sets of 1D angle estimates. However, *ACMP* breaks down if two sources have the same arrival angle relative to either the x -axis or the y -axis, assuming the URA to lie in the x - y plane.

In contrast, for a uniform circular array (UCA) the recently developed *UCA-ESPRIT* [7, 8] algorithm provides closed-form, automatically paired 2D angle estimates as long as the azimuth and elevation angle of each signal arrival is unique. As illustrated in Figure 2, in the final stage of *UCA-ESPRIT*, the i -th eigenvalue of a matrix is of the form $\sin \theta_i e^{j\phi_i}$, where ϕ_i and θ_i are the azimuth and elevation angles of the i -th source. Note that $\sin \theta_i e^{j\phi_i} = u_i + jv_i$, where u_i and v_i are the direction cosines of the i -th source relative to the x and y axes, respectively. The eigenvalue for each source is thus unique such that *UCA-ESPRIT* does not have the aforementioned problem

¹ *ESPRIT* may also be employed in the case of an array composed of at least two translationally invariant subarrays.

² In *ESPRIT* the DOA's are extracted from eigenvalues which are roots of the characteristic polynomial of a matrix.

³ van der Veen *et al* do not actually give their method a name. In a later paper Vanpoucke *et al* label their method *ACMP*.

ACMP has when two sources have the same u_i or the same v_i . We here develop a closed-form 2D angle estimation algorithm for a URA that provides automatic pairing in a similar fashion. That is, in the final stage of new algorithm, referred to as *2D Unitary ESPRIT*, the real and imaginary parts of the i -th eigenvalue of a matrix are one-to-one related to u_i and v_i , respectively.

2D Unitary ESPRIT is developed as an extension of the recently proposed *Unitary ESPRIT* [9, 10] algorithm for a uniform linear array (ULA). *Unitary ESPRIT* exploits the conjugate centro-symmetry of the array manifold for a ULA to formulate each of the three primary stages of *ESPRIT* in terms of real-valued computations: (1) the computation of the signal eigenvectors, (2) the solution to the system of equations derived from these signal eigenvectors, and (3) the computation of the eigenvalues of the solution to the system of equations formed in stage 2. Note that Huarng & Yeh [11] and Linebarger *et al* [12] previously exploited the conjugate centro-symmetry of the ULA manifold to formulate the determination of the noise eigenvectors and subsequent spectral search required by *MUSIC* in terms of real-valued computation. The ability to formulate an *ESPRIT*-like algorithm for a ULA that only requires real-valued computations from start to finish, after an initial sparse unitary transformation, is critically important in developing a closed-form 2D angle estimation algorithm for a URA similar to *UCA-ESPRIT* for a UCA. *Unitary ESPRIT* is thus reviewed in Section 3 after a brief overview in Section 2 of C^N to $\mathbb{R}^N$ transformations facilitated by the conjugate centro-symmetry of the ULA manifold.

A reduced dimension beamspace version of *Unitary ESPRIT* is developed in Section 4. There are a number of advantages to working in beamspace: reduced computational complexity [13], decreased sensitivity to array imperfections [14], and lower SNR resolution thresholds [15]. In contrast to the *Beamspace ESPRIT* [16] algorithm of Xu *et al*, the beamspace version of *Unitary ESPRIT* exploits the real-valued nature of the beamspace manifold to formulate each of the three primary stages of *ESPRIT* in terms of real-valued computations as in *Unitary ESPRIT*, but in a reduced dimension space. Although the respective developments of *Unitary ESPRIT* and its beamspace counterpart proceed along markedly different lines, there is an interesting relationship between the two presented in Section 4.1. The relationship between *Beamspace ESPRIT* and the new beamspace version of *Unitary ESPRIT* is examined in Section 4.2.

2D Unitary ESPRIT is developed in Section 5. In addition to the ability to handle sources having the same arrival angle relative to either the x -axis or the y -axis, *2D Unitary ESPRIT* offers a number of advantages over other recently proposed *ESPRIT*-based closed-form 2D angle estimation

techniques including *ACMP*. First, except for the final eigenvalue decomposition of dimension equal to the number of sources, it is efficiently formulated in terms of real-valued computation throughout. Second, it is amenable to a reduced dimension beamspace implementation. In Section 6, we develop a beamspace version of *2D Unitary ESPRIT* as an extension of the beamspace version of *Unitary ESPRIT* presented in Section 4.

Another advantage of *2D Unitary ESPRIT* over *ACMP* is that the former is applicable to array configurations that do not exhibit identical subarrays, *e. g.*, two noncollinear ULA's. In contrast, *ACMP* requires an array of sensor triplets so that one can extract three identical subarrays from the overall array. *2D Unitary ESPRIT* only requires that the array exhibit invariances in two distinct directions. In Section 7, we show how *2D Unitary ESPRIT* may be simply adapted for the case of two orthogonal ULA's having a common phase center. *ACMP* is not applicable with such an array geometry.

Simulation results are presented in Section 8 verifying the efficacy of *2D Unitary ESPRIT* and its beamspace counterpart, and comparing their respective performances with the Cramer-Rao Lower Bound.

2 Real-Valued Processing with a ULA

All of the developments in this paper rely on some well known aspects of real-valued processing with a ULA which are quickly reviewed here [9, 10, 11, 12, 17]. Employing the center of the ULA as the phase reference, the array manifold is conjugate centro-symmetric. For example, if the number of elements comprising the ULA, N , is odd, there is a sensor located at the array center and the array manifold is

$$\mathbf{a}_N(\mu) = \left[e^{-j(\frac{N-1}{2})\mu}, \dots, e^{-j\mu}, 1, e^{j\mu}, \dots, e^{j(\frac{N-1}{2})\mu} \right]^T, \quad (1)$$

where $\mu = \frac{2\pi}{\lambda} \Delta_x u$ with λ equal to the wavelength, Δ_x is equal to the interelement spacing, and u equal to the direction cosine relative to the array axis. The conjugate centro-symmetry of $\mathbf{a}_N(\mu)$ is mathematically stated as $\mathbf{\Pi}_N \mathbf{a}_N(\mu) = \mathbf{a}_N^*(\mu)$, where

$$\mathbf{\Pi}_N = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{bmatrix} \in \mathbb{R}^{N \times N}. \quad (2)$$

As the inner product between any two conjugate centro-symmetric vectors is real-valued, any matrix whose rows are each conjugate centro-symmetric may be employed to transform the complex-valued

element space manifold, $\mathbf{a}_N(\mu)$, into a real-valued manifold. As noted by a numerous authors [9, 11, 12], the simplest matrices for accomplishing such are

$$\mathbf{Q}_{2K} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_K & j\mathbf{I}_K \\ \mathbf{\Pi}_K & -j\mathbf{\Pi}_K \end{bmatrix}, \quad (3)$$

if N is even, or

$$\mathbf{Q}_{2K+1} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_K & \mathbf{0} & j\mathbf{I}_K \\ \mathbf{0}^T & \sqrt{2} & \mathbf{0}^T \\ \mathbf{\Pi}_K & \mathbf{0} & -j\mathbf{\Pi}_K \end{bmatrix}, \quad (4)$$

if N is odd. $\mathbf{Q}_N^H$ is a sparse unitary matrix that transforms $\mathbf{a}_N(\mu)$ into an $N \times 1$ real-valued manifold, $\mathbf{d}_N(\mu) = \mathbf{Q}_N^H \mathbf{a}_N(\mu)$. For example, if the number of elements comprising the ULA is odd, the form in (4) is used and

$$\mathbf{d}_N(\mu) = \mathbf{Q}_N^H \mathbf{a}_N(\mu) = \sqrt{2} \times \left[\cos\left(\frac{N-1}{2}\mu\right), \dots, \cos(\mu), 1/\sqrt{2}, -\sin\left(\frac{N-1}{2}\mu\right), \dots, -\sin(\mu) \right]^T. \quad (5)$$

Let $\hat{\mathbf{R}}_{xx}$ denote the $N \times N$ complex-valued element space sample covariance matrix. Since the transformed manifold is real-valued, the signal eigenvectors required at the front end of *ESPRIT* may be computed as the “largest” eigenvectors of $\mathcal{R}e\{\mathbf{Q}_N^H \hat{\mathbf{R}}_{xx} \mathbf{Q}_N\}$. Note that in addition to the obvious computational reduction, taking the real part of the correlation matrix effects signal decorrelation [17] in the case of highly correlated or coherent sources. Alternatively, if $\mathbf{X}$ denotes the $N \times N_s$ element space data matrix containing N_s snapshots as columns, the signal eigenvectors may be computed as the “largest” left singular vectors of the real-valued matrix $\mathbf{Q}_N^H [\mathbf{X}, \mathbf{\Pi}_N \mathbf{X}^*] \mathbf{M}_{2N_s}$, where

$$\mathbf{M}_{2N_s} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{N_s} & j\mathbf{I}_{N_s} \\ \mathbf{I}_{N_s} & -j\mathbf{I}_{N_s} \end{bmatrix}. \quad (6)$$

Since $\mathbf{\Pi}_N \mathbf{Q}_N = \mathbf{Q}_N^*$, it follows that $\mathbf{Q}_N^H [\mathbf{X}, \mathbf{\Pi}_N \mathbf{X}^*] \mathbf{M}_{2N_s} = \sqrt{2} [\mathcal{R}e\{\mathbf{Y}\}, -\mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_N^H \mathbf{X}$. From a numerical point of view, the latter is preferable due to computational efficiency and robustness to dynamic range, especially if one employs an algorithm like the rank revealing URV decomposition [18].

Note that pre-multiplication of an $N \times 1$ vector by $\mathbf{Q}_N^H$ involves very little computation. In fact, it involves no multiplications (the scaling by $\sqrt{2}$ is unnecessary in computing the signal eigenvectors) and only N additions. In Section 4, we also consider the use of the N pt. DFT matrix, with appropriate scaling of the rows to make them each conjugate centro-symmetric [17], to transform the data into a real-valued beamspace. Although FFT's are fast, this approach ostensibly involves significantly more computation than the use of $\mathbf{Q}_N^H$. The utility of transforming to beamspace comes

into play when there is a-priori information on the general angular locations of the signal arrivals, as in a radar application, for example. In this case, one may only apply those rows of the DFT matrix that form beams encompassing the sector of interest. This yields a reduced dimension beamspace and leads to reduced computational complexity [13, 14, 15, 17]. This is possible due to the physical interpretation that the rows of the DFT matrix form beams pointed to different angles. There is no such physical interpretation for the rows of $\mathbf{Q}_N^H$ thereby precluding the possibility to work in a reduced dimension space.

Note that in this paper we do not address the problem of estimating the number of sources. We will assume an estimate is available via a procedure such as that described by Xu *et al* in [19] which explicitly exploits the conjugate centro-symmetry of the array manifold for a ULA.

3 Review of Unitary ESPRIT for ULA

As a precursor to developing an *ESPRIT* [1] based closed-form 2D angle estimation scheme for a URA, we first briefly review the recently proposed *Unitary ESPRIT* [9] algorithm for a uniform linear array (ULA) that only requires real-valued computations from start to finish after an initial sparse unitary transformation by $\mathbf{Q}_N^H$. As discussed above, if $\mathbf{X}$ denotes the $N \times N_s$ element space data matrix containing N_s snapshots as columns, the signal eigenvectors for *Unitary ESPRIT* may be computed as the “largest” left singular vectors of the real-valued matrix $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_N^H \mathbf{X}$. Assume that there are $d < N$ signal arrivals. Asymptotically, the $N \times d$ real-valued matrix of signal eigenvectors, $\mathbf{E}_S$, is related to the real-valued $N \times d$ DOA matrix, $\mathbf{D} = [\mathbf{d}(\mu_1), \mathbf{d}(\mu_2), \dots, \mathbf{d}(\mu_d)]$, as $\mathbf{E}_S = \mathbf{D}\mathbf{T}$, where $\mathbf{T}$ is an unknown $d \times d$ real-valued matrix.

Since $\mathbf{Q}_N^H$ is unitary, it follows that asymptotically (as the number of snapshots becomes infinitely large)

$$\mathbf{Q}_N \mathbf{E}_S = \mathbf{A} \mathbf{T}, \quad (7)$$

where $\mathbf{A} = [\mathbf{a}(\mu_1), \mathbf{a}(\mu_2), \dots, \mathbf{a}(\mu_d)]$, the $N \times d$ complex-valued element space DOA matrix. For a ULA, $\mathbf{A}$ satisfies the so-called invariance property [1]

$$\mathbf{J}_1 \mathbf{A} \Phi_\mu = \mathbf{J}_2 \mathbf{A} \quad \text{where: } \Phi_\mu = \text{diag}\{e^{j\mu_1}, e^{j\mu_2}, \dots, e^{j\mu_d}\}, \quad (8)$$

and $\mathbf{J}_1$ and $\mathbf{J}_2$ are the $(N-1) \times N$ matrices

$$\mathbf{J}_1 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \in \mathbb{R}^{(N-1) \times N} \quad (9)$$

$$\mathbf{J}_2 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{(N-1) \times N}. \quad (10)$$

$\mathbf{J}_1$ and $\mathbf{J}_2$ select the first and last $N-1$ components of an $N \times 1$ vector, respectively. Note that

$$\mathbf{\Pi}_{N-1} \mathbf{J}_2 \mathbf{\Pi}_N = \mathbf{J}_1. \quad (11)$$

From (7), we have $\mathbf{A} = \mathbf{Q}_N \mathbf{E}_S \mathbf{T}^{-1}$ which when substituted in (8) yields the relation

$$(\mathbf{J}_1 \mathbf{Q}_N \mathbf{E}_S) \mathbf{\Psi} = \mathbf{J}_2 \mathbf{Q}_N \mathbf{E}_S, \quad \text{where: } \mathbf{\Psi} = \mathbf{T}^{-1} \mathbf{\Phi}_\mu \mathbf{T}. \quad (12)$$

Thus, the eigenvalues of the $d \times d$ solution $\mathbf{\Psi}$ to the above $(N-1) \times d$ matrix equation are $e^{j\mu_i}$, $i = 1, \dots, d$, where $\mu_i = \frac{2\pi}{\lambda} \Delta_x u_i$. At this point, we have an *ESPRIT* based method for estimating the arrival angles of plane waves incident at a ULA for which the first stage of determining signal eigenvectors may be efficiently formulated in terms of real-valued computations. We now show that the second and third stages, computing the solution to $(\mathbf{J}_1 \mathbf{Q}_N \mathbf{E}_S) \mathbf{\Psi} = \mathbf{J}_2 \mathbf{Q}_N \mathbf{E}_S$ and the eigenvalues of $\mathbf{\Psi}$, respectively, may also be efficiently formulated in terms of real-valued computations.

For the second stage, note that $\mathbf{\Pi}_N \mathbf{Q}_N = \mathbf{Q}_N^*$ so that $\mathbf{\Pi}_{N-1} \mathbf{J}_2 \mathbf{Q}_N = \mathbf{\Pi}_{N-1} \mathbf{J}_2 \mathbf{\Pi}_N \mathbf{\Pi}_N \mathbf{Q}_N = \mathbf{J}_1 \mathbf{Q}_N^*$, where we have invoked (11). Since $\mathbf{E}_S$ is real-valued, it follows that the system of equations in (12) may be expressed as

$$\mathbf{C}_1 \mathbf{\Psi} = \mathbf{\Pi}_{N-1} \mathbf{C}_1^*, \quad \text{where: } \mathbf{C}_1 = \mathbf{J}_1 \mathbf{Q}_N \mathbf{E}_S. \quad (13)$$

The TLS ⁴ solution to (13) is $\mathbf{\Psi} = -\mathbf{W}_{12} \mathbf{W}_{22}^{-1}$, where $\begin{bmatrix} \mathbf{W}_{12} \\ \mathbf{W}_{22} \end{bmatrix}$ is a complex-valued $2d \times d$ matrix containing the “smallest” right singular vectors of $[\mathbf{C}_1, \mathbf{\Pi}_{N-1} \mathbf{C}_1^*]$. To reformulate this step in terms of real-valued computations, we exploit the special structure of $[\mathbf{C}_1, \mathbf{\Pi}_{N-1} \mathbf{C}_1^*]$ to convert it to a real-valued matrix of the same dimension through pre- and post-multiplication by the unitary matrices $\mathbf{Q}_{N-1}$ and $\mathbf{M}_{2d}$, respectively, where $\mathbf{M}_{2d}$ is defined by (6) with N_s replaced by d . This yields

$$\mathbf{Z} = \mathbf{Q}_{N-1}^H [\mathbf{C}_1; \mathbf{\Pi}_{N-1} \mathbf{C}_1^*] \mathbf{M}_{2d}. \quad (14)$$

⁴When $\text{range}\{\mathbf{B}\} \subseteq \text{range}\{\mathbf{A}\}$, the TLS solution to $\mathbf{A}\mathbf{X}=\mathbf{B}$ is the same as the LS solution, assuming infinite precision.

The fact that $\mathbf{Z}$ is real-valued is verified by alternatively expressing it as $\mathbf{Z} = \sqrt{2}[\mathcal{R}e\{\mathbf{G}\}, -\mathcal{I}m\{\mathbf{G}\}]$, where $\mathbf{G} = \mathbf{Q}_{N-1}^H \mathbf{C}_1$. It is easily shown that the right singular vectors of $\mathbf{Z}$ are simply related to those of $[\mathbf{C}_1, \mathbf{\Pi}_{N-1} \mathbf{C}_1^*]$ through the unitary transformation $\mathbf{M}_{2d}$. Specifically, if $\begin{bmatrix} \mathbf{V}_{12} \\ \mathbf{V}_{22} \end{bmatrix}$ is a real-valued $2d \times d$ matrix containing the “smallest” right singular vectors of $\mathbf{Z}$, then

$$\begin{bmatrix} \mathbf{W}_{12} \\ \mathbf{W}_{22} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_d & j\mathbf{I}_d \\ \mathbf{I}_d & -j\mathbf{I}_d \end{bmatrix} \begin{bmatrix} \mathbf{V}_{12} \\ \mathbf{V}_{22} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{V}_{12} + j\mathbf{V}_{22} \\ \mathbf{V}_{12} - j\mathbf{V}_{22} \end{bmatrix} \quad (15)$$

This shows how the TLS solution $\Psi = -\mathbf{W}_{12}\mathbf{W}_{22}^{-1}$ may be computed in terms of the right singular vectors of the real-valued matrix $\mathbf{Z}$ in (14).

To formulate the final stage of ESPRIT in terms of real-valued computation, observe that

$$\begin{aligned} \Psi &= -\mathbf{W}_{12}\mathbf{W}_{22}^{-1} \\ &= -(\mathbf{V}_{12} + j\mathbf{V}_{22})(\mathbf{V}_{12} - j\mathbf{V}_{22})^{-1} \\ &= -\left((- \mathbf{V}_{12}\mathbf{V}_{22}^{-1}) - j\mathbf{I}_d\right)\left((- \mathbf{V}_{12}\mathbf{V}_{22}^{-1}) + j\mathbf{I}_d\right)^{-1} \\ &= f(-\mathbf{V}_{12}\mathbf{V}_{22}^{-1}). \end{aligned} \quad (16)$$

where $f(x)$ denotes the linear fractional transformation

$$f(x) = -\frac{x-j}{x+j}. \quad (17)$$

It follows from the Cayley-Hamilton theorem, that if ω is an eigenvalue of the real-valued matrix $-\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$, then $f(\omega) = -(\omega-j)/(\omega+j)$ is an eigenvalue of $-\mathbf{W}_{12}\mathbf{W}_{22}^{-1}$ and the associated eigenvectors are the same. This shows how the desired complex eigenvalues of $\Psi = -\mathbf{W}_{12}\mathbf{W}_{22}^{-1}$ may be determined in terms of the eigenvalues of a real-valued matrix.

Now, asymptotically, the eigenvalues of $\Psi = -\mathbf{W}_{12}\mathbf{W}_{22}^{-1}$ are $e^{j\mu_i}$, $i = 1, \dots, d$. Let ω_i be an eigenvalue of $-\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$. It follows from the above development that $e^{j\mu_i} = -(\omega_i - j)/(\omega_i + j)$. Solving for ω_i yields

$$\omega_i = \frac{1}{j} \frac{e^{j\mu_i} - 1}{e^{j\mu_i} + 1} = \tan\left(\frac{\mu_i}{2}\right). \quad (18)$$

This reveals a spatial frequency warping identical to the temporal frequency warping incurred in designing a digital filter from an analog filter via the bilinear transformation! Consider $d = \lambda/2$ so that $\mu = \frac{2\pi}{\lambda}\Delta_x u = \pi u$. In this case, there is a one-to-one mapping between $-1 < u_i < 1$, corresponding to the range of possible values for a direction cosine, and $-\infty < \omega_i < \infty$. *Unitary ESPRIT* is summarized below.

Summary of Unitary ESPRIT

1. Compute $\mathbf{E}_s$ via the d "largest" left singular vectors of $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_N^H \mathbf{X}$.
2. Compute $\begin{bmatrix} \mathbf{V}_{12} \\ \mathbf{V}_{22} \end{bmatrix}$ via the d "smallest" right singular vectors of $\mathbf{Z} = [\mathcal{R}e\{\mathbf{G}\}, -\mathcal{I}m\{\mathbf{G}\}]$, where $\mathbf{G} = (\mathbf{Q}_{N-1}^H \mathbf{J}_1 \mathbf{Q}_N) \mathbf{E}_s$.
3. Compute $\omega_i, i = 1, \dots, d$, as the eigenvalues of the $d \times d$ real-valued matrix $-\mathbf{V}_{12} \mathbf{V}_{22}^{-1}$.
4. Compute the spatial frequency estimates as $\mu_i = 2 \tan^{-1}(\omega_i), i = 1, \dots, d$.

4 DFT Beamspace ESPRIT for ULA

As an alternative to *Unitary ESPRIT*, we here develop a version of *ESPRIT* for a ULA that works in DFT beamspace. Similar to *Unitary ESPRIT*, and in contrast to the *Beamspace ESPRIT* algorithm of Xu *et al* [16], the algorithm to be developed, referred to as *DFT Beamspace ESPRIT*, involves only real-valued computation from start to finish after the initial transformation to beamspace.

Reduced dimension processing in beamspace is facilitated when one has a-priori information on the general angular locations of the signal arrivals, as in a radar application, for example. In this case, one may only apply those rows of the DFT matrix that form beams encompassing the sector of interest, thereby yielding reduced computational complexity. If there is no a-priori information, one may examine the DFT spectrum and apply the algorithm to be developed to a small set of DFT values around each spectral peak above a particular threshold. In a more general setting, one may simply apply *DFT Beamspace ESPRIT* via parallel processing to each of a number of sets of successive DFT values corresponding to overlapped sectors. Note, though, that in the development to follow, we will employ all N DFT beams for the sake of notational simplicity and so that we can relate *DFT Beamspace ESPRIT* to *Unitary ESPRIT*.

Applying the conjugate centro-symmetrized version of the $m - th$ row of the N pt. DFT matrix

$$\tilde{\mathbf{w}}_m^H = e^{j(\frac{N-1}{2})m\frac{2\pi}{N}} [1, e^{-jm\frac{2\pi}{N}}, e^{-j2m\frac{2\pi}{N}}, \dots, e^{-j(N-1)m\frac{2\pi}{N}}], \quad (19)$$

the $m - th$ component of the DFT beamspace manifold is

$$b_m(\mu) = \tilde{\mathbf{w}}_m^H \mathbf{a}_N(\mu) = \frac{\sin \left[\frac{N}{2} \left(\mu - m \frac{2\pi}{N} \right) \right]}{\sin \left[\frac{1}{2} \left(\mu - m \frac{2\pi}{N} \right) \right]}. \quad (20)$$

Note that we can perform a front end FFT (effectively implementing the Vandermonde form of the rows of the DFT matrix) and achieve conjugate symmetrized beamforming a-posteriori through simple scaling of the DFT values (see (19)). The $N \times 1$ real-valued beamspace manifold is then

$$\mathbf{b}_N(\mu) = \tilde{\mathbf{W}}_N^H \mathbf{a}_N(\mu) = [b_0(\mu), b_1(\mu), \dots, b_{N-1}(\mu)]^T, \quad (21)$$

where $\tilde{\mathbf{W}}_N^H$ denotes the conjugate centro-symmetrized N pt. DFT matrix whose rows are given by (19).

Comparing $b_{m+1}(\mu) = \frac{\sin[\frac{N}{2}(\mu - (m+1)\frac{2\pi}{N})]}{\sin[\frac{1}{2}(\mu - (m+1)\frac{2\pi}{N})]}$ with $b_m(\mu)$ in (20), the numerator of $b_{m+1}(\mu)$ is observed to be the negative of that of $b_m(\mu)$. Thus, two successive components of the beamspace manifold are related as

$$\sin\left[\frac{1}{2}\left(\mu - m\frac{2\pi}{N}\right)\right] b_m(\mu) + \sin\left[\frac{1}{2}\left(\mu - (m+1)\frac{2\pi}{N}\right)\right] b_{m+1}(\mu) = 0. \quad (22)$$

Trigonometric manipulations lead to

$$\tan\left(\frac{\mu}{2}\right) \left\{ \cos\left(m\frac{\pi}{N}\right) b_m(\mu) + \cos\left((m+1)\frac{\pi}{N}\right) b_{m+1}(\mu) \right\} = \sin\left(m\frac{\pi}{N}\right) b_m(\mu) + \sin\left((m+1)\frac{\pi}{N}\right) b_{m+1}(\mu). \quad (23)$$

Compiling all $N - 1$ equations in vector form yields an invariance relationship for the beamspace manifold similar to that for the element space manifold:

$$\tan\left(\frac{\mu}{2}\right) \Gamma_1 \mathbf{b}(\mu) = \Gamma_2 \mathbf{b}(\mu) \quad (24)$$

where

$$\Gamma_1 = \begin{bmatrix} 1 & \cos\left(\frac{\pi}{N}\right) & 0 & \dots & 0 & 0 \\ 0 & \cos\left(\frac{\pi}{N}\right) & \cos\left(\frac{2\pi}{N}\right) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \cos\left((N-2)\frac{\pi}{N}\right) & \cos\left((N-1)\frac{\pi}{N}\right) \end{bmatrix} \in \mathbb{R}^{(N-1) \times N} \quad (25)$$

$$\Gamma_2 = \begin{bmatrix} 0 & \sin\left(\frac{\pi}{N}\right) & 0 & \dots & 0 & 0 \\ 0 & \sin\left(\frac{\pi}{N}\right) & \sin\left(\frac{2\pi}{N}\right) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sin\left((N-2)\frac{\pi}{N}\right) & \sin\left((N-1)\frac{\pi}{N}\right) \end{bmatrix} \in \mathbb{R}^{(N-1) \times N} \quad (26)$$

With d sources, the beamspace DOA matrix is $\mathbf{B} = [\mathbf{b}(\mu_1), \mathbf{b}(\mu_2), \dots, \mathbf{b}(\mu_d)]$. The beamspace manifold relation in (24) translates into the beamspace DOA matrix relation

$$\Gamma_1 \mathbf{B} \Omega_\mu = \Gamma_2 \mathbf{B}, \text{ where: } \Omega_\mu = \text{diag} \left\{ \tan\left(\frac{\mu_1}{2}\right), \dots, \tan\left(\frac{\mu_d}{2}\right) \right\}. \quad (27)$$

Now, the appropriate signal eigenvectors for the algorithm presently under development may be computed as the “largest” left singular vectors of the real-valued matrix $\tilde{\mathbf{W}}_N^H[\mathbf{X}, \Pi_N \mathbf{X}^*]\mathbf{M}_{2N} = \sqrt{2}[\mathcal{R}e\{\mathbf{Y}\}, -\mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \tilde{\mathbf{W}}_N^H \mathbf{X}$. Asymptotically, the $N \times d$ matrix of signal eigenvectors, $\mathbf{E}_S$, satisfies $\mathbf{E}_S = \mathbf{B}\mathbf{T}$, where $\mathbf{T}$ is an unknown $d \times d$ real-valued matrix. Substituting $\mathbf{B} = \mathbf{E}_S \mathbf{T}^{-1}$ into (27) yields

$$\Gamma_1 \mathbf{E}_S \Psi = \Gamma_2 \mathbf{E}_S, \quad \text{where: } \Psi = \mathbf{T}^{-1} \Omega_\mu \mathbf{T}. \quad (28)$$

Thus, the eigenvalues of the $d \times d$ solution Ψ to the $(N-1) \times d$ matrix equation above are $\tan(\mu_i/2)$, $i = 1, \dots, d$. The algorithm based on this development, *DFT Beamspace ESPRIT*, is summarized below.

Summary of DFT Beamspace ESPRIT

1. Compute $\mathbf{E}_S$ via the d “largest” left singular vectors of $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \tilde{\mathbf{W}}_N^H \mathbf{X}$.
2. Compute Ψ as the solution to the $(N-1) \times d$ matrix equation $(\Gamma_1 \mathbf{E}_S) \Psi = (\Gamma_2 \mathbf{E}_S)$.
3. Compute ω_i , $i = 1, \dots, d$, as the eigenvalues of the $d \times d$ real-valued matrix Ψ .
4. Compute spatial frequency estimates as $\mu_i = 2 \tan^{-1}(\omega_i)$, $i = 1, \dots, d$.

4.1 Relationship Between Unitary ESPRIT and DFT Beamspace ESPRIT

To relate *Unitary ESPRIT* and *DFT Beamspace ESPRIT*, consider the following sequence of manipulations:

$$\mathbf{b}_N(\mu) = \tilde{\mathbf{W}}_N^H \mathbf{a}_N(\mu) = \tilde{\mathbf{W}}_N^H \mathbf{Q}_N \mathbf{Q}_N^H \mathbf{a}_N(\mu) = \tilde{\mathbf{W}}_N^H \mathbf{Q}_N \mathbf{d}_N(\mu). \quad (29)$$

Substituting (29) into (24), we find that $\mathbf{d}_N(\mu)$, defined in (5), satisfies a relation similar to (24):

$$\tan\left(\frac{\mu}{2}\right) \Upsilon_1 \mathbf{d}_N(\mu) = \Upsilon_2 \mathbf{d}_N(\mu) \quad (30)$$

where Υ_1 and Υ_2 are the $(N-1) \times d$ real-valued matrices

$$\Upsilon_1 = \Gamma_1 \tilde{\mathbf{W}}_N^H \mathbf{Q}_N \quad \text{and} \quad \Upsilon_2 = \Gamma_2 \tilde{\mathbf{W}}_N^H \mathbf{Q}_N. \quad (31)$$

Thus, the second stage of the *Unitary ESPRIT* algorithm summarized at the end of Section 2 could be alternatively posed as finding Ψ as the solution to the $(N-1) \times d$ matrix equation $(\Upsilon_1 \mathbf{E}_S) \Psi = \Upsilon_2 \mathbf{E}_S$. Employing the TLS method of solution, one would compute $\begin{bmatrix} \mathbf{V}_{12} \\ \mathbf{V}_{22} \end{bmatrix}$ via the

d “smallest” right singular vectors of the real-valued matrix $[\Upsilon_1 \mathbf{E}_S, \Upsilon_2 \mathbf{E}_S]$, and the rest of the algorithm would be the same. Note, though, that Υ_1 and Υ_2 are not sparse like either $\mathbf{J}_1$ and $\mathbf{J}_2$ or Γ_1 and Γ_2 . For example, for $N = 4$ elements,

$$\Upsilon_1 = \begin{bmatrix} 1 & 3 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \Upsilon_2 = \begin{bmatrix} -1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -3 \\ 1 & -1 & 1 & -3 \end{bmatrix}.$$

This concurs with the previous assertion that because there is no physical interpretation of the rows of $\mathbf{Q}_N^H$ in terms of forming beams pointed to different angles, one cannot work with a subset of the rows of $\mathbf{Q}_N^H$.

Again, the utility of *DFT Beamspace ESPRIT* over *Unitary ESPRIT* is in scenarios where one employs a subset of the rows of $\tilde{\mathbf{W}}_N^H$, the number of which depends on the width of the sector of interest and may be substantially less than N , to transform from element space to beamspace. Employing the appropriate subblocks of Γ_1 and Γ_2 as selection matrices, the algorithm is the same as that summarized previously except for the reduced dimensionality. For example, if one employed three successive rows of $\tilde{\mathbf{W}}_N^H$ associated with the DFT bin indices, m , $m+1$, and $m+2$, respectively, to form three beams in estimating the angles of two closely-spaced signal arrivals, as in the low-angle radar tracking scheme described by Zoltowski and Lee [20], the appropriate 3×2 selection matrices are

$$\Gamma_1 = \begin{bmatrix} \cos\left(m\frac{\pi}{N}\right) & \cos\left((m+1)\frac{\pi}{N}\right) & 0 \\ 0 & \cos\left((m+1)\frac{\pi}{N}\right) & \cos\left((m+2)\frac{\pi}{N}\right) \end{bmatrix} \quad \& \quad \Gamma_2 = \begin{bmatrix} \sin\left(m\frac{\pi}{N}\right) & \sin\left((m+1)\frac{\pi}{N}\right) & 0 \\ 0 & \sin\left((m+1)\frac{\pi}{N}\right) & \sin\left((m+2)\frac{\pi}{N}\right) \end{bmatrix}.$$

In this case, one would compute the $d = 2$ “largest” eigenvectors of a 3×3 real-valued matrix, solve a 2×2 real-valued system of equations, and compute the 2 eigenvalues of the resulting 2×2 matrix solution.

4.2 Relationship Between DFT Beamspace ESPRIT and Beamspace ESPRIT

In [16], Xu *et al* develop a beamspace version of ESPRIT that is applicable whenever the $N_b \times N$ beamforming matrix, $\mathbf{F}^H$, exhibits an invariance property similar to that exhibited by the element space DOA matrix in (8). Here N_b denotes the number of beams. That is, if $\mathbf{F}$ satisfies $\mathbf{J}_1 \mathbf{F} \Theta = \mathbf{J}_2 \mathbf{F}$, where Θ is an $N_b \times N_b$ diagonal matrix, then Xu *et al* provide prescriptions for constructing $(N_b - 1) \times N_b$ matrices Σ_1 and Σ_2 satisfying $e^{j\mu} \Sigma_1 \mathbf{b}(\mu) = \Sigma_2 \mathbf{b}(\mu)$, where $\mathbf{b}(\mu)$ is the $N_b \times 1$ beamspace manifold $\mathbf{b}(\mu) = \mathbf{F}^H \mathbf{a}(\mu)$. This facilitates the use of ESPRIT in beamspace ultimately

yielding as eigenvalues the quantities $e^{j\mu_i}$, $i = 1, \dots, d$ as in standard ESPRIT, except via processing in a reduced dimensional space.

Xu *et al* note that a beamforming matrix $\mathbf{F}^H$ composed of N_b rows of the N pt. DFT matrix satisfies a relationship of the form $\mathbf{J}_1 \mathbf{F} \Theta = \mathbf{J}_2 \mathbf{F}$ thereby facilitating the use of *Beamspace ESPRIT*. To see the relationship between *DFT Beamspace ESPRIT* and *Beamspace ESPRIT*, substitute the expression for $\tan(\mu/2)$ in (18) into the invariance relationship for $\mathbf{b}(\mu)$ in (24). This yields, after some manipulation,

$$(e^{j\mu} - 1)\Gamma_1 \mathbf{b}(\mu) = j(e^{j\mu} + 1)\Gamma_2 \mathbf{b}(\mu). \Rightarrow e^{j\mu}(\Gamma_1 - j\Gamma_2)\mathbf{b}(\mu) = (\Gamma_1 + j\Gamma_2)\mathbf{b}(\mu).$$

Thus, in the case where $\mathbf{F}^H$ is composed of conjugate centro-symmetrized rows of the N pt. DFT matrix, the appropriate matrices Σ_1 and Σ_2 required in the execution of *Beamspace ESPRIT* are $\Sigma_1 = \Gamma_1 - j\Gamma_2$ and $\Sigma_2 = \Sigma_1^*$. For this case then, this provides an alternative method for constructing Σ_1 and Σ_2 as opposed to the method prescribed by Xu *et al* in [16] which involves a singular value decomposition.

Note, though, that even if through centro-symmetrization one determines the signal eigenvectors via real-valued computation as discussed previously, the second and third stages of *Beamspace ESPRIT* require complex-valued computation ultimately yielding as eigenvalues $e^{j\mu_i}$, $i = 1, \dots, d$. Aside from the increased computation complexity relative to *DFT Beamspace ESPRIT*, this does not facilitate an extension for the URA yielding automatically paired azimuth and elevation angle estimates.

5 2D Unitary ESPRIT for URA

We now develop an extension of *Unitary ESPRIT* for a uniform rectangular array (URA) of $N \times M$ elements lying in the x - y plane and equi-spaced by Δ_x in the x direction and Δ_y in the y direction. In addition to $\mu = \frac{2\pi}{\lambda} \Delta_x u$, where u is the direction cosine variable relative to the x -axis, we define the spatial frequency variable $\nu = \frac{2\pi}{\lambda} \Delta_y v$, where v is the direction cosine variable relative to the y -axis.

In this development, in addition to representing the array manifold as an $NM \times 1$ vector, denoted $\mathbf{a}(\mu, \nu)$, it will be convenient to represent it as an $N \times M$ matrix, denoted $\mathcal{A}(\mu, \nu)$, as well. The two forms are related through the operators $\text{vec}(\cdot)$ and $\text{mat}(\cdot)$ as $\mathbf{a}(\mu, \nu) = \text{vec}(\mathcal{A}(\mu, \nu))$ and $\mathcal{A}(\mu, \nu) = \text{mat}(\mathbf{a}(\mu, \nu))$. The operator $\text{vec}(\cdot)$ maps an $N \times M$ matrix to an $NM \times 1$ vector by

stacking the columns of the matrix. The operator $\text{mat}(\cdot)$ performs the inverse mapping, mapping an $NM \times 1$ vector into an $N \times M$ matrix such that $\text{mat}(\text{vec}(\mathbf{X})) = \mathbf{X}$. An important property of the vec operator that will prove useful throughout the development is

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}), \quad (32)$$

where $\otimes$ denotes the Kronecker matrix product.

In matrix form, the array manifold may be expressed as

$$\mathcal{A}(\mu, \nu) = \mathbf{a}_N(\mu) \mathbf{a}_M^T(\nu), \quad (33)$$

where $\mathbf{a}_M(\nu)$ is defined by (1) with N replaced by M and μ replaced by ν . Recall that $\mathbf{a}_N(\mu)$ satisfies $e^{j\mu} \mathbf{J}_1 \mathbf{a}_N(\mu) = \mathbf{J}_2 \mathbf{a}_N(\mu)$, where $\mathbf{J}_1$ and $\mathbf{J}_2$ are the $(N-1) \times N$ selection matrices defined in (9) and (10), respectively. It follows that $\mathcal{A}(\mu, \nu)$ in (33) satisfies the invariance relation

$$e^{j\mu} \mathbf{J}_1 \mathcal{A}(\mu, \nu) = \mathbf{J}_2 \mathcal{A}(\mu, \nu). \quad (34)$$

Using the property of the vec operator in (32), we find that the $NM \times 1$ array manifold in vector form satisfies

$$e^{j\mu} \mathbf{J}_{\mu 1} \mathbf{a}(\mu, \nu) = \mathbf{J}_{\mu 2} \mathbf{a}(\mu, \nu) \quad (35)$$

where $\mathbf{J}_{\mu 1}$ and $\mathbf{J}_{\mu 2}$ are the $(N-1)M \times NM$ selection matrices:

$$\mathbf{J}_{\mu 1} = \mathbf{I}_M \otimes \mathbf{J}_1 \quad \text{and} \quad \mathbf{J}_{\mu 2} = \mathbf{I}_M \otimes \mathbf{J}_2. \quad (36)$$

This represents $(N-1)M$ equations obtained by comparing the respective phases of each adjacent pair of elements parallel to the x -axis.

Similarly, to set up the invariance relation relative to the y -axis, observe that

$$e^{j\nu} \mathcal{A}(\mu, \nu) \mathbf{J}_3^T = \mathcal{A}(\mu, \nu) \mathbf{J}_4^T, \quad (37)$$

where the $(M-1) \times M$ matrices $\mathbf{J}_3$ and $\mathbf{J}_4$ select the first and last $M-1$ components of an $M \times 1$ vector, respectively, such that $e^{j\nu} \mathbf{J}_3 \mathbf{a}_M(\nu) = \mathbf{J}_4 \mathbf{a}_M(\nu)$. $\mathbf{J}_3$ and $\mathbf{J}_4$ are defined similar to (9) and (10), except that they are $(M-1) \times M$. Using the property of the vec operator in (32), we find that the $NM \times 1$ array manifold in vector form satisfies the following invariance with respect to ν :

$$e^{j\nu} \mathbf{J}_{\nu 1} \mathbf{a}(\mu, \nu) = \mathbf{J}_{\nu 2} \mathbf{a}(\mu, \nu), \quad (38)$$

where $\mathbf{J}_{\nu 1}$ and $\mathbf{J}_{\nu 2}$ are the $N(M-1) \times NM$ selection matrices:

$$\mathbf{J}_{\nu 1} = \mathbf{J}_3 \otimes \mathbf{I}_N \quad \text{and} \quad \mathbf{J}_{\nu 2} = \mathbf{J}_4 \otimes \mathbf{I}_N. \quad (39)$$

This represents all possible $N(M-1)$ equations obtained by comparing the respective phases of each adjacent pair of elements parallel to the y -axis.

Since $\mathbf{a}_N(\mu)$ and $\mathbf{a}_M(\nu)$ are both conjugate centro-symmetric, $\mathbf{\Pi}_N \mathcal{A}(\mu, \nu) \mathbf{\Pi}_M = \mathcal{A}^*(\mu, \nu)$. Applying the *vec* operator to both sides of this relation and using the property in (32), we obtain $(\mathbf{\Pi}_M \otimes \mathbf{\Pi}_N) \mathbf{a}(\mu, \nu) = \mathbf{a}^*(\mu, \nu)$. Recognizing that $\mathbf{\Pi}_M \otimes \mathbf{\Pi}_N = \mathbf{\Pi}_{NM}$, it follows that $\mathbf{a}(\mu, \nu)$ is conjugate centro-symmetric. We may thus pre-multiply by the sparse unitary matrix $\mathbf{Q}_{MN}$ to obtain the $NM \times 1$ real-valued manifold

$$\mathbf{d}(\mu, \nu) = \mathbf{Q}_{MN}^H \mathbf{a}(\mu, \nu). \quad (40)$$

Let $\mathbf{X}$ be an $NM \times N_s$ matrix composed of N_s snapshots of data as columns. Viewing the array output at a given snapshot as a matrix, we effectively apply the *vec* operator to form an $NM \times 1$ vector and place it as a column of $\mathbf{X}$. Similar to the 1D case, the $NM \times d$ matrix of signal eigenvectors, $\mathbf{E}_S$, may be computed as the “largest” left singular vectors of the real-valued matrix $\mathbf{Q}_{NM}^H [\mathbf{X}, \mathbf{\Pi}_{NM} \mathbf{X}^*] \mathbf{M}_{2N} = \sqrt{2} [\mathcal{R}e\{\mathbf{Y}\}, -\mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_{NM}^H \mathbf{X}$. Asymptotically, $\mathbf{E}_S$ is related to the real-valued $NM \times d$ DOA matrix, $\mathbf{D} = [\mathbf{d}(\mu_1, \nu_1), \mathbf{d}(\mu_2, \nu_2), \dots, \mathbf{d}(\mu_d, \nu_d)]$, as $\mathbf{E}_S = \mathbf{D} \mathbf{T}$, where $\mathbf{T}$ is an unknown $d \times d$ real-valued matrix. Since $\mathbf{Q}_{NM}^H$ is unitary, it follows that asymptotically

$$\mathbf{Q}_{NM} \mathbf{E}_S = \mathbf{A} \mathbf{T}, \quad (41)$$

where $\mathbf{A} = [\mathbf{a}(\mu_1, \nu_1), \mathbf{a}(\mu_2, \nu_2), \dots, \mathbf{a}(\mu_d, \nu_d)]$, the $NM \times d$ complex-valued element space DOA matrix. From (35), it follows that

$$\mathbf{J}_{\mu 1} \mathbf{A} \mathbf{\Phi}_\mu = \mathbf{J}_{\mu 2} \mathbf{A}, \quad \text{where: } \mathbf{\Phi}_\mu = \text{diag}\{e^{j\mu_1}, e^{j\mu_2}, \dots, e^{j\mu_d}\}. \quad (42)$$

Substituting $\mathbf{A} = \mathbf{Q}_{NM} \mathbf{E}_S \mathbf{T}^{-1}$ into (42) yields the relation

$$(\mathbf{J}_{\mu 1} \mathbf{Q}_{NM} \mathbf{E}_S) \mathbf{\Psi}_\mu = \mathbf{J}_{\mu 2} \mathbf{Q}_{NM} \mathbf{E}_S, \quad \text{where: } \mathbf{\Psi}_\mu = \mathbf{T}^{-1} \mathbf{\Phi}_\mu \mathbf{T}. \quad (43)$$

Continuing the development similar to the 1D case, note that $\mathbf{J}_{\mu 1}$ and $\mathbf{J}_{\mu 2}$ satisfy a property similar to (11): $\mathbf{\Pi}_{(N-1)M} \mathbf{J}_{\mu 2} \mathbf{\Pi}_{NM} = \mathbf{J}_{\mu 1}$. Invoking this relationship and the property $\mathbf{\Pi}_{NM} \mathbf{Q}_{NM} =$

$\mathbf{Q}_{NM}^*$, we have $\Pi_{(N-1)M} \mathbf{J}_{\mu 2} \mathbf{Q}_{NM} = \Pi_{(N-1)M} \mathbf{J}_{\mu 2} \Pi_{NM} \Pi_{NM} \mathbf{Q}_{NM} = \mathbf{J}_{\mu 1} \mathbf{Q}_{NM}^*$. Since $\mathbf{E}_S$ is real-valued, it follows that the system of equations in (43) may be expressed as

$$\mathbf{C}_{\mu 1} \Psi_{\mu} = \Pi_{(N-1)M} \mathbf{C}_{\mu 1}^*, \quad \text{where: } \mathbf{C}_{\mu 1} = \mathbf{J}_{\mu 1} \mathbf{Q}_{NM} \mathbf{E}_S. \quad (44)$$

Let $\begin{bmatrix} \mathbf{U}_{12} \\ \mathbf{U}_{22} \end{bmatrix}$ be the $2d \times d$ matrix containing the “smallest” right⁵ singular vectors of the real-valued matrix

$$\begin{aligned} \mathbf{Z}_{\mu} &= \mathbf{Q}_{(N-1)M}^H [\mathbf{C}_{\mu 1} : \Pi_{(N-1)M} \mathbf{C}_{\mu 1}^*] \mathbf{M}_{2d} \\ &= \sqrt{2} [\mathcal{R}e\{\mathbf{G}_{\mu}\} : -\mathcal{I}m\{\mathbf{G}_{\mu}\}], \quad \text{where: } \mathbf{G}_{\mu} = (\mathbf{Q}_{(N-1)M}^H \mathbf{J}_{\mu 1} \mathbf{Q}_{NM}) \mathbf{E}_S. \end{aligned} \quad (45)$$

It follows from previous developments that the $d \times d$ real-valued matrix $-\mathbf{U}_{12} \mathbf{U}_{22}^{-1}$ may be spectrally decomposed as

$$-\mathbf{U}_{12} \mathbf{U}_{22}^{-1} = \mathbf{T}^{-1} \Omega_{\mu} \mathbf{T}, \quad \text{where: } \Omega_{\mu} = \text{diag} \left\{ \tan \left(\frac{\mu_1}{2} \right), \dots, \tan \left(\frac{\mu_d}{2} \right) \right\}. \quad (46)$$

A similar development relative to estimating $\nu_i, i = 1, \dots, d$, ultimately yields the following result. Let $\begin{bmatrix} \mathbf{V}_{12} \\ \mathbf{V}_{22} \end{bmatrix}$ denote the $2d \times d$ matrix containing the “smallest” right singular vectors of the real-valued matrix

$$\begin{aligned} \mathbf{Z}_{\nu} &= \mathbf{Q}_{N(M-1)}^H [\mathbf{C}_{\nu 1} : \Pi_{(N-1)M} \mathbf{C}_{\nu 1}^*] \mathbf{M}_{2d} \\ &= \sqrt{2} [\mathcal{R}e\{\mathbf{G}_{\nu}\} : -\mathcal{I}m\{\mathbf{G}_{\nu}\}], \quad \text{where: } \mathbf{G}_{\nu} = (\mathbf{Q}_{N(M-1)}^H \mathbf{J}_{\nu 1} \mathbf{Q}_{NM}) \mathbf{E}_S \end{aligned} \quad (47)$$

and $\mathbf{C}_{\nu 1} = \mathbf{J}_{\nu 1} \mathbf{Q}_{NM} \mathbf{E}_S$. The $d \times d$ real-valued matrix $-\mathbf{V}_{12} \mathbf{V}_{22}^{-1}$ may be spectrally decomposed as

$$-\mathbf{V}_{12} \mathbf{V}_{22}^{-1} = \mathbf{T}^{-1} \Omega_{\nu} \mathbf{T}, \quad \text{where: } \Omega_{\nu} = \text{diag} \left\{ \tan \left(\frac{\nu_1}{2} \right), \dots, \tan \left(\frac{\nu_d}{2} \right) \right\}. \quad (48)$$

Now, to achieve automatic pairing of μ and ν spatial frequencies, the following critical observations are made. First, the $d \times d$ matrix of eigenvectors $\mathbf{T}$ in the spectral decomposition of $-\mathbf{U}_{12} \mathbf{U}_{22}^{-1}$ in (46) is the same as that appearing in the spectral decomposition of $-\mathbf{V}_{12} \mathbf{V}_{22}^{-1}$ in (48). Second, this is the same real-valued matrix $\mathbf{T}$ appearing in (41) which is unique as long as no two sources have exactly the same azimuth and elevation angles. Finally, $-\mathbf{U}_{12} \mathbf{U}_{22}^{-1}$ and $-\mathbf{V}_{12} \mathbf{V}_{22}^{-1}$ are real-valued, as are the diagonal matrices Ω_{μ} and Ω_{ν} . These observations lead to the main result, namely

$$-\mathbf{U}_{12} \mathbf{U}_{22}^{-1} + j(-\mathbf{V}_{12} \mathbf{V}_{22}^{-1}) = \mathbf{T}^{-1} \{\Omega_{\mu} + j\Omega_{\nu}\} \mathbf{T}. \quad (49)$$

⁵We depart from the convention of using $\mathbf{U}$ to denote the matrix of left singular vectors here since the right singular vectors of $\mathbf{Z}_{\mu}$ are associated with the estimation of $u_i, i = 1, \dots, d$. $\mathbf{V}$ is used to denote the matrix of right singular vectors of $\mathbf{Z}_{\nu}$ since these are associated with the estimation of $v_i, i = 1, \dots, d$.

Thus, the eigenvalues of $-\mathbf{U}_{12}\mathbf{U}_{22}^{-1} + j(-\mathbf{V}_{12}\mathbf{V}_{22}^{-1})$ are $\tan(\mu_i/2) + j\tan(\nu_i/2)$, $i = 1, \dots, d$. The algorithm based on this development is referred to as *2D Unitary ESPRIT* and is summarized below.

Summary of 2D Unitary ESPRIT

1. Compute $\mathbf{E}_s$ via the d "largest" left singular vectors of $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_{NM}^H \mathbf{X}$.
2. Compute $\begin{bmatrix} \mathbf{U}_{12} \\ \mathbf{U}_{22} \end{bmatrix}$ via the d "smallest" right singular vectors of $\mathbf{Z}_\mu = [\mathcal{R}e\{\mathbf{G}_\mu\}, -\mathcal{I}m\{\mathbf{G}_\mu\}]$, where $\mathbf{G}_\mu = (\mathbf{Q}_{(N-1)M}^H \mathbf{J}_{\mu 1} \mathbf{Q}_{NM}) \mathbf{E}_s$.
3. Compute $\begin{bmatrix} \mathbf{V}_{12} \\ \mathbf{V}_{22} \end{bmatrix}$ via the d "smallest" right singular vectors of $\mathbf{Z}_\nu = [\mathcal{R}e\{\mathbf{G}_\nu\}, -\mathcal{I}m\{\mathbf{G}_\nu\}]$, where $\mathbf{G}_\nu = (\mathbf{Q}_{N(M-1)}^H \mathbf{J}_{\nu 1} \mathbf{Q}_{NM}) \mathbf{E}_s$.
4. Compute λ_i , $i = 1, \dots, d$, as the eigenvalues of the $d \times d$ matrix $-\mathbf{U}_{12}\mathbf{U}_{22}^{-1} + j(-\mathbf{V}_{12}\mathbf{V}_{22}^{-1})$.
5. Compute spatial frequency estimates: $\mu_i = 2 \tan^{-1}(\mathcal{R}e\{\lambda_i\})$, $\nu_i = 2 \tan^{-1}(\mathcal{I}m\{\lambda_i\})$, $i = 1, \dots, d$.

Note that the maximum number of sources *2D Unitary ESPRIT* can handle is $\min\{M(N-1), N(M-1)\}$, assuming that at least $d+1$ snapshots are available. If only a single snapshot is available, one can extract $d+1$ or more identical rectangular subarrays out of the overall array to get the effect of multiple snapshots, thereby decreasing the maximum number of sources that can be handled.

5.1 2D Unitary ESPRIT vs. ACMP

Note that *2D Unitary ESPRIT* provides closed-form, automatically paired 2D angle estimates as long as the spatial frequency coordinate pairs (μ_i, ν_i) , $i = 1, \dots, d$, are distinct. That is, no additional effort is needed if a pair or more of sources have the same μ_i or ν_i . This is in contrast to the Algebraically Coupled Matrix Pencil (*ACMP*) method of van der Veen *et al* which also provides closed-form, automatically paired 2D angle estimates but breaks down if two sources have either the same μ or ν spatial frequency coordinate. Note that in order to avoid the same problem as *ACMP* in this regard, one must solve the complex eigenvalue problem signified by (49). If one attempts to compute the real eigenvalues of $-\mathbf{U}_{12}\mathbf{U}_{22}^{-1}$ alone, for example, there is a degeneracy in the eigenvectors when two sources have the same μ spatial frequency coordinate thereby precluding the ability to determine $\mathbf{T}$.

Note that Vanpoucke *et al* propose a form of subarray averaging to overcome the problem of *ACMP* occurring when two sources have either the same μ or ν spatial frequency coordinate, but this decreases the maximum number of sources that can be handled and increases the computational complexity significantly.

Note that *ACMP* requires an array of sensor triplets so that one can extract three identical subarrays from the overall array. *2D Unitary ESPRIT* only requires that the array exhibit invariances in two distinct directions, as would be the case with two uniform linear arrays (ULA's), for example. In Section 7, we show how *2D Unitary ESPRIT* may be simply adapted for the case of two orthogonal ULA's having a common phase center. *ACMP* is not applicable with such an array geometry. Another advantage of *2D Unitary ESPRIT* over *ACMP* is that *2D Unitary ESPRIT* is efficiently formulated in terms of real-valued computations, except for the final $d \times d$ eigenvalue decomposition, while *ACMP* requires complex-valued computations throughout.

6 2D DFT Beamspace ESPRIT for URA

With 2D DFT beamforming (and attendant conjugate centro-symmetrization through simple scaling), the components of the beamspace array manifold are separable real-valued patterns of the form

$$b_{m,n}(\mu, \nu) = \frac{\sin \left[\frac{N}{2} \left(\mu - m \frac{2\pi}{N} \right) \right]}{\sin \left[\frac{1}{2} \left(\mu - m \frac{2\pi}{N} \right) \right]} \frac{\sin \left[\frac{M}{2} \left(\nu - n \frac{2\pi}{M} \right) \right]}{\sin \left[\frac{1}{2} \left(\nu - n \frac{2\pi}{M} \right) \right]}. \quad (50)$$

Note that the matrix form of the beamspace manifold, denoted $\mathcal{B}(\mu, \nu)$, is related to the matrix form of the array manifold via a 2D DFT as $\mathcal{B}(\mu, \nu) = \tilde{\mathbf{W}}_N^H \mathcal{A}(\mu, \nu) \tilde{\mathbf{W}}_M$, where $\tilde{\mathbf{W}}_N^H$ denotes the conjugate centro-symmetrized N pt. DFT matrix whose rows are given by (19) and $\tilde{\mathbf{W}}_M^H$ is defined similarly with N replaced by M . Substituting the form of $\mathcal{A}(\mu, \nu)$ in (33) into $\mathcal{B}(\mu, \nu) = \tilde{\mathbf{W}}_N^H \mathcal{A}(\mu, \nu) \tilde{\mathbf{W}}_M$ yields

$$\mathcal{B}(\mu, \nu) = \mathbf{b}_N(\mu) \mathbf{b}_M^T(\nu), \quad (51)$$

where $\mathbf{b}_N(\mu)$ is defined in (21) and $\mathbf{b}_M(\nu)$ is defined similarly with N replaced by M and μ replaced by ν . Given that $\mathbf{b}_N(\mu)$ satisfies the invariance relationship in (24), it follows that $\mathcal{B}(\mu, \nu)$ satisfies

$$\tan \left(\frac{\mu}{2} \right) \Gamma_1 \mathcal{B}(\mu, \nu) = \Gamma_2 \mathcal{B}(\mu, \nu). \quad (52)$$

where Γ_1 and Γ_2 are defined in (25) and (26). Using the property of the *vec* operator in (32), we find that the $NM \times 1$ beamspace manifold in vector form, $\mathbf{b}(\mu, \nu) = \text{vec}[B(\mu, \nu)]$, satisfies

$$\tan\left(\frac{\mu}{2}\right) \Gamma_{\mu 1} \mathbf{b}(\mu, \nu) = \Gamma_{\mu 2} \mathbf{b}(\mu, \nu), \quad (53)$$

where $\Gamma_{\mu 1}$ and $\Gamma_{\mu 2}$ are the $(N-1)M \times NM$ matrices:

$$\Gamma_{\mu 1} = \mathbf{I}_M \otimes \Gamma_1 \quad \text{and} \quad \Gamma_{\mu 2} = \mathbf{I}_M \otimes \Gamma_2. \quad (54)$$

(53) represents $(N-1)M$ equations obtained by comparing each pair of adjacent beams having the same μ pointing angle coordinate.

Similarly, the 1D beamspace manifold $\mathbf{b}_M(\nu)$ satisfies $\tan(\nu/2) \Gamma_3 \mathbf{b}_M(\nu) = \Gamma_4 \mathbf{b}_M(\nu)$, where Γ_3 and Γ_4 are defined similar to (25) and (26) with N replaced by M such that they are $(M-1) \times M$. It follows that

$$\tan\left(\frac{\nu}{2}\right) B(\mu, \nu) \Gamma_3^T = B(\mu, \nu) \Gamma_4^T. \quad (55)$$

Again, using the *vec* operator, we find that $\mathbf{b}(\mu, \nu)$ satisfies

$$\tan\left(\frac{\nu}{2}\right) \Gamma_{\nu 1} \mathbf{b}(\mu, \nu) = \Gamma_{\nu 2} \mathbf{b}(\mu, \nu), \quad (56)$$

where $\Gamma_{\nu 1}$ and $\Gamma_{\nu 2}$ are the $N(M-1) \times NM$ matrices:

$$\Gamma_{\nu 1} = \Gamma_3 \otimes \mathbf{I}_N \quad \text{and} \quad \Gamma_{\nu 2} = \Gamma_4 \otimes \mathbf{I}_N. \quad (57)$$

(56) represents $N(M-1)$ equations obtained by comparing each pair of adjacent beams having the same ν pointing angle coordinate.

Consider the $NM \times d$ real-valued beamspace DOA matrix $\mathbf{B} = [\mathbf{b}(\mu_1, \nu_1), \dots, \mathbf{b}(\mu_d, \nu_d)]$. (53) dictates that $\mathbf{B}$ satisfies

$$\Gamma_{\mu 1} \mathbf{B} \Omega_\mu = \Gamma_{\mu 2} \mathbf{B} \quad (58)$$

where Ω_μ is defined in (46). In turn, (56) dictates that $\mathbf{B}$ satisfies

$$\Gamma_{\nu 1} \mathbf{B} \Omega_\nu = \Gamma_{\nu 2} \mathbf{B} \quad (59)$$

where Ω_ν is defined in (48).

Now, viewing the array output at a given snapshot as an $N \times M$ matrix, we compute a 2D DFT, apply the *vec* operator, and place the resulting $NM \times 1$ vector as a column of an $NM \times N_s$ data matrix $\mathbf{Y}$. Recall that $\mathbf{X}$ denotes the $NM \times N_s$ data matrix prior to the 2D DFT. Using

the *vec* operator, the relationship between $\mathbf{Y}$ and $\mathbf{X}$ may be expressed as $\mathbf{Y} = (\tilde{\mathbf{W}}_M^T \otimes \tilde{\mathbf{W}}_N^H) \mathbf{X}$. The appropriate $NM \times d$ matrix of signal eigenvectors, $\mathbf{E}_S$, for the algorithm presently under development may be computed as the d “largest” left singular vectors of the real-valued matrix $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$. Asymptotically, $\mathbf{E}_S = \mathbf{B}\mathbf{T}$, where $\mathbf{T}$ is an unknown $d \times d$ *real-valued* matrix. Substituting $\mathbf{B} = \mathbf{E}_S \mathbf{T}^{-1}$ into (58) and (59) yields the signal eigenvector relations

$$\Gamma_{\mu 1} \mathbf{E}_S \Psi_\mu = \Gamma_{\mu 2} \mathbf{E}_S \quad \text{where:} \quad \Psi_\mu = \mathbf{T}^{-1} \Omega_\mu \mathbf{T} \quad (60)$$

$$\Gamma_{\nu 1} \mathbf{E}_S \Psi_\nu = \Gamma_{\nu 2} \mathbf{E}_S \quad \text{where:} \quad \Psi_\nu = \mathbf{T}^{-1} \Omega_\nu \mathbf{T}. \quad (61)$$

As in the extension of *Unitary ESPRIT* for a URA, automatic pairing of μ and ν spatial frequency estimates is facilitated by the fact that all of the quantities in (60) and (61) are real-valued. Thus, $\Psi_\mu + j\Psi_\nu$ may be spectrally decomposed as

$$\Psi_\mu + j\Psi_\nu = \mathbf{T}^{-1} \{\Omega_\mu + j\Omega_\nu\} \mathbf{T} \quad (62)$$

The algorithm based on this development, *2D DFT BeamSpace ESPRIT*, is summarized below.

Summary of 2D DFT BeamSpace ESPRIT

1. Compute a 2D DFT of the $N \times M$ matrix of array outputs at each snapshot (scale for conjugate centro-symmetrization), apply the *vec* operator, and place the result as a column of $\mathbf{Y}$.
2. Compute $\mathbf{E}_s$ via the d “largest” left singular vectors of $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$.
3. Compute Ψ_μ as the solution to the $(N-1)M \times d$ matrix equation $\Gamma_{\mu 1} \mathbf{E}_S \Psi_\mu = \Gamma_{\mu 2} \mathbf{E}_S$.
4. Compute Ψ_ν as the solution to the $N(M-1) \times d$ matrix equation $\Gamma_{\nu 1} \mathbf{E}_S \Psi_\nu = \Gamma_{\nu 2} \mathbf{E}_S$.
5. Compute λ_i , $i = 1, \dots, d$, as the eigenvalues of the $d \times d$ matrix $\Psi_\mu + j\Psi_\nu$.
6. Compute spatial frequency estimates: $\mu_i = 2 \tan^{-1}(\mathcal{R}e\{\lambda_i\})$, $\nu_i = 2 \tan^{-1}(\mathcal{I}m\{\lambda_i\})$, $i = 1, \dots, d$.

6.1 Reduced Dimension Example

As in the 1D case, the utility of *2D DFT BeamSpace ESPRIT* over *2D Unitary ESPRIT* is in scenarios where one works with a subset of 2D DFT beams that encompass some volume of space of interest. In fact, the ability to work in a reduced dimension beamspace is even of more value in the case of a URA since the total number of elements may be quite high. As an example, consider a scenario, similar to the low-angle radar tracking problem, in which we desire to estimate the

respective azimuth and elevation angles of each of two closely-spaced sources. To this end, we form four 2D DFT beams steered to the spatial frequency coordinate pairs $(m\frac{2\pi}{N}, n\frac{2\pi}{M})$, $((m+1)\frac{2\pi}{N}, n\frac{2\pi}{M})$, $(m\frac{2\pi}{N}, (n+1)\frac{2\pi}{M})$, and $((m+1)\frac{2\pi}{N}, (n+1)\frac{2\pi}{M})$, respectively, as depicted in Figure 3. Recalling that the components of the beamspace manifold have the form in (50), the 4×1 beamspace manifold for this case is

$$\mathbf{b}(\mu, \nu) = [b_{m,n}(\mu, \nu), b_{m+1,n}(\mu, \nu), b_{m,n+1}(\mu, \nu), b_{m+1,n+1}(\mu, \nu)]^T. \quad (63)$$

In this case, $\mathbf{E}_S$ is 4×2 and may be constructed from the two “largest” eigenvectors of the real part of the 4×4 matrix formed from the inter-beam correlations. The 2×2 matrices Ψ_μ and Ψ_ν would be computed as the corresponding solutions to the 4×2 respective matrix equations $\Gamma_{\mu 1} \mathbf{E}_S \Psi_\mu = \Gamma_{\mu 2} \mathbf{E}_S$ and $\Gamma_{\nu 1} \mathbf{E}_S \Psi_\nu = \Gamma_{\nu 2} \mathbf{E}_S$, where

$$\begin{aligned} \Gamma_{\mu 1} &= \begin{bmatrix} \cos\left(m\frac{\pi}{N}\right) & \cos\left((m+1)\frac{\pi}{N}\right) & 0 & 0 \\ 0 & 0 & \cos\left(m\frac{\pi}{N}\right) & \cos\left((m+1)\frac{\pi}{N}\right) \end{bmatrix} \\ \Gamma_{\mu 2} &= \begin{bmatrix} \sin\left(m\frac{\pi}{N}\right) & \sin\left((m+1)\frac{\pi}{N}\right) & 0 & 0 \\ 0 & 0 & \sin\left(m\frac{\pi}{N}\right) & \sin\left((m+1)\frac{\pi}{N}\right) \end{bmatrix} \\ \Gamma_{\nu 1} &= \begin{bmatrix} \cos\left(n\frac{\pi}{M}\right) & 0 & \cos\left((n+1)\frac{\pi}{M}\right) & 0 \\ 0 & \cos\left(n\frac{\pi}{M}\right) & 0 & \cos\left((n+1)\frac{\pi}{M}\right) \end{bmatrix} \\ \Gamma_{\nu 2} &= \begin{bmatrix} \sin\left(n\frac{\pi}{M}\right) & 0 & \sin\left((n+1)\frac{\pi}{M}\right) & 0 \\ 0 & \sin\left(n\frac{\pi}{M}\right) & 0 & \sin\left((n+1)\frac{\pi}{M}\right) \end{bmatrix}. \end{aligned}$$

In the final stage of the algorithm, $\tan(\mu_i/2) + j \tan(\nu_i/2)$, $i = 1, 2$, would be computed as the eigenvalues of a 2×2 matrix.

6.2 Comparison with UCA-ESPRIT

As discussed in Section 1, *UCA-ESPRIT* [7, 8] is a recently developed closed-form 2D angle estimation scheme for a uniform circular array (UCA). As indicated in Figure 2, in the final stage of *UCA-ESPRIT*, the i -th eigenvalue of a matrix has the form $u_i + jv_i$, where u_i and v_i are the direction cosines of the i -th source relative to the x and y axes, respectively, assuming the UCA to lie in the x - y plane. This is in contrast to *2D DFT Beamspace ESPRIT* where there is spatial frequency warping such that the final eigenvalues are of the form $\tan(\mu_i/2) + j \tan(\nu_i/2)$, $i = 1, \dots, d$. A notable difference between the development of *UCA-ESPRIT* and that of *2D DFT Beamspace ESPRIT* is that in the former the sampled aperture pattern was assumed to be approximately equal to the

continuous aperture pattern [7, 8], while no such approximation was made in the latter case. We here briefly show that if a similar approximation is made in the development of *2D DFT Beamspace ESPRIT*, the final eigenvalues yielded by the resulting approximate *2D DFT Beamspace ESPRIT* algorithm are identical in form to those yielded by *UCA-ESPRIT*.

Aside from averting spatial frequency warping, this form of the eigenvalue has a nice geometrical interpretation in that it may be expressed as $u_i + jv_i = \sin \theta_i e^{j\phi_i}$, where ϕ_i and θ_i are the azimuth and elevation angles of the i -th source, respectively. This is illustrated in Figure 2. θ_i varies between 0° and 90° so that $\sin \theta_i$ varies between 0 and 1, while ϕ_i varies between 0° and 360° . Thus, one can immediately glean the azimuth angle of the i -th source from the polar angle of the i -th eigenvalue. The corresponding elevation angle is the arcsine of the magnitude of the i -th eigenvalue. If the eigenvalue is at the origin, the source is at boresite. If the eigenvalue is on the unit circle, the source is in the same plane as the array. Also, we may use the fact that an eigenvalue should be located on or within the unit circle to screen out false alarms.

Assume the interelement spacing in either direction to be less than or equal to a half-wavelength. In this case, in the vicinity of the mainlobe and first few sidelobes, $b_{m,n}(\mu, \nu) \approx \frac{\sin[\frac{N}{2}(\mu - m\frac{2\pi}{N})]}{\frac{1}{2}(\mu - m\frac{2\pi}{N})} \frac{\sin[\frac{M}{2}(\nu - n\frac{2\pi}{M})]}{\frac{1}{2}(\nu - n\frac{2\pi}{M})}$. Substituting $\mu = \frac{2\pi}{\lambda}\Delta_x u$ and $\nu = \frac{2\pi}{\lambda}\Delta_y v$, define

$$b_{m,n}^a(u, v) = \frac{\sin\left[\frac{N}{2}\left(\frac{2\pi}{\lambda}\Delta_x u - m\frac{2\pi}{N}\right)\right]}{\frac{1}{2}\left(\frac{2\pi}{\lambda}\Delta_x u - m\frac{2\pi}{N}\right)} \frac{\sin\left[\frac{M}{2}\left(\frac{2\pi}{\lambda}\Delta_y v - n\frac{2\pi}{M}\right)\right]}{\frac{1}{2}\left(\frac{2\pi}{\lambda}\Delta_y v - n\frac{2\pi}{M}\right)}. \quad (64)$$

This is the far field pattern that would result with a continuous rectangular aperture of dimension $N\Delta_x$ by $M\Delta_y$. The superscript a denotes approximate pattern. Similar to the development for the sampled aperture pattern, observe that $b_{m,n}^a(u, v)$ and $b_{m+1,n}^a(u, v)$ are related as

$$\left(\frac{2\pi}{\lambda}\Delta_x u - m\frac{2\pi}{N}\right) b_{m,n}^a(u, v) + \left(\frac{2\pi}{\lambda}\Delta_x u - (m+1)\frac{2\pi}{N}\right) b_{m+1,n}^a(u, v) = 0, \quad (65)$$

which may be rearranged as

$$u \{b_{m,n}^a(u, v) + b_{m+1,n}^a(u, v)\} = \frac{\lambda}{N\Delta_x} \{mb_{m,n}^a(u, v) + (m+1)b_{m+1,n}^a(u, v)\}. \quad (66)$$

Similarly, $b_{m,n}^a(u, v)$ and $b_{m,n+1}^a(u, v)$ are related as

$$\left(\frac{2\pi}{\lambda}\Delta_y v - n\frac{2\pi}{M}\right) b_{m,n}^a(u, v) + \left(\frac{2\pi}{\lambda}\Delta_y v - (n+1)\frac{2\pi}{M}\right) b_{m,n+1}^a(u, v) = 0, \quad (67)$$

which may be rearranged as

$$v \{b_{m,n}^a(u, v) + b_{m,n+1}^a(u, v)\} = \frac{\lambda}{M\Delta_y} \{nb_{m,n}^a(u, v) + (n+1)b_{m,n+1}^a(u, v)\}. \quad (68)$$

For the sake of brevity, consider again the case of four 2D DFT beams to estimate the respective azimuth and elevation angles of each of two closely-spaced sources. In this case, the 4×1 beamspace manifold is $\mathbf{b}^a(u, v) = [b_{m,n}^a(u, v), b_{m+1,n}^a(u, v), b_{m,n+1}^a(u, v), b_{m+1,n+1}^a(u, v)]^T$. Given the relations above, it is readily deduced that $u\Gamma_{u1}^a \mathbf{b}^a(u, v) = \Gamma_{u2}^a \mathbf{b}^a(u, v)$ and $v\Gamma_{v1}^a \mathbf{b}^a(u, v) = \Gamma_{v2}^a \mathbf{b}^a(u, v)$, where

$$\Gamma_{u1}^a = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \Gamma_{u2}^a = \frac{\lambda}{N\Delta_x} \begin{bmatrix} m & (m+1) & 0 & 0 \\ 0 & 0 & m & (m+1) \end{bmatrix}$$

$$\Gamma_{v1}^a = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \Gamma_{v2}^a = \frac{\lambda}{M\Delta_y} \begin{bmatrix} n & 0 & (n+1) & 0 \\ 0 & n & 0 & (n+1) \end{bmatrix}.$$

Asymptotically, the 4×2 real-valued matrix of signal eigenvectors, $\mathbf{E}_S$, satisfies $\mathbf{E}_S = \mathbf{B}\mathbf{T}$, where $\mathbf{B} = [\mathbf{b}(u_1, v_1), \mathbf{b}(u_2, v_2)]$ and $\mathbf{T}$ is an unknown 2×2 *real-valued* matrix. Expediting the development, it follows that $\Gamma_{u1}^a \mathbf{E}_S \Psi_u = \Gamma_{u2}^a \mathbf{E}_S$, where $\Psi_u = \mathbf{T}^{-1} \Omega_u \mathbf{T}$ and $\Omega_u = \text{diag}\{u_1, u_2\}$. Also, $\Gamma_{v1}^a \mathbf{E}_S \Psi_v = \Gamma_{v2}^a \mathbf{E}_S$, where $\Psi_v = \mathbf{T}^{-1} \Omega_v \mathbf{T}$ and $\Omega_v = \text{diag}\{v_1, v_2\}$. Thus, $u_1 + jv_1$ and $u_2 + jv_2$ are the two eigenvalues of $\Psi_u + j\Psi_v$.

The point is that with $d \leq \lambda/2$ the sampled aperture pattern is very well approximated by the continuous aperture pattern in the vicinity of the mainlobe and first few sidelobes. Thus, if only a relatively small number of beams is selected, the modified version of *2D DFT Beamspace ESPRIT* sketched above yields the direction cosines directly without spatial warping.

7 2D DFT Beamspace ESPRIT for Cross Array

Consider an array composed of an N element ULA aligned with the x-axis and an M element ULA aligned with the y-axis. The center of each leg is assumed to be at the origin so that they have a common phase center. To ease the development and for the sake of notational simplicity, we will assume M and N are both even so that the two legs do not share a common element at the origin. However, with slight modification, the adaptation of *2D DFT Beamspace ESPRIT* for a cross array developed subsequently may also be employed when M and/or N are odd. Also, due to space limitations, we here only present the appropriate adaptation of *2D DFT Beamspace ESPRIT*. *2D Unitary ESPRIT* may also be suitably adapted but would require a slightly more complicated development.

Let $\mathbf{x}(\ell)$ and $\mathbf{y}(\ell)$ be the $N \times 1$ and $M \times 1$ snapshot vectors output by the two respective legs at time ℓ . The $(N + M) \times 1$ composite snapshot vector is formed as $\mathbf{z}(\ell) = \begin{bmatrix} \mathbf{x}(\ell) \\ \mathbf{y}(\ell) \end{bmatrix}$. These are

stacked as the columns of an $(N + M) \times N_s$ matrix $\mathbf{Z}$. The array manifold for such an array is

$$\mathbf{a}(\mu, \nu) = \begin{bmatrix} \mathbf{a}_N(\mu) \\ \mathbf{a}_M(\nu) \end{bmatrix}, \quad (69)$$

where $\mathbf{a}_N(\mu)$ and $\mathbf{a}_M(\nu)$ are each conjugate centro-symmetric as defined previously. Note that it is only because the two legs have a common phase center that we are able to express the array manifold in this form. If this is not the case, as with an L-shaped array, for example, either the upper $N \times 1$ or lower $M \times 1$ block of $\mathbf{a}(\mu, \nu)$ would not be conjugate centro-symmetric and it would not be possible to convert $\mathbf{a}(\mu, \nu)$ to a real-valued manifold through a simple matrix transformation.

Transformation to beamspace is accomplished via

$$\mathbf{F} = \begin{bmatrix} \tilde{\mathbf{W}}_N & \mathbf{O} \\ \mathbf{O} & \tilde{\mathbf{W}}_M \end{bmatrix}. \quad (70)$$

The beamspace manifold is

$$\mathbf{b}(\mu, \nu) = \mathbf{F}^H \mathbf{a}(\mu, \nu) = \begin{bmatrix} \mathbf{b}_N(\mu) \\ \mathbf{b}_M(\nu) \end{bmatrix}, \quad (71)$$

where $\mathbf{b}_N(\mu)$ and $\mathbf{b}_M(\nu)$ are as defined previously. In practice, transformation to beamspace is accomplished via an N pt. DFT of the x -axis leg and an M pt. DFT of the y -axis leg, with a-posteriori conjugate centro-symmetrization via simple scaling of each DFT value.

Let $\mathbf{E}_S$ be the $(N + M) \times d$ matrix of signal eigenvectors computed as the d "largest" left singular vectors of $[\mathcal{R}e\{\mathbf{H}\}, \mathcal{I}m\{\mathbf{H}\}]$, where $\mathbf{H} = \mathbf{F}^H \mathbf{Z}$. (Alternatively, $\mathbf{E}_S$ may be determined as the d "largest" eigenvectors of $\mathcal{R}e\{\mathbf{F}^H \mathbf{Z} \mathbf{Z}^H \mathbf{F}\}$.) Asymptotically, $\mathbf{E}_S = \mathbf{B} \mathbf{T}$, where $\mathbf{B} = [\mathbf{b}(\mu_1, \nu_1), \dots, \mathbf{b}(\mu_d, \nu_d)]$ and $\mathbf{T}$ is an unknown $d \times d$ real-valued matrix. Define the following matrices:

$$\Delta_{\mu 1} = \left[\underbrace{\Gamma_1}_N \vdots \underbrace{\mathbf{O}}_M \right]^{N-1} \quad \text{and} \quad \Delta_{\mu 2} = \left[\underbrace{\Gamma_2}_N \vdots \underbrace{\mathbf{O}}_M \right]^{N-1} \quad (72)$$

$$\Delta_{\nu 1} = \left[\underbrace{\mathbf{O}}_N \vdots \underbrace{\Gamma_3}_M \right]^{M-1} \quad \text{and} \quad \Delta_{\nu 2} = \left[\underbrace{\mathbf{O}}_N \vdots \underbrace{\Gamma_4}_M \right]^{M-1} \quad (73)$$

where Γ_3 and Γ_4 are defined similar to (25) and (26) with N replaced by M . The following signal eigenvector relations follow quite readily from previous developments:

$$\Delta_{\mu 1} \mathbf{E}_S \Psi_\mu = \Delta_{\mu 2} \mathbf{E}_S \quad \text{where:} \quad \Psi_\mu = \mathbf{T}^{-1} \Omega_\mu \mathbf{T} \quad (74)$$

$$\Delta_{\nu 1} \mathbf{E}_S \Psi_\nu = \Delta_{\nu 2} \mathbf{E}_S \quad \text{where:} \quad \Psi_\nu = \mathbf{T}^{-1} \Omega_\nu \mathbf{T}. \quad (75)$$

As with *2D DFT Beamspace ESPRIT*, automatic pairing of μ and ν spatial frequency estimates is facilitated by the fact that all of the quantities in (74) and (75) are real-valued. Thus, $\Psi_\mu + j\Psi_\nu$

may be spectrally decomposed as

$$\Psi_\mu + j\Psi_\nu = \mathbf{T}^{-1} \{\Omega_\mu + j\Omega_\nu\} \mathbf{T} \quad (76)$$

The algorithm based on these observations is similar in form to *2D DFT Beamspace ESPRIT* for a URA.

8 Simulations

Simulations were conducted employing an 8×8 URA (*i.e.*, $N = M = 8$) with $\Delta_x = \Delta_y = \lambda/2$. The source scenario consisted of $d = 3$ equi-powered, uncorrelated sources located at $(u_1, v_1) = (0, 0)$, $(u_2, v_2) = (1/8, 0)$, and $(u_3, v_3) = (0, 1/8)$, where u_i and v_i are the direction cosines of the i -th source relative to the x and y axes, respectively. Sources 1 and 2 were separated by a half-beamwidth, *i.e.*, half the Rayleigh resolution limit, as were sources 2 and 3. Sources 1 and 2 have the same v coordinate, while sources 2 and 3 have the same u coordinate. If the *ACMP* algorithm of van der Veen *et al* was applied in this scenario, it would provide a faulty estimate of the number of sources as well as faulty source direction estimates.

A given trial run at a given SNR level (per source per element) involved $N_s = 64$ snapshots. The noise was *i.i.d.* from element to element and from snapshot to snapshot. RMS error defined as

$$RMSE_i = \sqrt{E\{(\hat{u}_i - u_i)^2\} + E\{(\hat{v}_i - v_i)^2\}}, \quad i = 1, 2, 3, \quad (77)$$

was employed as the performance metric. Let $(\hat{u}_{i_k}, \hat{v}_{i_k})$ denote the coordinate estimates of the i -th source obtained from a particular algorithm at the k -th run. Sample performance statistics were computed from $K = 500$ independent trials as

$$\widehat{RMSE}_i = \sqrt{\frac{1}{K} \sum_{k=1}^K \{(\hat{u}_{i_k} - u_i)^2 + (\hat{v}_{i_k} - v_i)^2\}}, \quad i = 1, 2, 3. \quad (78)$$

The bias of *2D Unitary ESPRIT* for $N_s = 64$ snapshots over the range of SNR's simulated was found to be negligible, as was the bias of *2D DFT Beamspace ESPRIT*. This facilitated comparison with the Cramer Rao Lower Bound (CRLB). The performance of *2D Unitary ESPRIT* relative to *2D MUSIC* was also compared, as was the relative performance of *2D DFT Beamspace ESPRIT*. The CRLB and the theoretically predicted performance of *2D MUSIC* were computed according to formulas provided in [8] and are plotted in Figures 4(a), 4(b), and 4(c) for sources 1, 2, and 3, respectively.

Note that *2D MUSIC* essentially achieved the CRLB over the range of SNR's simulated so that its theoretically predicted RMSE curve is coincident with the CRLB curve. Of course, *2D MUSIC* requires the localization of 3 peaks of a 2D spectrum. In element space, determining the value of the *2D MUSIC* spectrum at a given point involves the calculation of an inner product of the form $\mathbf{a}^H(\mu, \nu) \mathbf{P}^\perp \mathbf{a}(\mu, \nu)$, where $\mathbf{P}^\perp$ is 64×64 . This kind of calculation has to be done repeatedly in performing a localized Newton-Raphson search around each spectral peak.

The respective RMSE's of *2D Unitary ESPRIT* and *2D DFT Beam-space ESPRIT* for sources 1, 2, and 3 are plotted in Figures 4(a), 4(b), and 4(c), respectively. In accordance with the summary of *2D Unitary ESPRIT* at the end of Section 3.0, the computations required for a single run were: (i) 64 additions per each of 64 snapshots to transform from complex-valued space to real-valued space, (ii) calculation of the 3 "largest" left singular vectors of a 64×128 real-valued matrix, (iii) calculation of the solution to two systems of equations of the form $\mathbf{A}\mathbf{X} = \mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ are both 64×3 and real-valued, and (iv) calculation of the eigenvalues of a 3×3 complex-valued matrix. The performance of *2D Unitary ESPRIT* is observed to be very close to the CRLB for SNR's greater than or equal to -6 dB, although it does not achieve the CRLB even at the rather high SNR level of 12 dB. (Keep in mind that there are 64 elements and that the SNR is that per element.) Observe that on a logarithmic scale, the small gap between the performance of *2D Unitary ESPRIT* and the CRLB is fairly constant as a function of SNR for SNR's above -6 dB.

To demonstrate the efficacy of working in a reduced dimension beam-space, *2D DFT Beam-space ESPRIT* employed a 3×3 set of 9 beams with mainlobes rectangularly spaced in the u - v plane and centered at $(u, v) = (0, 0)$. In accordance with the summary of *2D DFT Beam-space ESPRIT* at the end of Section 4.0, the computations required for a single run were: (i) 9 sets of 64 multiplications and 63 additions for each of 64 snapshots to transform from element space to beam-space, (ii) calculation of the 3 "largest" left singular vectors of a 9×128 real-valued matrix, (iii) calculation of the solution to two systems of equations of the form $\mathbf{A}\mathbf{X} = \mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ are both 9×3 and real-valued, and (iv) calculation of the eigenvalues of a 3×3 complex-valued matrix. A scatter plot of the 3 eigenvalues obtained from *2D DFT Beam-space ESPRIT* for each of 200 independent runs at an SNR of 3 dB is displayed in Figure 4(d). For SNR's greater than or equal to -6 dB, the performance of *2D DFT Beam-space ESPRIT* is observed to be only slightly worse than that of *2D Unitary ESPRIT* despite the dramatic reduction in computational complexity. Similar to *2D Unitary ESPRIT*, the gap between the performance of *2D DFT Beam-space ESPRIT* and the CRLB

is fairly constant as a function of SNR over the range of SNR's simulated.

An interesting observation is that for SNR's lower than -9 dB, *2D DFT Beamspace ESPRIT* outperformed *2D Unitary ESPRIT*. This is in accordance with observations made by Xu *et. al.* [16] in comparing the performance of their version of *Beamspace ESPRIT* with that of *ESPRIT* in element space. At low SNR's Xu *et. al.* argued that the better performance of the former over that latter is due to fact that *Beamspace ESPRIT* exploits a-priori information on the source locations by forming beams pointed in the general directions of the sources. This argument is applicable here as well.

The difference in performance between *2D Unitary ESPRIT* or *2D DFT Beamspace ESPRIT* and the CRLB, and the fact that *2D MUSIC* achieves the CRLB for the range of SNR's simulated, suggests a strategy wherein the 2D angle estimates provided by either *2D Unitary ESPRIT* or *2D DFT Beamspace ESPRIT* are used as starting points for localized Newton searches of the *2D MUSIC* spectrum to achieve uniformly minimum variance unbiased estimates (UMVUE's). Note that the computational burden of performing these localized searches of the *2D MUSIC* spectrum may be reduced substantially by operating in beamspace and exploiting the conjugate centro-symmetry of the URA manifold.

9 Conclusions

2D Unitary ESPRIT is a closed form 2D angle estimation algorithm for use in conjunction with a URA and is easily adapted for other dual invariance arrays including a cross array. *2D DFT Beamspace ESPRIT* is an efficient beamspace implementation of *2D Unitary ESPRIT* facilitating reduced dimension processing and attendant reduction in computational complexity. The 2D angle estimates provided by either *2D Unitary ESPRIT* or *2D DFT Beamspace ESPRIT* may be used as starting points for localized Newton searches of the *2D MUSIC* spectrum, the *ML* algorithm, or the *Multiple Invariance ESPRIT* algorithm. Due to space limitations, performance analysis of either *2D Unitary ESPRIT* or *2D DFT Beamspace ESPRIT* is not included here, but would follow in the same vein as the performance analysis of *UCA-ESPRIT* in [22]. Note that *2D Unitary ESPRIT* may also be employed in a variety of applications other than 2D angle estimation including 2D harmonic retrieval for image analysis, for example.

10 Bibliography

- [1] R. Roy and T. Kailath, "ESPRIT-Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, pp.984-995, July 1989.
- [2] A. L. Swindlehurst, B. Ottersten, G. Xu, R. H. Roy, and T. Kailath, "Multiple Invariance ESPRIT", *IEEE Trans. Signal Processing*, vol. 40, pp. 867-881, Apr. 1992.
- [3] A. L. Swindlehurst and T. Kailath, "Azimuth/elevation direction finding using regular array geometries", *IEEE Trans. Aerospace and Electronic Systems*, vol. 29, pp. 145-156, Jan. 1993.
- [4] M. P. Clark and L. L. Scharf, "Two-Dimensional Modal Analysis Based on Maximum Likelihood", *IEEE Trans. Signal Processing*, vol. 42, pp. 1443-1452, June 1994.
- [5] M. D. Zoltowski and D. Stavrinos, "Sensor Array Signal Processing via a Procrustes Rotations Based Eigenanalysis of the ESPRIT Data Pencil," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, pp. 832-861, Jun. 1989.
- [6] A.J. van der Veen, P.B. Ober, E.D. Deprettere, "Azimuth and Elevation Computation in High Resolution DOA Estimation", *IEEE Trans. Signal Processing*, vol. 40, pp. 1828-1832, July 1992.
- [7] M.D. Zoltowski and C.P. Mathews, "Closed-Form 2D Angle Estimation with Uniform Circular Arrays Via Phase Mode Excitation and ESPRIT," *27th Asilomar IEEE Conf. on Signals, Systems, and Computers*, vol. 1, pp. 169-173, Nov. 1993.
- [8] C.P. Mathews and M.D. Zoltowski, "Eigenstructure Techniques for 2-D Angle Estimation with Uniform Circular Arrays," scheduled to appear in *IEEE Trans. on Signal Processing*, September 1994.
- [9] M. Haardt and M.E. Ali-Hackl, "Unitary ESPRIT: How to Exploit Additional Information Inherent in the Rotational Invariance Structure", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Adelaide, Australia, Apr. 1994.
- [10] M. Haardt and J.A. Nossék, "Unitary ESPRIT: How to Obtain Increased Estimation Accuracy with a Reduced Computational Burden", *Technical Report No. TUM-LNS-TR-94-3*, Institute

of Network Theory & Circuit Design, Technical University of Munich, D-80290 Munich, Germany, May 1994.

- [11] K.C. Huarng and C.C. Yeh, "A unitary transformation method for angle-of-arrival estimation", *IEEE Trans. Signal Processing*, vol. 39, pp. 975-977, Apr. 1991.
- [12] D.A. Linebarger, R.D. DeGroat, and E.M. Dowling, "Efficient Direction Finding Methods Employing Forward/Backward Averaging", scheduled to appear in *IEEE Trans. Signal Processing*, 1994.
- [13] K. Buckley and X.-L. Xu, "Spatial-Spectrum Estimation in a Location Sector," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-38, no. 11, pp. 1842-1852, Nov. 1990.
- [14] G. Bienvenu and L. Kopp, "Decreasing High Resolution Method Sensitivity by Conventional Beamforming Preprocessing," in *Proc. of 1984 IEEE Int'l Conf. on Acoust., Speech, and Signal Process.*, pp. 33.2.1-33.2.4, April 1984.
- [15] H. Lee and M. Wengrovitz, "Resolution Threshold of Beamspace MUSIC for Two Closely-Spaced Emitters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1545-1559, Sept. 1990.
- [16] G. Xu, S.D. Silverstein, R. H. Roy, and T. Kailath, "Beamspace ESPRIT," *IEEE Trans. Signal Processing*, vol. 42, pp. 349-356, Feb. 1994.
- [17] M.D. Zoltowski, G.M. Kautz, and S.D. Silverstein, "Beamspace Root-MUSIC", *IEEE Trans. Signal Processing*, vol. 41, pp. 344-364, Jan. 1993.
- [18] K.J.R. Liu, D.P. O'Leary, G.W. Stewart, and Y.J.J. Wu, "An Adaptive ESPRIT Based on URV decomposition", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. IV, pp. 37-40, Minneapolis, MN, Apr. 1993.
- [19] G. Xu, R.H. Roy, and T. Kailath, "Detection of Number of Sources via Exploitation of Centrosymmetry Property", *IEEE Trans. Signal Processing*, vol. 42, pp. 102-112, Jan. 1994.
- [20] M. D. Zoltowski and T. Lee, "Maximum Likelihood Based Sensor Array Signal Processing in the Beamspace Domain for Low-Angle Radar Tracking," *IEEE Trans. on Signal Processing*, vol. 39, pp. 656-671, Mar. 1991.

- [21] F. Vanpoucke, M. Moonen, and Y. Berthoumieu, "An Efficient Subspace Algorithm for 2-D Harmonic Retrieval", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Adelaide, Australia, Apr. 1994.
- [22] C. P. Mathews and M. D. Zoltowski, "Performance Analysis of the UCA-ESPRIT Algorithm for Circular Ring Arrays," scheduled to appear in *IEEE Trans. on Signal Processing*, September 1994.

u = dir. cosine wrt array/displacement axis

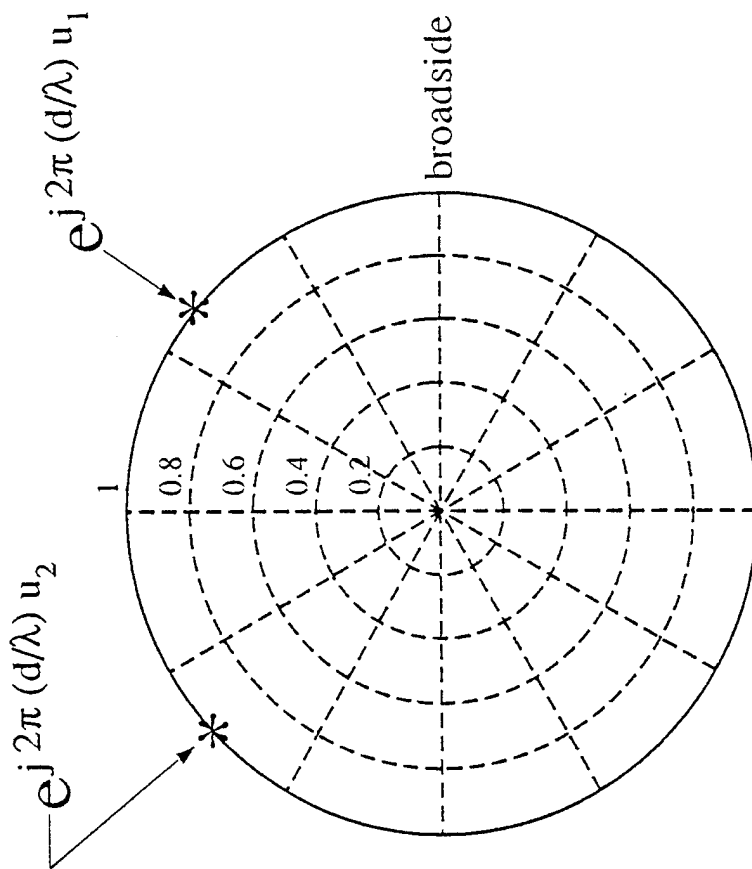


Figure 1. Illustrating the form of signal roots obtained via Root-MUSIC with ULA or ESPRIT with single invariance (roots are eigenvalues).

u = dir. cosine wrt x-axis; v = dir. cosine wrt y-axis

θ = elevation angle; ϕ = azimuth angle

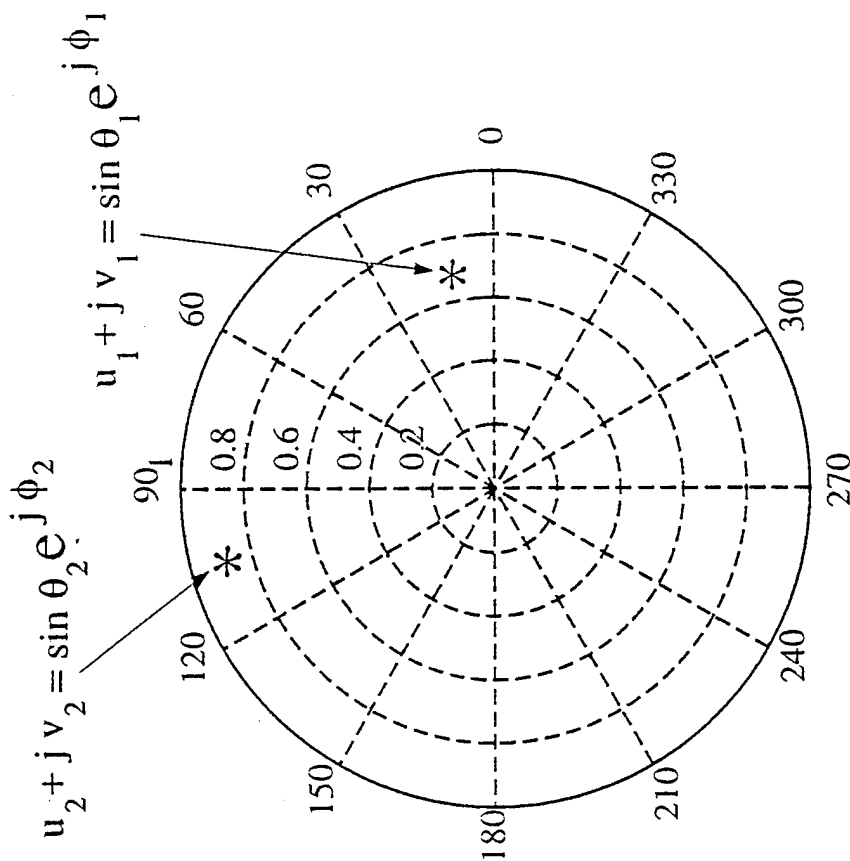


Figure 2. Illustrating the form of signal roots obtained via UCA-ESPRIT with circular array or approximate 2D DFT Beamspace ESPRIT with rectangular array.

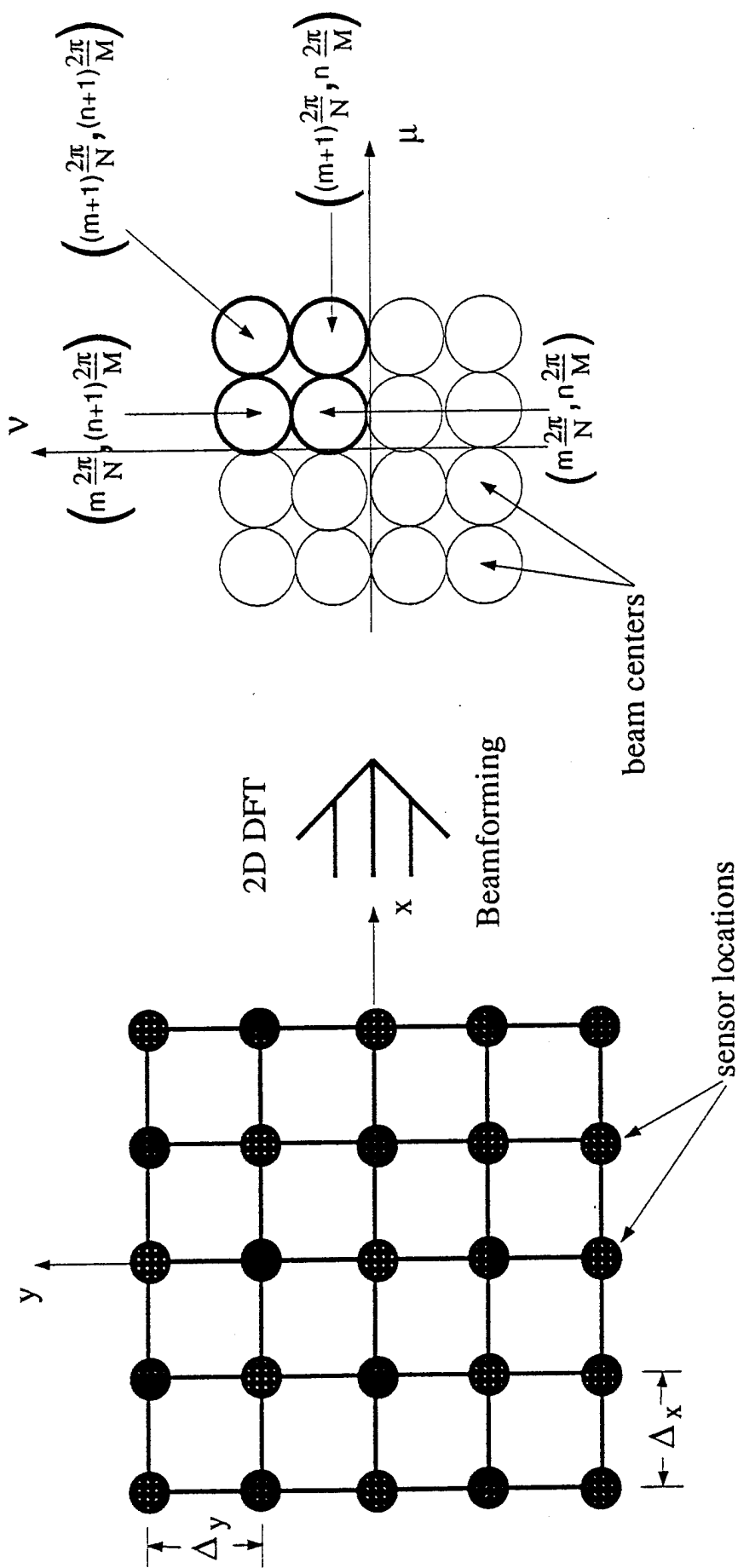


Figure 3. Illustration of transformation from element space to beamspace highlighting four beam example for 2D DFT Beamspace ESPRIT.

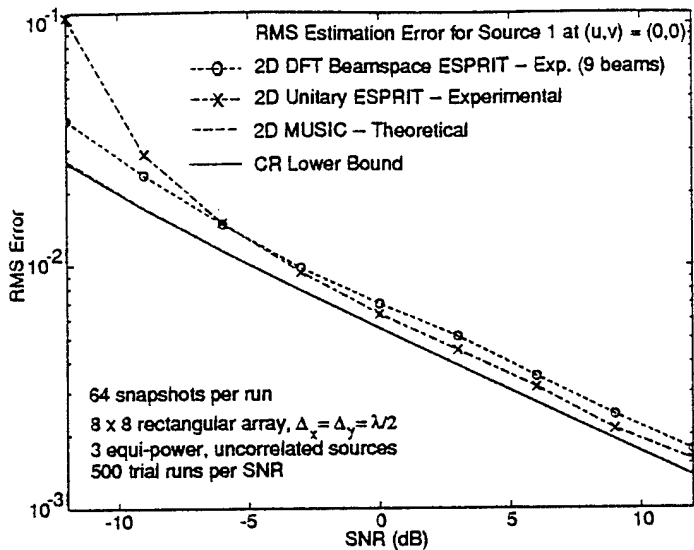


Figure 4: (a) RMSE for source 1 in simulation example.

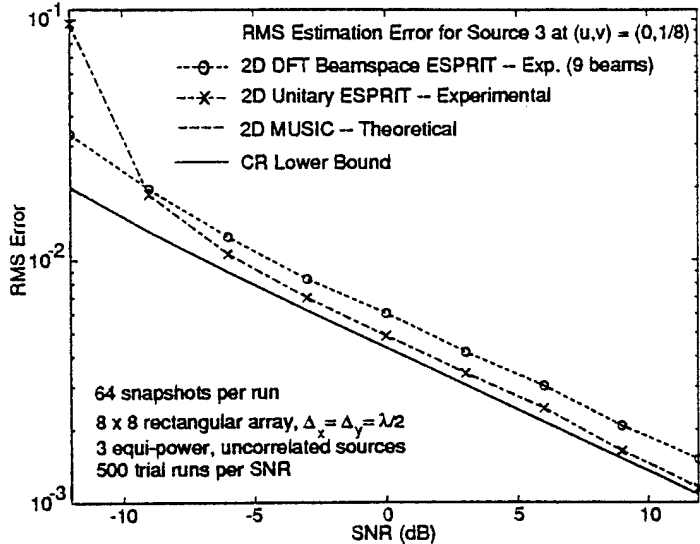


Figure 4: (c) RMSE for source 3 in simulation example.

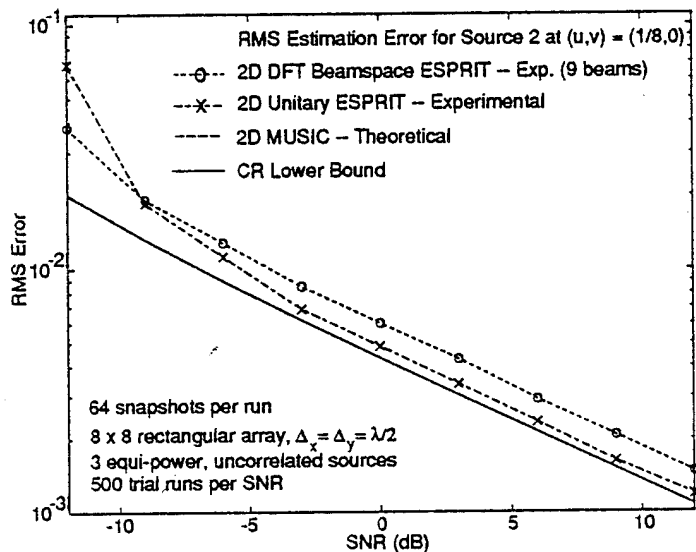


Figure 4: (b) RMSE for source 2 in simulation example.

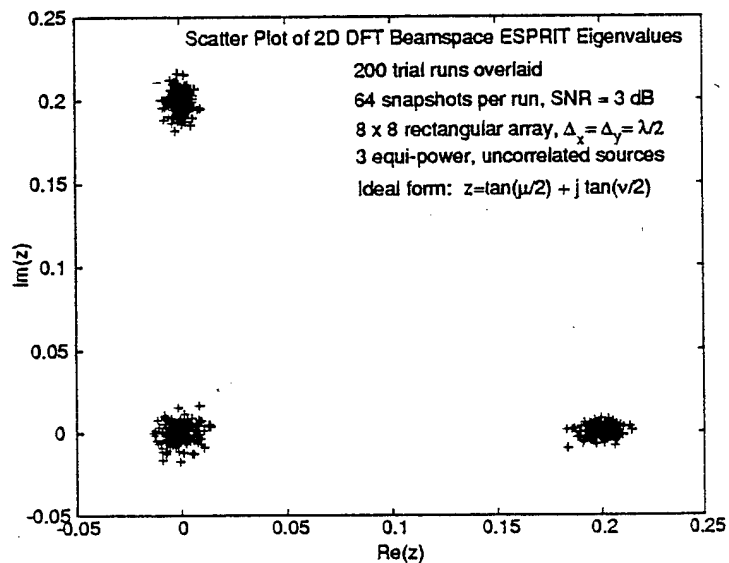


Figure 4: (d) Scatter plot of 2D DFT Beamspace ESPRIT eigenvalues.

2 Beamspace DOA Estimation Featuring Multirate Eigenvector Processing

A novel approach to angle of arrival estimation in beamspace is developed. Beamspace noise eigenvectors may be transformed to vectors in the element-space noise subspace. The transformed noise eigenvectors are bandpass, facilitating multirate processing involving modulation to baseband, filtering, and decimation. As these operations are linear, a matrix transformation applied to the eigenvectors may be constructed a priori. Incorporation of the technique into either the Root-MUSIC or ESPRIT prescriptions provides a computationally efficient procedure. Compared to past efforts to adapt Root-MUSIC and ESPRIT to beamspace, this approach circumvents the need for restrictive requirements on the form of the beamforming transformation. An asymptotic theoretical performance analysis is also included to provide an alternative to computationally intensive Monte-Carlo simulations. Simulation studies show the validity of the performance predictive expressions and verify that the procedure, when incorporated into the Root-MUSIC/ESPRIT formulations, produces a direction finding technique that nearly attains the Cramer-Rao bound.

2.1 Introduction

2.2 Array Signal Model

2.3 Development of DOA Estimators Featuring Multirate Eigenvector Processing

2.3.1 Multirate Noise Eigenvector Processing

2.3.2 Incorporation of Filter Deconvolution

2.3.3 Root-MUSIC Incorporating Multirate Eigenvector Processing

2.3.4 TLS-ESPRIT Incorporating Multirate Eigenvector Processing

2.3.5 Location of Extraneous Roots Created by Filtering

2.4 Theoretical Performance Analysis

2.4.1 Performance Analysis of Root-MUSIC Formulation

2.4.2 Performance Analysis of ESPRIT Formulation

2.5 Computer Simulations

2.6 Conclusions/Remarks

2.7 References

2.8 Appendix: Asymptotic Variance of ESPRIT Formulation

1. Introduction

Beamspace formulations of the eigenstructure class of direction finding sensor array processing algorithms offer a number of advantages over their element space counterparts. First, there is a computational benefit realized in the processing of data of a much smaller dimension. Second, a practical implementation to current phased array technology is allowed. Third, beamspace formulations exhibit a reduced sensitivity to sensor position perturbations and noise non-idealities [1]. Fourth, although suboptimal in high SNR situations, the inherent concentration over a specific spatial region of interest leads to noise reduction and, hence, enhanced ability for localization in the more critical case of low SNR [2, 3].

In the case of the Spectral MUSIC algorithm proposed by Schmidt [4], which is applicable to arbitrary array geometries, the Vandermonde nature of the element space array response to a plane wave signal for the common uniform linear-spaced array geometry facilitates a root-finding procedure for angle estimation [5] as a computationally attractive alternative to the spectral search. The beamspace formulation of Spectral MUSIC, however, does not directly offer a polynomial root-finding capability. By relating the beamspace manifold to the element space direction vector, a beamspace Root-MUSIC capability can be realized but the order of the resulting polynomial to be rooted is related to the number of sensors, N , as $2N - 2$. This represents such a considerable computationally intensive task for large arrays so as to preclude its use for the associated performance gains as noted in [6].

Recently, an efficient algorithm was proposed in [7] as a means of reducing the Root-MUSIC polynomial to order $2N_b - 2$, where N_b is the number of beams. This represents a tremendous computational savings if only a relatively few number of beams are formed to probe a spatial subband (sector) for sources. The approach in [7] was accomplished by requiring that the beamforming vectors possess common spatial nulls. We point out that, like the beamspace Root-MUSIC formulation in [7], an adaptation of ESPRIT to beamspace in [8] also required significant restrictions on the form of the beamforming vectors. Aside from this possibly over-restrictive requirement, two other problems associated with the beamspace Root-MUSIC algorithm were observed. First, the technique didn't exploit the spatially-confined region of operation in the rooting stage of the algorithm. That is, as the number of sensors comprising the array increases, the spatial extent of the beamforming window decreases with constant N_b but, yet, the rooting algorithm is still capable of localizing signals over all of visible space. Second, the approach involved the use of an $N_b \times N_b$ matrix transformation $\mathbf{Q}$ which was found to be highly ill-conditioned. For example,

the condition number of $\mathbf{Q}$ for an $N = 128$ element array operated upon by a spatial Discrete Fourier Transform (DFT) beamformer was computed for a varying number of beams and plotted in Figure 1. In contrast, the other curve in the figure ($\mathbf{Z}$ transformation) corresponds to an alternative approach that is the key result of this paper, having a similar implementation for the MUSIC setting but fundamentally different to the approach in [7]. Whereas the condition number associated with the $\mathbf{Z}$ transformation is relatively constant at a value near 3 for all beamspace dimensions, the corresponding value for the $\mathbf{Q}$ transformation is large for even a small number of beams, e.g., for a beamformer comprised of $N_b = 8$ spatial DFT beams, the condition number is approximately $8 \cdot 10^9$.

The main purpose of this paper is to develop a processing methodology that is based on the transformability of a beamspace noise eigenvector to an element-space counterpart as noted in passing in [3, 9]. In the intended application of beamspace processing, a spatial subband is probed so that the transformed beamspace noise eigenvectors are naturally bandlimited in a spatial sense. This banded characteristic allows for the application of classical multirate digital signal processing to isolate and spatially enlarge the spatial subband of interest. Note that this methodology departs from classical multirate processing in that the pertinent information lies in the in-band signal nulls instead of signal peaks. In the conventional mode of multirate processing, one has to be concerned with spectral peaks outside the “basebanded” subband being aliased into the subband thereby causing ambiguities. Lowpass filtering is implemented prior to decimation to avoid this condition. However, in array processing at the sensor level, this pre-filtering operation destroys the Vandermonde nature of the manifold thereby precluding rooting based DOA estimation techniques such as Root-MUSIC or ESPRIT. Here the goal is to preserve in-band signal nulls and the development will show that the ability to root is easily maintained. In addition, with respect to aliasing artifacts, out-of-band signals not sufficiently de-emphasized by the front-end beamforming give rise to out-of-band signal nulls which actually serve to suppress aliasing contributions resulting from decimation (see Figure 2 to be discussed shortly).

An important feature of this approach is that there are no restrictive requirements on the form of the beamforming vectors. Another advantage of this technique is that the angular separation between “in-band” signal roots is increased by the decimation factor, thereby easing the job of rooting. Another major advantage is that the technique is computationally robust as the $\mathbf{Z}$ matrix transformation applied to the beamspace noise eigenvectors is well conditioned, e.g., refer to Figure 1 where the condition number

of a $\mathbf{Z}$ transformation is shown for the same array length and a suitable decimation procedure.

As the eigenvector transformation-decimation procedure is general in nature, the technique may be applied to any eigenstructure direction finding algorithm. We here consider the Root-MUSIC and ESPRIT [10] formulations as these techniques are fairly representative of the eigenstructure class of angle estimators; application to other algorithms is straightforward.

The contents of this paper are as follows. Following a description of the data model, the beamspace noise eigenvector transformation-decimation technique is developed and applied to Root-MUSIC and ESPRIT ideology in Section 3. The theoretical performance of the MUSIC/ESPRIT formulations is developed, in terms of the estimation variance, in Section 4. Finally, the theoretical performance expressions are validated in simulations and the optimality of the technique is observed through a comparison study with the stochastic Cramer-Rao bound in a variety of experiments in Section 5.

2. Array Signal Model

The DOA estimation methodology described herein assumes a uniform linear array of sensors. An extension to the two-dimensional array geometry composed of a rectangular lattice of sensors is readily clear.

Assuming that K narrowband plane-wave signals, residing at a common center frequency, impinge upon an N -sensor array, the complex basebanded data snapshot vector at the m 'th sampling instant, $\mathbf{x}(m)$, is expressed as a superposition of signals embedded in additive noise as

$$\mathbf{x}(m) = \sum_{k=1}^K s_k(m) \mathbf{a}_N(\mu_k) + \mathbf{n}(m) \quad m = 1, \dots, M \quad (1)$$

In the above equation, the amplitudes of the K signals, $s_k(\cdot)$, $k = 1, \dots, K$, are modelled as zero-mean jointly Gaussian random variables with non-singular covariance $\mathbf{P}_s$, and $\mathbf{n}(\cdot)$ is a zero-mean complex gaussian noise vector with assumed covariance $\mathcal{E} [\mathbf{n}(m)\mathbf{n}^H(m)] = \sigma_n^2 \mathbf{I}_N$. The array response to a unit-amplitude signal arriving from the spatial location μ is represented by $\mathbf{a}_N(\mu)$, where $\mu = \frac{2\pi}{\lambda} d \sin(\theta)$, d is the sensor spacing, λ is the wavelength, and θ is the conical angle of arrival. In accordance with a uniform sensor placement, the structure of the array manifold vector has the form

$$\mathbf{a}_N(\mu) = [1, e^{j\mu}, e^{j2\mu}, \dots, e^{j(N-1)\mu}]^T. \quad (2)$$

The associated sensor covariance matrix, assuming that the noise is uncorrelated with the signal set, is

simply

$$\mathbf{R}_x = \mathbf{A} \mathbf{P}_s \mathbf{A}^H + \sigma_n^2 \mathbf{I}_N, \quad (3)$$

where $\mathbf{A}$ is the matrix of element-space manifold vectors, $\mathbf{A} = [\mathbf{a}_N(\mu_1) : \mathbf{a}_N(\mu_2) : \dots : \mathbf{a}_N(\mu_K)]$. In the event that the noise exhibits a colored character, we assume that the noise correlation matrix is known.

The benefits of beamforming as a pre-processing operation prior to DOA estimation is well known in the literature [1, 2, 3, 7]. Here we transform the element-space data to an N_b dimensional beamspace in a digital or analog fashion. This operation is mathematically modelled as

$$\mathbf{y}(m) = \mathbf{W}^H \mathbf{x}(m) \quad m = 1, \dots, M \quad (4)$$

where the columns of the $N \times N_b$ beamforming matrix are orthonormalized so that $\mathbf{W}^H \mathbf{W} = \mathbf{I}_{N_b}$. Denoting the N_b -dimensional beamspace manifold vector as $\mathbf{b}(\mu)$, the associated beamspace covariance, $\mathbf{R}_y$, is

$$\mathbf{R}_y = \mathcal{E} [\mathbf{y}(m) \mathbf{y}(m)^H] = \mathbf{W}^H \mathbf{R}_x \mathbf{W} = \mathbf{B} \mathbf{P}_s \mathbf{B}^H + \sigma_n^2 \mathbf{I}_{N_b}, \quad (5)$$

where $\mathbf{B} = [\mathbf{b}(\mu_1) : \mathbf{b}(\mu_2) : \dots : \mathbf{b}(\mu_K)]$ and $\mathbf{b}(\mu) = \mathbf{W}^H \mathbf{a}_N(\mu)$.

As the ideal covariance matrix is not accessible in practice, an M -sample estimate is employed as

$$\hat{\mathbf{R}}_y = \sum_{m=1}^M \mathbf{y}(m) \mathbf{y}^H(m), \quad (6)$$

where we assume that $M > K$. We also assume that $N_b > K$ for proper operation of the DOA estimators. As the beamspace dimension, N_b , is usually chosen to be small in relation to N , to yield a computationally attractive algorithm displaying enhanced localization performance of low SNR signals [7, 9], the assumption $K < N_b$ may seem too restrictive. However, through judicious selection of beamforming vectors, we merely assume that fewer than N_b signals are effectively present in the beamspace data; signals that are not located within the spatial sector of interest are sufficiently de-emphasized by the beamforming operation.

The eigendecomposition of $\hat{\mathbf{R}}_y$ provides the signal and noise subspace descriptors as necessitated by the DOA architectures considered in this paper. Notationally, $\hat{\mathbf{R}}_y$ is decomposed as $(\hat{\lambda}_i, \hat{\mathbf{e}}_i)$ where $\hat{\lambda}_i$, $i=1, \dots, N_b$, are the eigenvalues arranged in decreasing order, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_{N_b} > 0$, with associated eigenvectors $\hat{\mathbf{e}}_i$. Thus $\{\hat{\mathbf{e}}_i, i = 1, \dots, K\}$ span a K -dimensional (signal) subspace used as an estimate of the true subspace spanned by the columns of $\mathbf{B}$, and the remaining $N_b - K$ eigenvectors span an estimate of

the orthogonal (noise) subspace. The number of signals, K , is assumed to be available, possibly estimated via a procedure such as that described in [11].

An alternative procedure for the estimation of the noise or signal subspace is the decomposition of the real part of $\hat{\mathbf{R}}_y$ as discussed in [7, 9]. By simply referencing the phase of the beamforming vectors and the element space manifold to the array center, i.e., through the scaling of (2) by the multiplicative factor $\exp(-j\mu\frac{N-1}{2})$, and requiring a symmetric magnitude taper in the beamforming vectors, the beamspace manifold $\mathbf{b}(\mu)$ is real-valued. Thus $\text{Re}\{\mathbf{R}_y\} = \mathbf{B}\text{Re}\{\mathbf{P}_s\}\mathbf{B}^T + \sigma_n^2\mathbf{I}_{N_b}$. The advantages of processing only the real part of $\hat{\mathbf{R}}_y$ are a computational savings and a signal decorrelation effect to improve the angle estimation accuracy in correlated signal scenes [9]. Note that the forthcoming discussion of DOA estimation employing eigenvector decimation places no restrictions on how the signal or noise eigenvectors are estimated.

3. Development of DOA Estimators Featuring Multirate Eigenvector Processing

In this section, we develop the beamspace Root-MUSIC and TLS-ESPRIT DOA estimators incorporating multirate eigenvector processing. In Section A, we discuss the basis of the multirate processing technique of beamspace noise eigenvectors and present some computational reductions in Section B. Finally the techniques are applied to obtain Root-MUSIC and TLS-ESPRIT DOA estimation algorithms in Sections C and D, respectively.

A. Multirate Noise Eigenvector Processing

The critical relation motivating the development of the algorithms presented in this paper is that a beamspace noise eigenvector can be transformed to a noise eigenvector in element space as noted in [3, 9]. Defining

$$\mathbf{v}_i = \mathbf{W} \mathbf{e}_i, \quad (7)$$

where $\mathbf{e}_i$, $i > K$, is a noise eigenvector of the ideal beamspace covariance, we see that $\mathbf{v}_i$ is indeed an eigenvector lying in the noise subspace of $\mathbf{R}_x$ as evidenced by

$$\mathbf{0} = \mathbf{B}^H \mathbf{e}_i = (\mathbf{W}^H \mathbf{A})^H \mathbf{e}_i = \mathbf{A}^H (\mathbf{W} \mathbf{e}_i) = \mathbf{A}^H \mathbf{v}_i \quad i > K. \quad (8)$$

Since $\mathbf{A}$ is an $N \times K$ matrix composed of the element space direction vectors which collectively span the signal subspace, $\mathbf{v}_i = \mathbf{W} \mathbf{e}_i$, $i=K+1, \dots, N_b$, lies in the element space noise subspace. Also, given that $\mathbf{e}_i$ is

unit-length, $\mathbf{v}_i$ is unit-length as guaranteed by the orthonormality of the columns of $\mathbf{W}$. Note, however, that no direct relationship exists between the beamspace and element-space signal subspace eigenvectors and that the $N_b - K$ transformed noise eigenvectors only partially describe the N -dimensional element-space noise subspace.

We now focus the development of the multirate eigenvector prescription to the MUSIC algorithm. Employing the transformed noise eigenvectors which partially describe the element-space noise subspace, the associated MUSIC null spectrum [4] is appropriately described as

$$\mathbf{S}_{MU}(\mu) = \sum_{k=K+1}^{N_b} |\mathbf{a}_N^H(\mu) \mathbf{v}_k|^2, \quad -\pi \leq \mu \leq \pi \quad (9)$$

For the structure of the array manifold given in (2), it is observed that each term in (9) simply has the form of an N -point spatial Discrete Time Fourier Transform (DTFT) of a transformed noise eigenvector,

$$V_k(\mu) = \mathbf{a}_N^H(\mu) \mathbf{v}_k = \sum_{n=1}^N v_k(n) e^{j\mu(n-1)} \quad -\pi \leq \mu \leq \pi \quad (10)$$

where $v_k(n)$ represents the n 'th entry in the vector $\mathbf{v}_k$.

By selecting the set of beamforming vectors to interrogate some sector of space while attenuating signals that lie elsewhere, the spectrum of the transformed eigenvectors are naturally spatially band-limited. This can be seen by viewing the null spectrum of a single transformed noise eigenvector as shown in Figure 2. The parameters associated with the figure are as follows. $N=128$ half-wavelength spaced sensors were employed in conjunction with a spatial DFT beamformer consisting of eight consecutive beams centered in space at $\sin \theta = 25/N$. For reference purposes, the spatial responses of the $N_b = 8$ beams are plotted in Figure 3. There were two equi-powered signals located near mid-band at 10.6° and 11.5° ; the locations are labelled on the figure. In addition, a high-strength signal was placed at a distant location of $\sin \theta = 69/N$. A single beamspace noise eigenvector of the ideal covariance was employed to generate the plot in Figure 2. Note that the "extraneous" null within the band will fill in when all of the transformed noise eigenvectors are employed. Although in-band nulls are of interest, the main point of the figure is that the spectrum exhibits an elevated response in the spatial region where the beams are directed and a suppressed response in the region neighboring the distant signal.

The bandpass nature of the null spectra suggests a multirate procedure wherein the spatial band surrounding $\sin \theta = 25/N$ is spatially basebanded and the corresponding spatial sequence is decimated. Consider decimation by an integer factor D that is less than or equal to the maximum allowable value.

For the example employing N_b spatial DFT beams, the maximum decimation factor is $D_{max} = N/N_b$.² The sequence associated with the k 'th decimated eigenvector is (recall $\mathbf{v}_k = \mathbf{W} \mathbf{e}_k$)

$$v_D^{(k)}(i+1) = v_k(Di+1) \quad i = 0, 1, \dots, N/D - 1.$$

From classical multirate theory, the spatial spectrum associated with the k 'th decimated eigenvector is

$$V_D^{(k)}(\mu) = \sum_{\ell=0}^{D-1} V_k\left(\frac{\mu - 2\pi\ell}{D}\right). \quad (11)$$

Keep in mind the periodicity in the variable μ , i.e., $V(\mu + 2\pi n) = V(\mu)$ for integer n . Assuming that the spectrum has negligible amplitude outside of the region of interest, i.e., $V_k(\mu) \approx 0$, $|\mu| > \pi/D$, only the $\ell = 0$ term contributes to the sum leading to $V_D^{(k)}(\mu) \approx V_k(\mu/D)$ for $-\pi < \mu < \pi$.

In the usual application of multirate processing, one must be concerned with the aliasing of signals into the band of interest; here we must insure that aliasing does not result in the "filling in" of signal nulls within the band of interest. Note that signals that lie outside of the spatial band of interest do not affect the spectrum, i.e., in fact, the reduced amplitude in the neighboring region as seen in Figure 2 will result in a smaller aliasing contribution. However, the presence of the large distant signals may increase the perceived dimension of the signal subspace, $\hat{K}$, in the decomposition of the sample covariance matrix so that their presence is undesired.

If the front-end beamformers have high sidelobes, a spatial filter prior to decimation might be necessary to insure that the null spectrum is not distorted due to aliasing, i.e., the "signal" nulls are not lost or shifted appreciably. The filter should incorporate a sufficient stopband attenuation to limit the degree of aliasing. However, a larger stopband attenuation requires a larger filter length. As the ultimate intention of multirate processing is to reduce the dimension of the transformed/decimated noise eigenvectors, a shorter-length filter is desired. Note that the length of the noise eigenvectors after decimation is $\lceil \frac{N+L-1}{D} \rceil$, where L is the filter length, D is the decimation factor which is less than or equal to $D_{max} = N/N_b$, and $\lceil x \rceil$ refers to the smallest integer greater than or equal to x .

As there is no need for a linear phase requirement, an IIR filter may be employed. The absence of a linear phase requirement in IIR designs should result in a smaller filter length, L , where L is taken as some appropriate effective length of the associated impulse response. Note, however, that the classic IIR low-pass filter designs such as Butterworth, Chebyshev, Elliptic, etc., yield poles that are very near

²Although the terminology "sampling rate alteration" applies for non-integer D_{max} , we will still refer to the rate conversion operation as decimation.

the unit circle so that the associated impulse responses are relatively long. It was determined that these classic IIR designs offer little or no advantages in terms of lengths vs. band specifications as compared to such FIR techniques as the Hamming, Hanning, or Blackman windowed low-pass filters (LPF). Also note that a high degree of passband ripple may not pose a significant problem as there is a procedure, to be discussed shortly, for the removal of the residual ripple after decimation.

A major factor in determining an appropriate filter length is the width of the transition band. The simplest means of increasing the width of the transition band, and, hence, shortening the filter length, is to decimate by a factor that is less than the maximum allowable limit D_{max} . This would increase the distance between the edges of the beamforming sector, i.e., the region encompassed by the mainlobes of the N_b beams, and the spatial location $\mu = \pi/D$, i.e., the location that is scaled-up to the spectral edge ($\mu = \pi$) after decimation. Thus, by designing a filter with a transition band that lies within a spatial zone that is exterior to the passband of the beamforming sector, the aliasing effects are essentially confined to this region which is disregarded in the end.

Another approach is to simply allow the passband edge to extend within the beamforming sector as it has been shown in [7, 9] that beamspace DOA architectures tend to perform rather poorly in terms of estimation bias/variance at the edges of the beamforming sector. This effect is attributable to the reduction in the total signal power, proportional to $\mathbf{b}^H(\mu) \mathbf{b}(\mu)$, as the signal nears the edge of the spatial subband. Thus the transition band of the filter may be designed to encompass perhaps 25-50% of the total beamforming sector in which case one would have to allow a corresponding overlap appropriate amongst subbands probed in succession or in parallel. Due to the characteristic shape of the noise eigenvector spectra, the aliasing effects primarily originate just outside of the pre-decimation subband defined over $\mu \in [-\pi/D, \pi/D]$. Thus, specifying that the transition band of the filter be centered at π/D , the aliasing will be primarily present in the edges of the beamforming window and this is disregarded.

Returning to the $N_b = 8$ beam example, an $N = 128$ element Hamming-windowed LPF with a transition band defined over the region $\mu \in [6.5\pi/N, 9.5\pi/N]$, where $\mu = 8\pi/N$ is both the edge of the beamforming window and the edge of the pre-decimation subband, proved to be a reasonable design. A sketch of the passband associated with this low pass filter design can be found in Figure 4. The filter response is superimposed over the MUSIC null spectrum associated with the use of all spatially basebanded transformed noise eigenvectors to show another feature of this filter selection: the interlacing

of the nulls which results in a dramatical reduction in the effects of aliasing. As the out-of-band nulls of the basebanded beamspace MUSIC null spectrum are at known data-independent spatial positions corresponding to the common null locations of the beam set of Figure 3, the filter parameters can be selected to produce the null interlacing effect as seen in Figure 4. Also note that the use of all beamspace noise eigenvectors in a MUSIC formulation resulted in the removal of the non-signal in-band spatial null that was present in the single transformed noise eigenvector spectrum of Figure 2. The resulting filtered eigenvector MUSIC null spectrum is shown in Figure 5 and the corresponding decimated MUSIC null spectrum is shown in Figure 6.

With the modulation (spatial basebanding), filtering, and decimation operations notated by $\mathcal{M}$, $\mathcal{F}$, and $\mathcal{D}$, respectively, the decimated/transformed noise eigenvectors are then $\nu_i = \mathcal{D}\mathcal{F}\mathcal{M}\{\mathbf{W}\mathbf{e}_i\}$, $i > K$. As decimation, filtering, and modulation are linear operations, these may be performed a priori on the N_b columns of $\mathbf{W}$ as evidenced in

$$\nu_i = \mathcal{D}\mathcal{F}\mathcal{M}\left\{\sum_{k=1}^{N_b} \mathbf{w}_k e_i(k)\right\} = \sum_{k=1}^{N_b} [\mathcal{D}\mathcal{F}\mathcal{M}\{\mathbf{w}_k\}] e_i(k) = \sum_{k=1}^{N_b} \mathbf{z}_k e_i(k) = \mathbf{Z}\mathbf{e}_i, \quad (12)$$

where

$$\mathbf{Z} = [\mathbf{z}_1 : \mathbf{z}_2 : \dots : \mathbf{z}_{N_b}] = \mathcal{D}\mathcal{F}\mathcal{M}\{\mathbf{W}\}. \quad (13)$$

Hence, a matrix $\mathbf{Z}$ of dimension $N_Z \times N_b$, where $N_Z = \lceil \frac{N+L-1}{D} \rceil$, may be computed a priori and applied to the beamspace noise eigenvectors $\mathbf{e}_i$, $i = K+1, \dots, N_b$. In the more general case of sampling rate conversion where the desired "decimation" factor is not an integer but can be expressed as a ratio of two integers $D = M_D/M_I$, the corresponding matrix $\mathbf{Z}$ is computed as

$$\mathbf{Z} = \mathcal{D}_{M_D} \mathcal{F} \mathcal{I}_{M_I} \mathcal{M}\{\mathbf{W}\}, \quad (14)$$

where $\mathcal{D}_{M_D}$ represents a decimation operation by a factor of M_D and $\mathcal{I}_{M_I}$ refers to an interpolation operation by a factor of M_I . Note that the filter frequency design specifications are appropriately modified to reflect the positioning following the interpolator. Also, due to the modulation operation, the matrix $\mathbf{Z}$ can be employed for a common beam set steered to any sector of space. In this mode of operation, the estimates of the signal μ locations provided by the algorithm are relative to the center of the beamforming sector.

B. Incorporation of Filter Deconvolution

As the inclusion of a properly designed filter will result in negligible aliasing effects, it is possible to reduce the row-dimension of the matrix $\mathbf{Z}$, and hence the order of the polynomial that ultimately needs to be rooted. This computational advantage is accomplished through the deconvolution of the decimated filter sequence from each column of $\mathbf{Z}$ as substantiated in this section.

Denoting the spatial DTFT of the i 'th transformed and decimated beamspace noise eigenvector $\mathbf{v}_i$ defined in (12) as $V_{DFM}^{(i)}(\mu)$, we find, as similar to the form in Equation (10),

$$V_{DFM}^{(i)}(\mu) = \sum_{k=1}^{N_Z} \nu_i(k) e^{j\mu(k-1)} \quad (15)$$

The above form offers an alternative view of the decimation procedure where the spatial spectrum $V_{DFM}^{(i)}(\mu)$ is expressed in terms of the respective DTFT's of the filter and the i 'th modulated-transformed eigenvector. Defining the DTFT's

$$V_M^{(i)}(\mu) = \sum_{k=1}^N v_M^{(i)}(k) e^{j\mu(k-1)} \quad i > K \quad (16)$$

$$H(\mu) = \sum_{k=1}^L h(k) e^{j\mu(k-1)} \quad (17)$$

where $\mathbf{v}_M^{(i)} = \mathcal{M} \{ \mathbf{W} \mathbf{e}_i \} = \mathcal{M} \{ \mathbf{v}_i \}$ and $\mathbf{h} = [h(1), \dots, h(L)]^T$ is an $L \times 1$ vector composed of the entries of the filter impulse response. One can express $V_{DFM}^{(i)}(\mu)$ as

$$V_{DFM}^{(i)}(\mu) = \frac{1}{D} \sum_{\ell=0}^{D-1} H\left(\frac{\mu - 2\pi\ell}{D}\right) V_M^{(i)}\left(\frac{\mu - 2\pi\ell}{D}\right). \quad (18)$$

Notice that the form of (18) implies an integer-valued decimation factor D . Modifications for the more general cases where the sampling rate alteration is expressible as a non-reducible ratio of two integers, $D = M_D/M_I$, are readily incorporated into the procedure and will be addressed later in this section.

Assuming that aliasing effects are negligible, the $\ell = 0$ term (region surrounding baseband) dominates so that the following approximations hold

$$\begin{aligned} V_{DFM}^{(i)}(\mu) &\approx \frac{1}{D} H\left(\frac{\mu}{D}\right) V_M^{(i)}\left(\frac{\mu}{D}\right) \\ &\approx \frac{1}{D} \left[\sum_{\ell=0}^{D-1} H\left(\frac{\mu - 2\pi\ell}{D}\right) \right] V_M^{(i)}\left(\frac{\mu}{D}\right). \end{aligned} \quad (19)$$

Notice that the bracketted term in the latter approximation is simply the DTFT of the decimated impulse response of the filter sequence, $h_D(k) = h(Dk)$. Acceptance of the above approximations suggests that one is capable of removing the effects of the filter from the decimated null spectrum. Thus, we

may acquire the pertinent (signal) information associated with the eigenvector spectrum by viewing an alternate spectrum, denoted $V_{G^{-1}DFM}^{(i)}(\mu)$, given as

$$V_{G^{-1}DFM}^{(i)}(\mu) \approx \frac{V_{DFM}^{(i)}(\mu)}{\sum_{\ell=0}^{D-1} H\left(\frac{\mu-2\pi\ell}{D}\right)}. \quad (20)$$

Equivalently, the spectral division can be accomplished by deconvolving the decimated filter sequence out of the i 'th decimated eigenvector, $\mathbf{Z}\mathbf{e}_i$. As the deconvolution operation is also linear, one can simply deconvolve the decimated filter impulse response out from each column of $\mathbf{Z}$ in (13) to form a matrix $\mathbf{Z}'$. Denote the deconvolution operator as $\mathcal{G}^{-1}$ so that $\mathbf{Z}' = \mathcal{G}^{-1} \mathcal{DFM} \{ \mathbf{W} \}$. Recall that $\mathbf{Z}$ is an $N_Z \times N_b$ matrix where $N_Z = \lceil \frac{N+L-1}{D} \rceil$. Assuming that the deconvolution is exact, the size of $\mathbf{Z}'$ is $N_{Z'} \times N_b$, where $N_{Z'} = \lceil \frac{N+L-1}{D} \rceil - \lceil \frac{L}{D} \rceil + 1$. As the imperfect filtering introduces a small degree of aliasing, the deconvolution is not exact. Therefore, there exists a remainder term that must be considered such that the resultant process may not be causal. Numerically it is better to carry out the deconvolution by way of spectral division. In this case, the DTFT of a given column of $\mathbf{Z}$ is divided, point-wise, by the DTFT of the decimated filter sequence so that the inverse DTFT of the result provides the associated deconvolved column of $\mathbf{Z}'$. Depending upon the values of N and N_b , simulations have shown that possibly one or two extra points on either side of the $N_{Z'}$ points should be appended to each column of $\mathbf{Z}'$. A suitable criterion employed in simulation studies is that all points whose magnitudes greater than 5-10% of the maximum value should be included in $\mathbf{Z}'$.

Returning to the example cited earlier where the beamforming matrix corresponding to an $N = 128$ element ULA with $d = \lambda/2$ and $N_b = 8$ beams is operated on by an $L = 128$ length Hamming-windowed LPF and then maximally decimated, the dimensionality of the $\mathbf{Z}$ matrix is $N_Z \times N_b$, $N_Z = \lceil \frac{N+L-1}{D} \rceil = 16$. Assuming perfect deconvolution, the associated value of $N_{Z'}$ is 9. Adopting the 10% criteria in the selection of the row-dimension of $\mathbf{Z}'$, it was found that one extra row was needed. By way of spectral division employing the FFT/IFFT algorithms, the extra values were the last samples of the IFFT, which were wrapped-around to form the first row of $\mathbf{Z}'$.

In the case of non-integer decimation where the factor D is expressible as a ratio of two integers as $D = M_D/M_I$, a similar procedure can be implemented. Referring to Equation (18), the spectrum $V_M^{(i)}(\cdot)$ is replaced by the pre-filtered spectrum $V_{IM}^{(i)}(\cdot)$ defined by the DTFT of the i 'th transformed, modulated, and interpolated (M_I) noise eigenvector. The applicable decimation factor in (18) is then M_D . Note that the filter frequency-band specifications are selected to reflect the presence of the interpolation stage. As

a result, for the matrix $\mathbf{Z}$ defined by $\mathbf{Z} = \mathcal{D}_{M_D} \mathcal{F} \mathcal{I}_{M_I} \mathcal{M} \{ \mathbf{W} \}$, the N_b columns of the matrix $\mathbf{Z}'$ are found by deconvolving the decimated filter impulse response (decimated by the factor M_D) out from the corresponding columns of $\mathbf{Z}$.

The reduced row dimension of $\mathbf{Z}'$ relative to that of $\mathbf{Z}$ ultimately results in a computational savings for DOA estimation at the expense of a slight degradation in performance as will be shown in a subsequent section. The application of multirate eigenvector processing to the MUSIC algorithm is analyzed in Section C while an application to the TLS-ESPRIT algorithm is considered in Section D. The two algorithms are considered as representative of the class of eigenstructure DOA estimators. Extensions to other DOA estimation algorithms are easily accomplished.

C. Root-MUSIC Incorporating Multirate Eigenvector Processing

The multirate eigenvector technique is simply incorporated into the MUSIC algorithm of Schmidt [4]. As the transformed beamspace eigenvectors, $\mathbf{W}\mathbf{e}_i$, $i > K$, are orthogonal to the element-space manifold vectors corresponding to a signal arrival angle, $\mathbf{a}_N(\mu_k)$, $k \leq K$, the following condition holds

$$\mathbf{Z}\mathbf{e}_i = \mathcal{D} \mathcal{F} \mathcal{M} \{ \mathbf{W}\mathbf{e}_i \} \perp \mathcal{D} \mathcal{F} \mathcal{M} \{ \mathbf{a}_N(\mu_k) \} \quad i > K, k \leq K. \quad (21)$$

Assuming that the filter is ideal with a cutoff at the spatial location $\mu = \pi/D$, it is easily observed that the in-band signal nulls are preserved through the decimation operation such that

$$(\mathcal{D} \mathcal{F} \mathcal{M} \{ \mathbf{W}\mathbf{e}_i \})^H (\mathcal{D} \mathcal{F} \mathcal{M} \{ \mathbf{a}_N(\mu_k) \}) = (\mathbf{Z}\mathbf{e}_i)^H \mathbf{a}_{N_Z}(D\mu_k) = 0 \quad i > K, k \leq K. \quad (22)$$

If the filter is properly designed to limit aliasing yet pass all in-band signals, Equation (22) is a reasonably accurate approximation. Thus a suitable MUSIC null spectrum can be defined as

$$\mathcal{N}_{MU}(\mu) = \sum_{k=K+1}^{N_b} \left| \mathbf{a}_{N_Z}^H(D\mu) (\mathbf{Z}\hat{\mathbf{e}}_k) \right|^2 = \mathbf{a}_{N_Z}^H(D\mu) \mathbf{Z} \hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Z}^H \mathbf{a}_{N_Z}(D\mu), \quad (23)$$

where the estimated noise eigenvectors comprise $\hat{\mathbf{E}}_n = [\hat{\mathbf{e}}_{K+1} : \hat{\mathbf{e}}_{K+2} : \dots : \hat{\mathbf{e}}_{N_b}]$ and $\mathbf{a}_{N_Z}(D\mu)$ is an N_Z -dimensional element space manifold vector, where $N_Z = \lceil \frac{N+L-1}{D} \rceil$. Due to the Vandermonde structure of $\mathbf{a}_{N_Z}(D\mu)$, the spectral search for the estimation of the DOA angles can be converted to the rooting of polynomial a la Root-MUSIC. The true angles, $\hat{\theta}_k$, are then computed from $\hat{z}_k$ via $\hat{\theta}_k = \sin^{-1}(\arg\{\hat{z}_k\}\lambda/2\pi dD)$, $k \leq K$. The resulting algorithm is summarized below. Note that the Root-MUSIC algorithm employing the deconvolved version of $\mathbf{Z}$, $\mathbf{Z}' = \mathcal{G}^{-1} \mathcal{D} \mathcal{F} \mathcal{M} \{ \mathbf{W} \}$, is defined in a similar way where $\mathbf{Z}'$ and $N_{Z'}$ are substituted for $\mathbf{Z}$ and N_Z , respectively.

Summary of Root-MUSIC Application Algorithm

1. form the $N_Z \times N_b$ decimated-filtered-modulated beamforming matrix a-priori: $\mathbf{Z} = \mathcal{DFM} \{ \mathbf{W} \}$
2. EVD of $\hat{\mathbf{R}}_y = \sum_{m=1}^M \mathbf{y}(m)\mathbf{y}^H(m)/M$, where $\mathbf{y}(m) = \mathbf{W}^H \mathbf{x}(m)$ $m=1, \dots, M$.
3. estimate number of sources, K , and place $N_b - \hat{K}$ "smallest" eigenvectors as columns of $\hat{\mathbf{E}}_n$
4. with $p_k = \sum_{i=0}^k \mathbf{P}(N_Z - k + i, i + 1)$, $k=0, 1, \dots, N_Z-1$, where $\mathbf{P} = \mathbf{Z} \hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Z}^H$, and construct

$$p(z) = p_0 + p_1 z + \dots + p_{N_Z-1} z^{N_Z-1} + \dots + p_1^* z^{2N_Z-3} + p_0^* z^{2N_Z-2}$$

5. root $p(z)$, select $\hat{K}$ signal roots: $\hat{\theta}_k = \sin^{-1}(\arg\{\hat{z}_k\} \lambda / 2\pi d D)$ $k = 1, 2, \dots, \hat{K}$

Comparing the above prescription to that delineated in [7], the $N_Z \times N_b$ transformation $\mathbf{Z}$ replaces an $N_b \times N_b$ matrix $\mathbf{Q}$. The only disadvantage is a slight increase in computation as the polynomial to be rooted is slightly higher in order. However, the dimension N_Z can be selected to be only slightly larger than N_b if the deconvolution operation, $\mathcal{G}^{-1}$, is incorporated. The advantages of using the $\mathbf{Z}$ approach over that of $\mathbf{Q}$ are robustness to the computational accuracy of the rooting algorithm (due to the increase in angular separation between signal roots) and removal of the over-restrictive structural requirements of the type of beamformer employed. In addition, the condition number of $\mathbf{Q}$ is astronomical, between 10^5 and 10^{25} for the array parameters employed in generating Figure 1 while $\mathbf{Z}$ is extremely well-conditioned.

The accuracy of the $\mathbf{Z}$ and $\mathbf{Z}'$ transformations was assessed by observing the signal root locations when the ideal sample covariance is decomposed for use in the Root-MUSIC algorithm. The parameters of the array, beamformer, and decimator are those presented earlier in the example of Figures 2-6. The resulting root locations are shown in Figure 7 and the actual signal root locations for the two transformation types are included in the figure. The extremely accurate signal-root placement associated with the use of $\mathbf{Z}$ suggests that the orthogonality criterion $\mathbf{Z} \mathbf{e}_i \perp \mathbf{a}_{N_Z}(D\mu_k)$, $i > K$, $k \leq K$, is valid. Also note that the effects of the filter can be removed via deconvolution without appreciably affecting the performance of the algorithm as indicated by the locations of the roots associated with the use of $\mathbf{Z}'$.

To visualize the removal of the passband ripple as induced by the filter when deconvolution is employed, an example involving an FIR filter designed via the Parks-McClellan [12] algorithm with a "large" passband ripple was analyzed. In addition, to verify the validity of the general multirate procedure, an

$N = 90$ sensor array with $N_b = 6$ beams was used in a scenario involving decimation by a non-integer fraction $D = 11.25 = 45/4$ which is less than the maximum allowable value of $D_{max} = N/N_b = 15$. The filter was designed to be of length 270; note that the filtering is accomplished at the output of the interpolator stage ($D_I = 4$). The sub-maximal decimation factor allowed for a wide filter transition band, $(1/4)(5/N)\pi \leq \mu \leq (1/4)(11/N)\pi$, which, combined with a frequency band weighting favoring a high stopband attenuation, resulted in a 67 dB stopband attenuation with a 1.8 dB passband ripple. Plots of the spatial responses of the filter (dashed line) and interpolated beamformers (solid lines) are presented in Figure 8. The beamforming weight vectors were interpolated, by a factor of 4, to allow a visual comparison with the filter response curve.

Figure 9 shows the response of the $N_b = 6$ transformed, filtered, and decimated beamforming vectors along with the spectrum of the decimated filter. Note that the decimated filter magnitude spectrum (dashed curve) appears to follow the shape of the beam peaks.

The spectral MUSIC algorithm was employed with an ideal noise-only beamspace covariance matrix to compare the effects of using $\mathbf{Z}$ or $\mathbf{Z}'$. As this situation is effected using $\mathbf{E}_n \mathbf{E}_n^H = \mathbf{I}$, the MUSIC spectrum characterizes the imparted distortion to a white noise input spectrum by the inclusion of filtering or filtering followed by deconvolution. Figure 10 shows the MUSIC spatial spectra for a noise-only input employing the $\mathbf{Z}$ and $\mathbf{Z}'$ techniques. The results show that the deconvolution operation was effective in removing the filter shape from the spectrum leaving only a slight ripple that is representative of the finite spatial window associated with the beamformer. Again, the deviation at the edges of the spatial spectrum from the anticipated constant level is expected: the beamforming sector does not extend to the edge of the band at $\mu = \pi/D$.

D. TLS-ESPRIT Incorporating Multirate Eigenvector Processing

As with a previous beamspace Root-MUSIC algorithm [7], the beamspace ESPRIT formulation of Xu, et.al. [8] requires a rather restrictive specification on the form of the beamforming vectors. As we will see in this section, the ULA geometry allows an ESPRIT application of the transformed-decimated beamspace eigenvector approach of section B.

Given the N_b - K transformed and decimated noise eigenvectors, define an $N_Z \times ([N_Z - N_b] + K)$ matrix $\mathbf{E}_Z$, whose columns form a subspace that is orthogonal to that formed from the vectors $\mathbf{Z}\mathbf{e}_i$, $i > K$. An efficient means of computing $\mathbf{E}_Z$, is by way of a QR decomposition of $\mathbf{Z}\mathbf{E}_n$. Note that the

standard ESPRIT approach employs a matrix whose K columns span an estimate of the signal subspace; here we have a set of vectors in $\mathbf{E}_Z$, whose span encompasses the (decimated element-space) subspace, since $N_Z > N_b$. Assuming aliasing effects to be negligible, we have

$$\text{span} \{ \mathbf{a}_{N_Z}(D\mu_k), k = 1, \dots, K \} \subset \text{range} \{ \mathbf{E}_Z \}. \quad (24)$$

Although beamspace signal eigenvectors are not transformable to their element space counterparts, there is an alternative means of finding a set of vectors that are related to the beamspace signal eigenvectors and also span the orthogonal subspace of $\text{span} \{ \mathbf{Z}\mathbf{e}_i, i = K + 1, \dots, N_b \}$. The $N_Z \times N_b$ matrix transformation $\mathbf{Z}$ has full column rank so that the orthogonal subspace of $\text{span} \{ \mathbf{Z}\mathbf{e}_i, i = K + 1, \dots, N_b \}$ is expressible as a collection of $N_Z - N_b$ spanning vectors which are orthogonal to the columns of $\mathbf{Z}$ as well as K vectors lying in the column space of $\mathbf{Z}$. A permissible set of vectors which span the orthogonal subspace are the columns of

$$\mathbf{E}_Z = \left[\mathbf{Z}(\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{e}_1 : \dots : \mathbf{Z}(\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{e}_K : \beta_1 : \dots : \beta_{N_Z - N_b} \right], \quad (25)$$

where $\{ \beta_1, \dots, \beta_{N_Z - N_b} \}$ is a set of vectors that span the subspace orthogonal to the column space of $\mathbf{Z}$. Notice that the set of vectors in (25) are not orthogonal but still are adequate for use in ESPRIT. In addition to the computational savings in avoiding a QR-decomposition, construction of $\mathbf{E}_Z$, according to (25) also allows one to derive the theoretical angle estimation performance using available asymptotic expressions for the beamspace eigenvector statistics as we shall see in Section 4.

We will return to the “over-specification” issue of the decimated signal subspace in this section and show that judicious beamforming and filter design allows for proper operation of a suitably defined ESPRIT algorithm. Assuming that the beamforming and filtering operations produce little aliasing effects so that Equation (24) is a reasonably accurate approximation, we may define a TLS-ESPRIT procedure to estimate the directions of the K signal arrivals based upon the Vandermonde form of $\mathbf{a}_{N_Z}(\cdot)$. The algorithm is summarized as follows.

Summary of TLS-ESPRIT Application Algorithm

1. form $N_Z \times N_b$ decimated-filtered-modulated beamforming matrix a priori: $\mathbf{Z} = \mathcal{DFM} \{ \mathbf{W} \}$.

Form a set of vectors, $\beta_i, i = 1, \dots, N_Z - N_b$, that span a subspace orthogonal to $\text{range} \{ \mathbf{Z} \}$.

2. EVD of $\hat{\mathbf{R}}_y = \sum_{m=1}^M \mathbf{y}(m) \mathbf{y}^H(m) / M$, where $\mathbf{y}(m) = \mathbf{W}^H \mathbf{x}(m)$, $m = 1, \dots, M$.

3. estimate number of sources, K , and form the matrix $\mathbf{E}_Z$, composed of vectors which span the estimated decimated signal subspace: $\mathbf{E}_Z = \left[\mathbf{Z}(\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{e}_1 : \dots : \mathbf{Z}(\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{e}_{\hat{K}} : \beta_1 : \dots : \beta_{N_Z - N_b} \right]$.
4. form $(N_Z - 1) \times 2(N_Z - N_b + \hat{K})$ matrix $\hat{\mathbf{E}}_{xy} \doteq \left[\hat{\mathbf{E}}_1 | \hat{\mathbf{E}}_2 \right]$ where $\hat{\mathbf{E}}_1$ and $\hat{\mathbf{E}}_2$ are the first and last $N_Z - 1$ rows of $\hat{\mathbf{E}}_Z$, and compute the $2(N_Z - N_b + \hat{K}) \times 2(N_Z - N_b + \hat{K})$ EVD $\hat{\mathbf{E}}_{xy}^H \hat{\mathbf{E}}_{xy} = \hat{\mathbf{E}} \hat{\Sigma} \hat{\mathbf{E}}^H$
5. partition $\hat{\mathbf{E}}$ into $(N_Z - N_b + \hat{K}) \times (N_Z - N_b + \hat{K})$ submatrices: $\hat{\mathbf{E}} = \begin{bmatrix} \hat{\mathbf{E}}_{11} & \hat{\mathbf{E}}_{12} \\ \hat{\mathbf{E}}_{21} & \hat{\mathbf{E}}_{22} \end{bmatrix}$.
6. compute the $(N_Z - N_b + \hat{K}) \times (N_Z - N_b + \hat{K})$ EVD $-\hat{\mathbf{E}}_{12} \hat{\mathbf{E}}_{22}^{-1} = \mathbf{T} \Phi \mathbf{T}^{-1}$
7. for those $\hat{K}$ nearly unit-magnitude eigenvalues $\lambda_i = \Phi_{ii}$, estimate the corresponding signal arrival directions as $\hat{\theta}_k = \sin^{-1}(\text{angle}\{\lambda_i\} \lambda / 2\pi d D)$

Location of Extraneous Roots Created by Filtering

A major concern is that the extra column dimension of $\mathbf{E}_Z$, over the K -dimensional signal subspace will result in the declaration of ambiguous signals. First of all, note that we've already at this point estimated the number of signal arrivals. Here, an argument is presented that suggests that the extraneous roots will not lie near the unit circle. This claim is also verified via a simulation example presented in Section 4.

First, note that in the case of ideal decimation where the filter exhibits a perfect low-pass nature, Equation (24) applies. From the summary above, recall that the k 'th diagonal element of Φ has unit magnitude, $\Phi_{kk} = e^{jD\mu_k}$. Now consider the inclusion of a linear filter in the decimation operation. The aliasing effects caused by decimation will result in an ESPRIT signal eigenvalue that will not have a unit magnitude characteristic, even if the ideal beamspace covariance matrix is available. However, a judicious filter and beamformer design will result in an approximate unit-magnitude eigenvalue characteristic.

In addition to ESPRIT eigenvalues directly corresponding to signals, assume that there is an extraneous unit magnitude eigenvalue, λ_* , i.e.,

$$\Gamma_1 \mathbf{E}_Z - \lambda_* \Gamma_2 \mathbf{E}_Z = 0.$$

This suggests that, in addition to the Vandermonde components arising from the true signals, a Vandermonde vector corresponding to the angle $D\mu_*$ also lies in the decimated signal subspace. Equivalently,

this implies $\mathbf{a}_{N_z}^H(D\mu_*)$ is orthogonal to the range of $\mathbf{Z}\mathbf{E}_z$, so that

$$\mathbf{a}_{N_z}^H(D\mu_*) [\mathbf{Z}\mathbf{E}_n\mathbf{E}_n^H\mathbf{Z}^H] \mathbf{a}_{N_z}(D\mu_*) = 0.$$

Thus the spectrum of every transformed and decimated beamspace noise eigenvector exhibits a null at the spatial location $D\mu_*$. By design, there are no common in-band beamformer nulls and the filter response is also non-zero across the spatial sector of interest so that λ_* must be an ESPRIT eigenvalue associated with a signal arrival.

Refer to Figure 6 where a Hamming-weighted LPF was employed as the decimation filter applied to noise eigenvectors generated from an $N_b = 8$ spatial DFT beamformer. The filter has an associated spatial response that is relatively flat across the subband and there are no common in-band nulls in the set of beamforming vectors. Note that the only nulls in the MUSIC null spectrum correspond to signal arrival angles. The behavior at the edges of the band is expected from the presence of a root near π at a radius of 0.9 as shown in Figure 7. As a result of the relationship between the ESPRIT eigenvalues and the roots generated from Root-MUSIC, it is anticipated that an extraneous ESPRIT eigenvalue will lie in the complex plane near the unit circle at π and that all other non-signal eigenvalues will be sufficiently displaced from the unit circle. This is acceptable since these eigenvalues are discarded anyway as a result of previous discussion. In summary, an ESPRIT eigenvalue with a nearly unit magnitude suggests the presence of a signal at an associated spatial angle as long as the filter and beamforming vectors are judiciously designed.

4. Theoretical Performance Analysis

As noted in Section 2, the use of conjugate centro-symmetric beamforming architectures in conjunction with uniformly-spaced linear arrays with phase referencing at the array center results in a purely real-valued beamspace manifold. The real-valued property of the manifold allows one to decompose only the real part of the sample covariance matrix to determine the signal or noise subspaces as noted in [7, 9]. In addition to the obvious computational advantages of a real-valued decomposition, a performance benefit is realized through the decorrelation of correlated signals as taking the real part of the beamspace sample covariance matrix is equivalent to applying a single forward/backward average in element-space prior to beamforming [7, 9]. In uncorrelated signal environments, the real and complex-valued procedures result in similar performances in terms of estimation variance; however, the bias is, in general, smaller with the use

of real covariance processing. As a result of these advantages as well as the applicability of either approach with regard to the Root-MUSIC and ESPRIT based procedures incorporating eigenvector decimation, we derive the theoretical performance of the two algorithmic approaches for the case of real-covariance processing. Extension for the case of complex processing is readily determined.

Define $\Delta \mathbf{e}_i = \hat{\mathbf{e}}_i - \mathbf{e}_i$, $i = 1, \dots, K$, as the error in the i 'th eigenvector due to the use of a sample estimate of the covariance matrix where $\hat{\mathbf{e}}_i$ and $\mathbf{e}_i$ are the i 'th eigenvectors obtained from the beamspace sample covariance matrix and the ideal covariance, respectively, under some common uniqueness criterion. The distribution of $\Delta \mathbf{e}_i$ was shown to be asymptotically Gaussian with zero mean and covariance [9]

$$\mathcal{E}\{M \Delta \mathbf{e}_k \Delta \mathbf{e}_\ell^T\} = \sum_{\substack{m=1 \\ m \neq k}}^{N_b} \sum_{\substack{n=1 \\ n \neq \ell}}^{N_b} \frac{\Gamma_{mnlk}}{(\lambda_k - \lambda_m)(\lambda_\ell - \lambda_n)} \mathbf{e}_m \mathbf{e}_n^T, \quad k, \ell = 1, \dots, K \quad (26)$$

$$\begin{aligned} \Gamma_{mnlk} = \frac{1}{2} \{ & \lambda_k \lambda_\ell \delta_{m\ell} \delta_{nk} + \lambda_k \lambda_m \delta_{mn} \delta_{k\ell} + (\mathbf{e}_m^T \mathbf{R}_I \mathbf{e}_\ell)(\mathbf{e}_k^T \mathbf{R}_I \mathbf{e}_n)(1 - \delta_{m\ell})(1 - \delta_{kn}) \\ & + (\mathbf{e}_m^T \mathbf{R}_I \mathbf{e}_n)(\mathbf{e}_k^T \mathbf{R}_I \mathbf{e}_\ell)(1 - \delta_{mn})(1 - \delta_{k\ell}) \} \end{aligned} \quad (27)$$

$$\mathbf{R}_I = \mathcal{I}m\{\mathbf{R}\} = \mathbf{B} \mathcal{I}m\{\mathbf{P}_S\} \mathbf{B}^T. \quad (28)$$

To allow for the use of previous MUSIC [6, 9] and ESPRIT [15] performance analyses, it is assumed that the aliasing effects are negligible. As noted earlier, the assumption is valid when the decimation operation includes a judiciously designed filter or the use of front-end beamformers with very low out-of-band responses. The condition may be verified by observing the placement of the (signal) MUSIC roots/ESPRIT eigenvalues in the case of a known ideal covariance. Once again, the Root-MUSIC signal locations for the motivational example shown in Figure 7 confirm the validity of the assumption, particularly in the case where deconvolution is not employed.

A. Performance Analysis of Root-MUSIC Formulation

The asymptotic variance of the Root-MUSIC estimator is readily obtained using available results when assuming orthogonality between the transformed-filtered-decimated beamspace noise eigenvectors and the decimated element-space manifold, i.e., $\mathbf{Z} \mathbf{e}_i \perp \mathbf{a}_{N_Z}(\theta_k) \quad k = 1, \dots, K \quad i = K + 1, \dots, N_b$. By observing that the spectral and Root-MUSIC formulations offer the same asymptotic performance in terms of the variance as shown in [6], the expression for the spectral MUSIC estimate variance employing real-covariance processing in [9] can be easily amended to the case at hand. Specifically, the null spectrum

can be written as

$$\begin{aligned}\mathcal{N}_{MU}(\theta) &= \mathbf{a}_{N_Z}^H(\theta) \left\{ \sum_{i=K+1}^{N_b} (\mathbf{Z}\hat{\mathbf{e}}_i)(\mathbf{Z}\hat{\mathbf{e}}_i)^H \right\} \mathbf{a}_{N_Z}(\theta) \\ &= \mathbf{a}_{N_b}^H(\theta) \mathbf{Z} \left\{ \mathbf{I}_{N_b} - \sum_{i=K+1}^{N_b} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^T \right\} \mathbf{Z}^H \mathbf{a}_{N_b}(\theta).\end{aligned}\quad (29)$$

Observing the results in [9], the asymptotic variance of the Root-MUSIC estimator is easily shown to be expressed as

$$\mathcal{A}\text{Var}\{\hat{\theta}_i\} = \frac{1}{M \dot{\mathbf{a}}_{N_Z}^H(\theta_i) \mathbf{Z} \mathbf{E}_n \mathbf{E}_n^T \mathbf{Z}^H \dot{\mathbf{a}}_{N_Z}(\theta_i)} \sum_{k=1}^K \frac{\lambda_k \sigma_n^2}{(\lambda_k - \sigma_n^2)^2} \left| \mathbf{e}_k^T \mathbf{Z}^H \mathbf{a}_{N_Z}(\theta_i) \right|^2 \quad i = 1, \dots, K, \quad (30)$$

where M is the number of snapshots, $\dot{\mathbf{a}}_{N_Z}(\theta_i)$ is the derivative of $\mathbf{a}_{N_Z}(\theta)$ with respect to θ evaluated at $\theta = \theta_i$, and $(\lambda_k, \mathbf{e}_k)$, $k=1, \dots, K$, are the signal eigenvalues and corresponding eigenvectors of the real part of the ideal beamspace covariance matrix.

B. Performance Analysis of ESPRIT Formulation

The alternate expression in Equation (25) for the decimated signal subspace involving the transformed beamspace signal eigenvectors and a non-random basis for the orthogonal subspace of the columns of $\mathbf{Z}$ allows for an asymptotic analysis of the ESPRIT formulation. The error in the matrix whose columns form a basis for the decimated signal subspace, $\Delta \mathbf{E}_{Z_s}$, is simply

$$\Delta \mathbf{E}_{Z_s} = \left[\mathbf{Z}(\mathbf{Z}^H \mathbf{Z})^{-1} \Delta \mathbf{e}_1 : \dots : \mathbf{Z}(\mathbf{Z}^H \mathbf{Z})^{-1} \Delta \mathbf{e}_K : \mathbf{0}_{N_Z \times (N_Z - N_b)} \right]. \quad (31)$$

In this form, the error is only a function of the error in the eigenvectors associated with signal eigenvalues of the beamspace covariance. This allows for an asymptotic variance analysis similar to that found in [15]. The analysis in [15] is valid for the Least-Squares (LS) and Total Least-Squares (TLS) versions of ESPRIT. The variance analysis, for real beamspace covariance processing, is included in Appendix A. The asymptotic variance associated with the i 'th angle estimate in the case of uncorrelated sources is

$$\begin{aligned}\mathcal{E}\{(\Delta \theta_i)^2\} &= \left[\frac{\lambda}{2\pi d D \cos \theta_i} \right]^2 \frac{1}{2M} \left[\sum_{k=1}^K \frac{\sigma_n^2 \lambda_k}{(\lambda_k - \sigma_n^2)^2} \left| \text{Im} \{x_i(k) \mathbf{E}_n^T \boldsymbol{\alpha}_i\} \right|^2 \right. \\ &\quad \left. + \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \frac{\lambda_k \lambda_\ell}{(\lambda_k - \lambda_\ell)^2} \left(\text{Im} \{x_i(\ell) \mathbf{e}_k^T \boldsymbol{\alpha}_i\}^2 - \text{Im} \{x_i(\ell) \mathbf{e}_\ell^T \boldsymbol{\alpha}_i\} \text{Im} \{x_i(k) \mathbf{e}_k^T \boldsymbol{\alpha}_i\} \right) \right],\end{aligned}\quad (32)$$

$$\boldsymbol{\alpha}_i = (\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{Z}^H [\boldsymbol{\Gamma}_1 - \mathbf{z}_i^* \boldsymbol{\Gamma}_2]^H ([\boldsymbol{\Gamma}_1 \mathbf{E}_{Z_s}]^\dagger)^H \mathbf{q}_i, \quad (33)$$

$$\mathbf{E}_n = \left[\mathbf{e}_{K+1} : \dots : \mathbf{e}_{N_b} \right], \quad (34)$$

where $\dagger$ denotes pseudoinverse, $\mathbf{x}_i$ and $\mathbf{q}_i$ are the right and left eigenvectors associated with the i 'th (signal) eigenvalue of $\mathbf{F} = (\mathbf{\Gamma}_1 \mathbf{E}_{Z_s})^\dagger \mathbf{\Gamma}_2 \mathbf{E}_{Z_s}$, and $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ are $(N_Z - 1) \times N_Z$ matrices that select the first and last $N_Z - 1$ rows of a matrix with N_Z rows, respectively. Note that the expressions contained in Appendix A may be applied to the more general case of correlated signals; only the result for the uncorrelated signal scenario is summarized here due to its simpler form.

5. Computer Simulations

A number of computer simulations were conducted to assess the validity of the noise eigenvector transformation/decimation techniques with regard to angle estimation. Specifically, the theoretical and empirical standard deviations of the Root-MUSIC and TLS-ESPRIT estimators were compared in a variety of source/processing scenarios. Also, the performance of the decimation approach was compared to the stochastic Cramer-Rao Lower Bound [4, 16].

Common to all experiments, 600 trials were employed to derive the empirical results and only $M = 16$ snapshots were used to estimate the beamspace covariance matrix. Although this situation can hardly be classified as asymptotic in the number of snapshots, the theoretical performance curves were observed to compare rather closely to the derived experimental results.

The empirical standard deviations were computed in a variety of scenarios involving one or two uncorrelated, closely-spaced signals. A MUSIC root or ESPRIT eigenvalue was classified as arising from a signal if the root/eigenvalue location was within a 0.15 radial distance from the unit circle and lying in an angular (decimated) region encompassing 85% of the unit circle, i.e., in the region $[-0.85\pi, 0.85\pi]$. All trial runs, including those unresolved situations where only one signal was observed in the neighborhood of a signal pair, were used to compute the location statistics.

Experiment 1: The simulation parameters of this experiment associated with the array, beamformer, and decimation components are similar to those outlined in the example of Section 3, namely, an $N = 128$ element ULA with half-wavelength spacing was operated on with an $N_b = 8$ channel spatial DFT beamformer. The spatial window was centered at broadside so that the spatial region $-N_b/N \leq \sin \theta \leq N_b/N$ was probed. An $L = 128$ length Hamming-weighted low-pass filter was employed in the decimation procedure configured for maximal decimation, i.e., $D = N/N_b$.

Two half-Rayleigh spaced signals of equal power were embedded in additive complex Gaussian noise so that a sensor level 10 dB SNR was achieved. To assess the effects of signal placement within the spatial

beamforming sector on the estimation variance, the center of the signal set was shifted from baseband ($\sin \theta = 0$) to the edge of the window ($\sin \theta = 8/N$). The empirical standard deviation of the two Root-MUSIC angle estimators, i.e., those formed using the matrix $\mathbf{Z}$ as well as the deconvolved version $\mathbf{Z}'$, were computed. Note that the dimension of $\mathbf{Z}$ was 16×8 while $\mathbf{Z}'$ was formed by adding one additional (remainder) row to the required $(N_b + 1) \times N_b$ matrix to form a 10×8 eigenvector transformation. The results are shown, along with the theoretical prediction as obtained from Equation (30) and the stochastic Cramer-Rao Lower bound [4, 16] in Figure 11.

Several comments relating to Figure 11 are in order. Although the number of snapshots is relatively small, the theoretical performance curve is still a reasonably accurate representation of the empirically derived result. The rippled nature of the variance curves is due to the limited number of beams that are implemented in the approach. This characteristic is the result of a varying spatial power gain as similar to that depicted in Figure 10. As noted in [7, 9], the degradation in performance near the band edge suggests the need for sub-band overlap if one is interested in the detection and localization of all signals across the visible spatial spectrum. The variance of the estimate at the extreme right edge is not shown as the experimental and theoretical curves exhibit an exponential rise. In the central region of the band, however, the eigenvector transformation-decimation technique is seen to produce an accurate estimate in this Root-MUSIC formulation as evidenced by the closeness of the results to the Cramer-Rao Bound. Note that the curves related to the theoretical variance associated with the use of $\mathbf{Z}$ and the Cramer-Rao Bound overlap.

Experiment 2: Employing the same decimation transformations as in Experiment 1, the variance of the Root-MUSIC estimators were observed for a varying SNR for two signals located at 10.6° and 11.5° , as used in the motivational example of Figures 2 through 7. The empirical and theoretical standard deviations were computed and are depicted in Figure 12.

Note that the theoretically derived curve, defined for the 16×8 transformation $\mathbf{Z}$, closely tracks the corresponding empirical counterpart at moderate to high SNR values. The deviation at the lower SNR values is attributed to the signal-merging effects in the resolution threshold regime of operation as noted in [9]. Although the stochastic Cramer-Rao Bound is based upon the statistics of the available beamspace data and does not assume the presence of any sub-optimal techniques such as decimation, the Root-MUSIC procedure incorporating decimation is readily observed to essentially offer the optimum

performance associated with un-biased estimators. Also, the similarity between the empirical variance curves corresponding to the competing approaches ($\mathbf{Z}$ versus $\mathbf{Z}'$) suggests that the computational savings associated with the smaller Root-MUSIC polynomial via the use of $\mathbf{Z}'$ is not obtained at the expense of a higher estimation variance. In fact, simulations have shown that the estimation variance is usually smaller for decimation architectures incorporating deconvolution. However, the imperfect deconvolution usually results in an induced estimate bias as will be observed in Experiment 3.

Experiment 3: The main purpose of this experiment is to show that the filtering operation in the decimator may not be warranted in certain situations. A single signal was positioned at 1° and the bias performance was studied for the use of two beamforming architectures. In one situation, $N_b = 6$ DFT beams were formed from an $N = 36$ element ULA. The beamspace to element-space eigenvector transformation was configured for maximal decimation, $D = 6$, with and without the use of a Parks-McClellan equiripple FIR filter exhibiting approximately 50 dB attenuation in the stopband region. In the other beamforming scenario, a practical application of $N_b = 6$ Taylor weighted beams [17], exhibiting a 50 dB sidelobe level, were spaced at the half-power points and employed in a similar scheme involving the use/absence of additional filtering in the decimation operation. Note that the latter approach will produce an angle estimate exhibiting a substandard resolution ability due to the attendant wider mainlobes relative to DFT beams. However, this methodology is often required in practice to reduce the deleterious effects of sidelobe clutter, i.e., the masking of signals within a given beam by a strong clutter signal in the sidelobes of the beam. The beam spacing/aperture weighting associated with this case results in no common spatial nulls amongst the beam set so that the application of past beamspace MUSIC [7] and ESPRIT [8] formulations is precluded.

The empirically derived mean location estimates were determined for a varying SNR for various schemes incorporating the two beamforming architectures and are plotted in Figure 13. Again, the purpose here is not to compare the two beamforming approaches, rather, it is to observe the effects on performance of the inclusion of a filter in the decimation operation. Also, the inclusion of a filter increases the order of the polynomial to be rooted thereby increasing computation and creating extraneous roots. With reference to Figure 13, note that the use of a filtering operation in the decimator with no additional deconvolution stage results in essentially an unbiased estimator for both beamforming architectures. As observed in the results, the Taylor-based sensor weighting provides sufficient attenuation so that a

negligible aliasing effect is incurred, i.e., the induced estimation bias is small. However, with the filter incorporated into the decimation operation, the imperfect deconvolution stage imparts a small bias of -0.02° . Thus the filtering operation is unnecessary as evidenced in the bias plot and a smaller standard deviation should be realized on account of the smaller dimension of the resulting Root-MUSIC polynomial.

Essentially the opposite is observed for the case of unweighted spatial DFT beamforming. Here the sidelobe levels are large so that aliasing effects are present as evidenced by the top curve indicating a 0.05° bias in the unfiltered mode of operation. With filtering as well as a deconvolution stage included in the decimation operation, a smaller bias of 0.025° is realized. The need for filtering is evident from observing the required dimension of the transformation $\mathbf{Z}'$. Comparing the necessary row dimension of the decimation transformation incorporating deconvolution, $\mathbf{Z}'$, for the unweighted DFT and Taylor beamformers, the required sizes were 10×6 and 7×6 , respectively. These required sizes were determined according to the criteria discussed in Section 3.

Experiment 4: In this experiment, we test the validity of the TLS-ESPRIT formulation of the noise eigenvector transformation-decimation procedure and verify the theoretical variance expression of Section 4i, Equation (32). The source/processing parameters are the same as those of Experiment 2.

The theoretical and empirical standard deviation were computed over a varying SNR and the results are depicted in Figure 14. The results show that the performance predictor of Section 4 accurately tracks the empirical results. Also, the variance associated with the decimation architecture incorporating a filter deconvolution stage outperforms the “undeconvolved” counterpart. To verify the conjecture that the quiescent locations of the extraneous eigenvalues are sufficiently away from the unit circle, the ESPRIT eigenvalues were calculated in the absence of noise and plotted in Figure 15. Note that only the eigenvalues interior to the unit circle are plotted as the closest exterior eigenvalue is located at a radius of 5.4 (associated with the $\mathbf{Z}$ transformation). Referring to Figure 15, in the absence of deconvolution, two “signal” eigenvalues appear at the correct location and the eigenvalue closest to the unit circle of the remaining is located at a radius of 0.62 and an angle very near π . When deconvolution is incorporated, the closest non-signal eigenvalue is located at π at a radius of 0.09. However, the signal eigenvalues exhibit a small bias at the perceived (translated) angular locations of 10.587° and 11.465° .

6. Conclusions/Remarks

We have developed a novel approach to angle estimation in the beamspace domain. The approach

offers a computationally attractive and non-restrictive procedure relative to the type of beamformer employed that is easily implemented in the MUSIC and ESPRIT algorithms. Theoretical expressions for the estimate variance were obtained in an asymptotical analysis and confirmed in a variety of simulations. Although the technique was applied to the uniform linear array geometry, an extension to a two-dimensional array to provide simultaneous azimuth/elevation angle estimates is evident and currently under investigation.

Appendix: Asymptotic Variance of ESPRIT Formulation

Given that z_i is a (signal) unit-magnitude eigenvalue of the matrix

$$\mathbf{F} = (\mathbf{\Gamma}_1 \mathbf{E}_{Z_s})^\dagger (\mathbf{\Gamma}_2 \mathbf{E}_{Z_s}) = [(\mathbf{\Gamma}_1 \mathbf{E}_{Z_s})^H (\mathbf{\Gamma}_1 \mathbf{E}_{Z_s})]^\dagger (\mathbf{\Gamma}_1 \mathbf{E}_{Z_s})^H (\mathbf{\Gamma}_2 \mathbf{E}_{Z_s}), \quad (35)$$

with $\mathbf{x}_i$ and $\mathbf{q}_i$ the corresponding right and left eigenvectors, Rao and Hari [15] showed that, to $o(M^{-1})$,

$$\Delta z_i = \mathbf{q}_i^H \Delta \mathbf{F} \mathbf{x}_i. \quad (36)$$

The error in $\mathbf{F}$, $\Delta \mathbf{F}$, due to the finite sample estimation of the beamspace covariance matrix is

$$\Delta \mathbf{F} = (\mathbf{\Gamma}_1 \mathbf{E}_{Z_s})^\dagger (\mathbf{\Gamma}_2 \Delta \mathbf{E}_{Z_s}) - (\mathbf{\Gamma}_1 \mathbf{E}_{Z_s})^\dagger (\mathbf{\Gamma}_1 \Delta \mathbf{E}_{Z_s}) \mathbf{F}, \quad (37)$$

which is applicable to either the Least Squares (LS) or Total Least Squares (TLS) versions of ESPRIT.

Substituting the form of $\Delta \mathbf{E}_{Z_s}$ in Equation (31) into Equation (36), one obtains

$$\mathcal{E}\{|\Delta z_i|^2\} = \boldsymbol{\alpha}_i^H \left[\sum_{k=1}^K \sum_{\ell=1}^K x_i(k) x_i^*(\ell) \mathcal{E}\{\Delta \mathbf{e}_k \Delta \mathbf{e}_\ell^T\} \right] \boldsymbol{\alpha}_i \quad (38)$$

$$(z_i^*)^2 \mathcal{E}\{(\Delta z_i)^2\} = \boldsymbol{\alpha}_i^H \left[\sum_{k=1}^K \sum_{\ell=1}^K x_i(k) x_i(\ell) \mathcal{E}\{\Delta \mathbf{e}_k \Delta \mathbf{e}_\ell^T\} \right] \boldsymbol{\alpha}_i^*, \quad (39)$$

where $\boldsymbol{\alpha}$ and the signal eigenvector error statistics were stated in Equations (33) and (26), respectively.

Following [15], these quantities are then substituted into

$$\mathcal{E}\{(\Delta \theta_i)^2\} = \left[\frac{\lambda}{2\pi d D \cos \theta_i} \right]^2 \left[\frac{\mathcal{E}\{|\Delta z_i|^2\} - \text{Re}\{(z_i^*)^2 \mathcal{E}\{(\Delta z_i)^2\}\}}{2} \right]. \quad (40)$$

to yield the desired theoretical asymptotic estimation variance.

In the case of uncorrelated signals, the asymptotic error in the signal subspace eigenvectors become

$$\mathcal{E}\{\Delta \mathbf{e}_k \Delta \mathbf{e}_\ell^T\} = \frac{\delta_{k\ell}}{2M} \sum_{\substack{m=1 \\ m \neq k}}^{N_b} \frac{\lambda_k \lambda_m}{(\lambda_k - \lambda_m)^2} \mathbf{e}_m \mathbf{e}_m^T - \frac{(1 - \delta_{k\ell})}{2M} \frac{\lambda_k \lambda_\ell}{(\lambda_k - \lambda_\ell)^2} \mathbf{e}_\ell \mathbf{e}_k^T. \quad (41)$$

After substituting and simplifying (the algebraic details are omitted here due to space limitations), the asymptotic variance of the ESPRIT angle estimate for uncorrelated sources reduces to

$$\begin{aligned} \mathcal{E}\{(\Delta\theta_i)^2\} = & \left[\frac{\lambda}{2\pi dD \cos \theta_i} \right]^2 \frac{1}{2M} \left[\sum_{k=1}^K \frac{\sigma_n^2 \lambda_k}{(\lambda_k - \sigma_n^2)^2} \left| \text{Im} \{x_i(k) \mathbf{E}_n^T \boldsymbol{\alpha}_i\} \right|^2 \right. \\ & \left. + \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \frac{\lambda_k \lambda_\ell}{(\lambda_k - \lambda_\ell)^2} \left(\text{Im} \{x_i(\ell) \mathbf{e}_k^T \boldsymbol{\alpha}_i\}^2 - \text{Im} \{x_i(\ell) \mathbf{e}_\ell^T \boldsymbol{\alpha}_i\} \text{Im} \{x_i(k) \mathbf{e}_k^T \boldsymbol{\alpha}_i\} \right) \right], \quad (42) \end{aligned}$$

where $\mathbf{E}_n$ is an $N_b \times (N_b - K)$ matrix composed of the noise eigenvectors associated with the ideal beamspace covariance.

References

- [1] G. Bienvenu and L. Kopp, "Decreasing High Resolution Method Sensitivity by Conventional Beamforming Preprocessing," in *Proc. of 1984 IEEE Int'l Conf. on Acoust., Speech, and Signal Process.*, pp. 33.2.1-33.2.4, April 1984.
- [2] C. L. Byrne and A. K. Steele, "Sector-Focussed Stability for High-Resolution Array Processing," in *Proc. of 1987 IEEE Int'l Conf. on Acoust., Speech, and Signal Process.*, pp. 54.11.1-54.11.4, April 1987.
- [3] H. Lee and M. Wengrovitz, "Resolution Threshold of Beamspace MUSIC for Two Closely-Spaced Emitters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-38, no. 9, pp. 1545-1559, Sept. 1990.
- [4] R. O. Schmidt, "A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation," Ph.D. dissertation, Stanford University, Stanford, CA, 1981.
- [5] A. J. Barabell, "Improving the Resolution Performance of Eigenstructure-Based Direction Finding Algorithms," in *Proc. of 1983 IEEE Int'l Conf. on Acoust., Speech, and Signal Process.*, pp. 336-339, May 1983.
- [6] B. D. Rao and K. V. S. Hari, "Performance Analysis of Root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-37, no. 12, pp. 1939-1949, December 1989.
- [7] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Beamspace Root-MUSIC," *IEEE Trans. on Signal Processing*, vol. 41, no. 1, pp. 344-364, Jan 1993.
- [8] G. Xu, S. D. Silverstein, R. Roy, and T. Kailath, "Parallel implementation and Performance of Beamspace ESPRIT," in *Proc. of 1984 IEEE Int'l Conf. on Acoust., Speech, and Signal Process.*, pp. 1497-1500, April 1991.
- [9] G. M. Kautz and M. D. Zoltowski, "Performance Analysis of MUSIC Employing Conjugate Symmetric Beamformers," Submitted to *IEEE Trans. on Signal Processing*, May 1993.
- [10] A. Paulraj, R. Roy, and T. Kailath, "Estimation of Signal Parameters via Rotational Invariance Techniques - ESPRIT," in *Proc. 19-th Asilomar Conf. on Signals, Systems and Computers*, pp. 83-89, Nov. 1985.
- [11] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, pp. 387-392, April 1985.
- [12] J. G. Proakis and D. G. Manolakis, *Introduction to Digital Signal Processing*, Macmillan, 1988.
- [13] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Second ed., John Wiley & Sons, 1984.

- [14] M. Kaveh and A. J. Barabell, "The Statistical Performance of the MUSIC and the Minimum-Norm Algorithms in Resolving Plane Waves in Noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 2, pp. 331-341, April 1986.
- [15] B. D. Rao and K. V. S. Hari, "Performance Analysis of ESPRIT and TAM in Determining the Direction of Arrival Plane Waves in Noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-37, no. 12, pp. 1990-1995, December 1989.
- [16] P. Stoica and A. Nehorai, "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-38, no. 10, pp. 1783-1795, October 1990.
- [17] T. T. Taylor, "Design of Line-Source Antennas for Narrow Beamwidth and Low Side Lobes," *IRE Transactions on Antennas and Propagation*, pp. 16-28, January 1955.

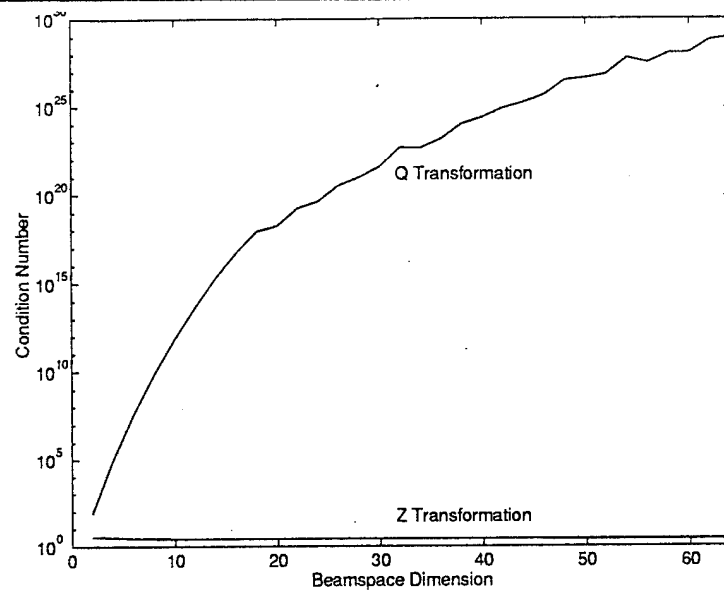


Figure 1: Condition Number vs. Number of Beams

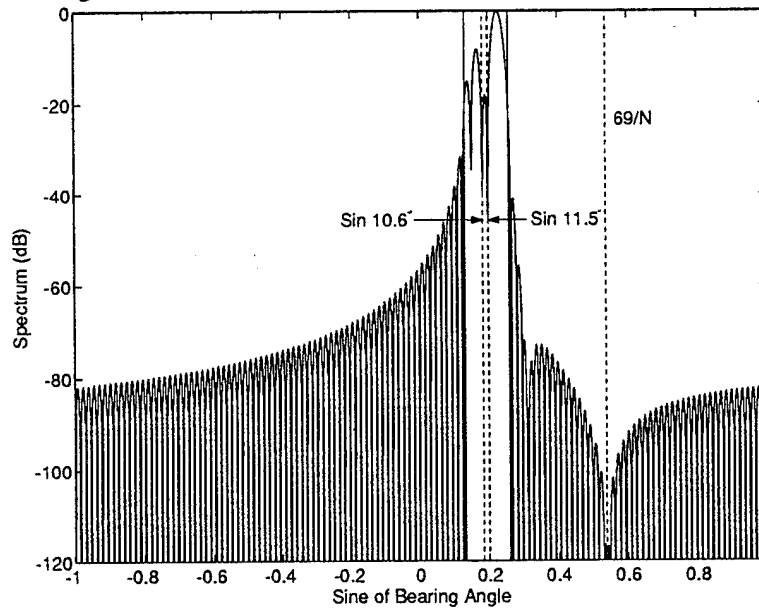


Figure 2: Spectrum of a Transformed Noise Eigenvector

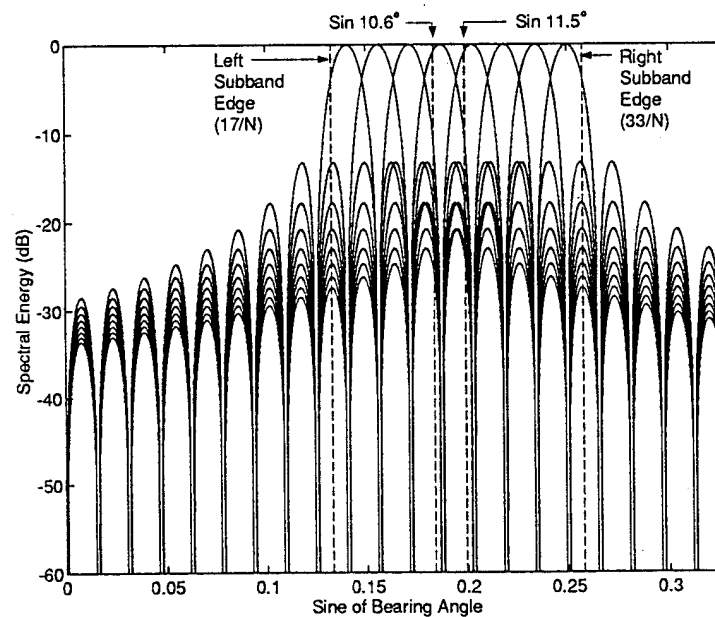


Figure 3: Angular Responses of Eight Successive DFT Beamforming Vectors

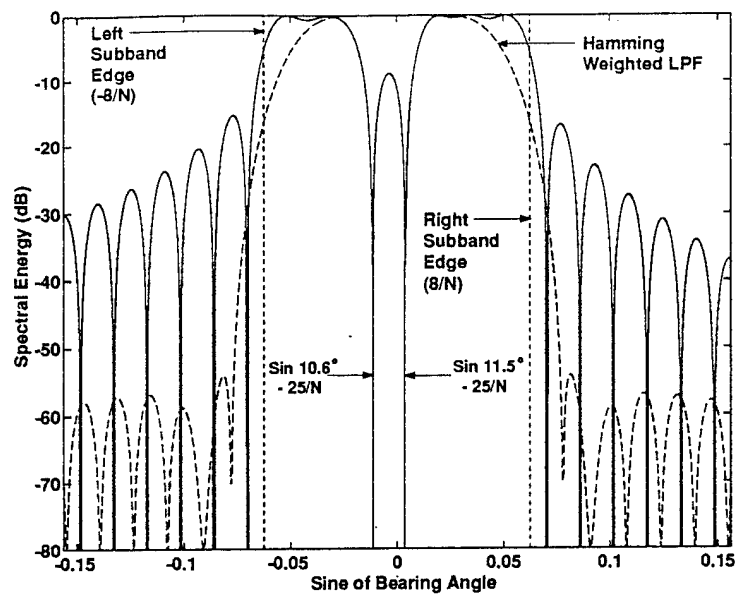


Figure 4: Null Spectrum After Modulation and Filter Response

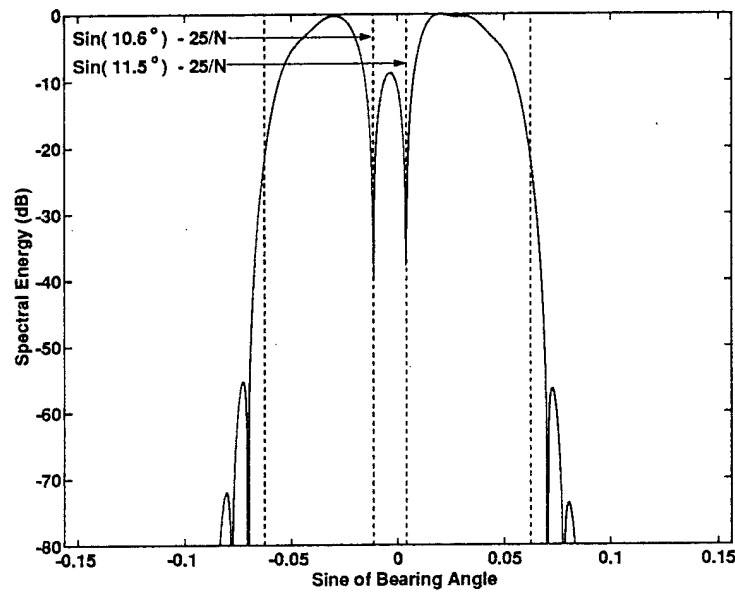


Figure 5: Null Spectrum After Hamming Window Based Bandpass Filtering

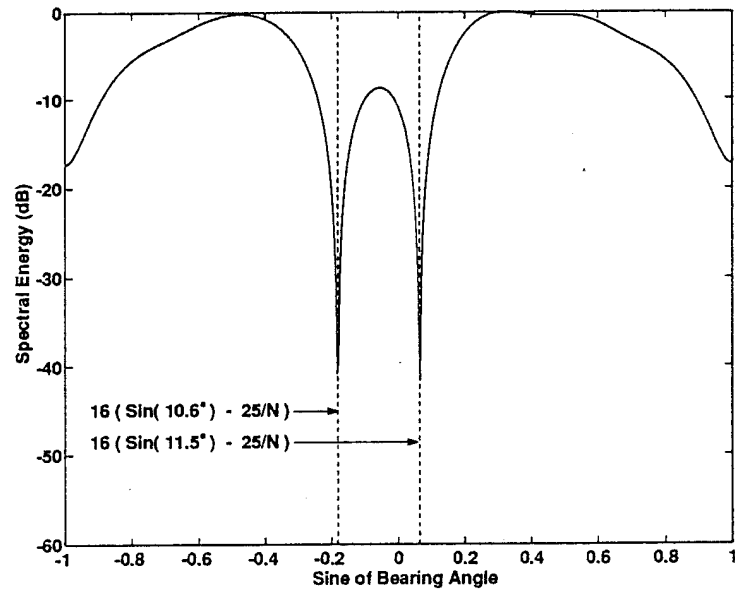


Figure 6: Null Spectrum After Decimation By Factor of $128/8 = 16$

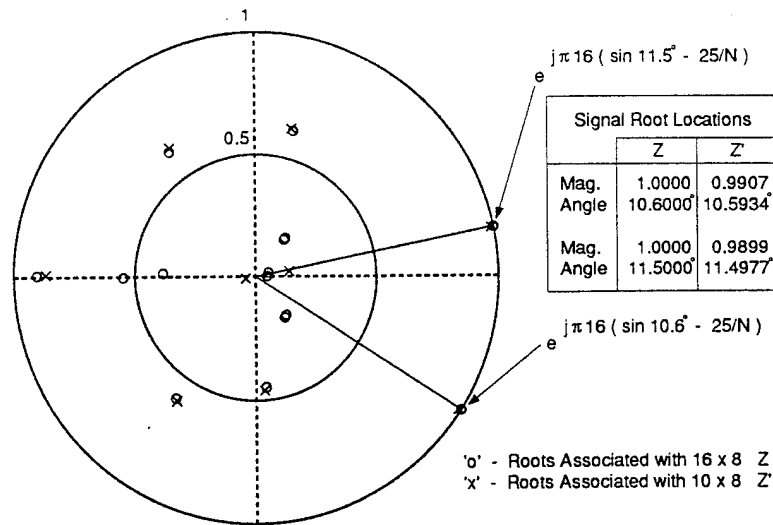


Figure 7: Roots Using Modulated-Filtered-Decimated Noise Eigenvectors

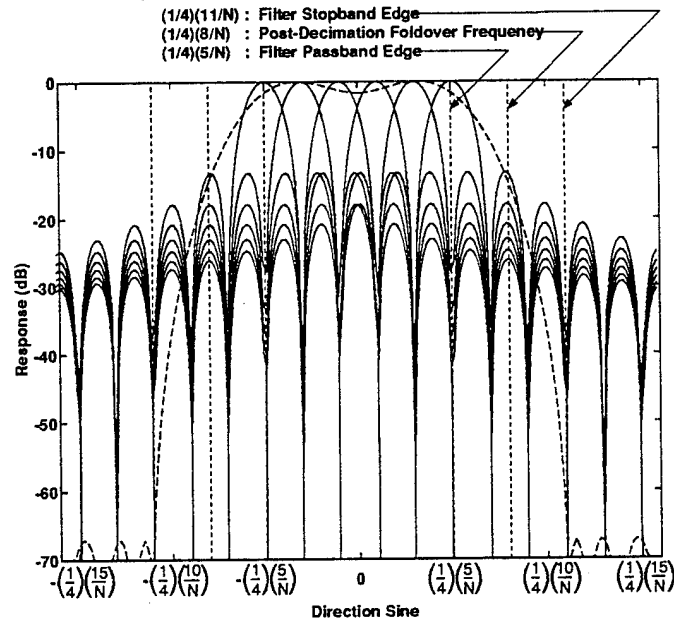


Figure 8: Spatial Responses of Equi-Ripple Filter and Interpolated Beam Set

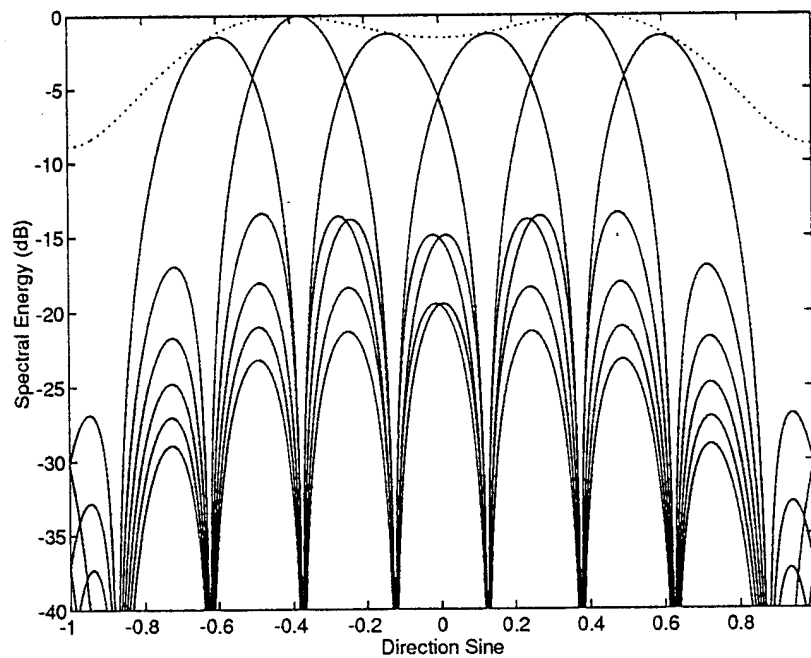


Figure 9: Decimated Filter/Beamformer Spectra

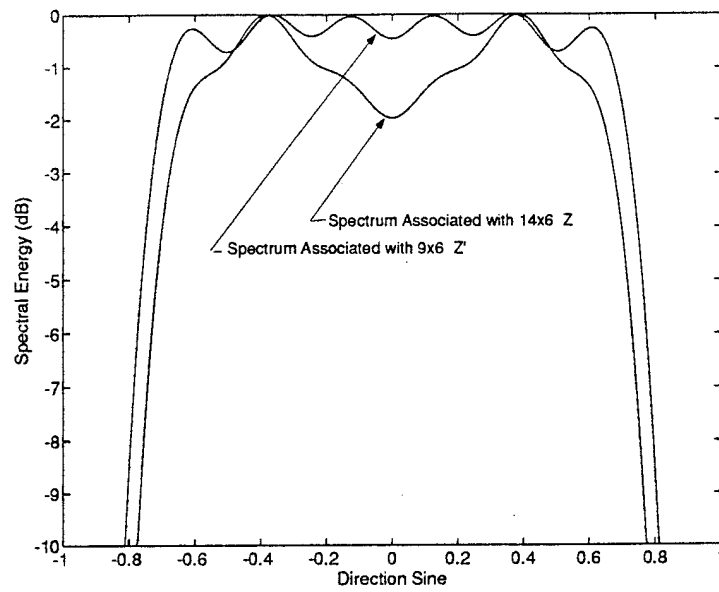


Figure 10: Effects of Filter Deconvolution on White Noise MUSIC Spectrum

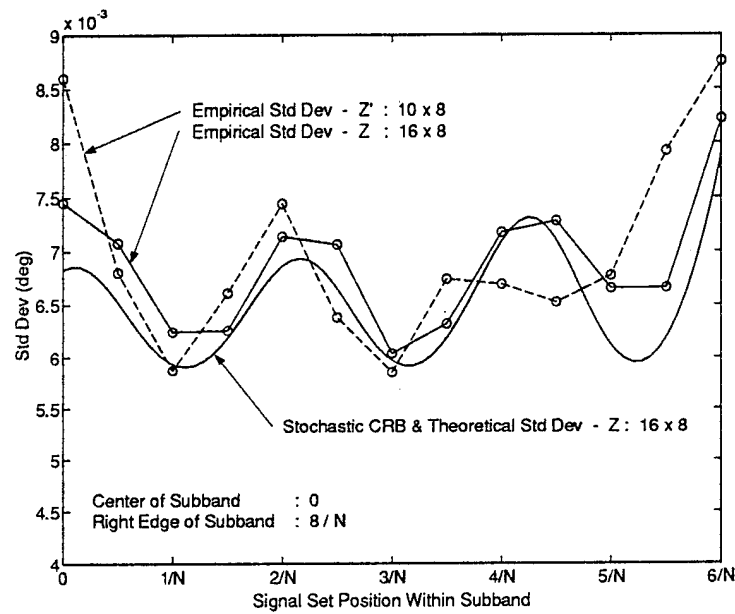


Figure 11: Experiment 1: Left Signal Standard Deviation vs. Signal Set Location Within Subband

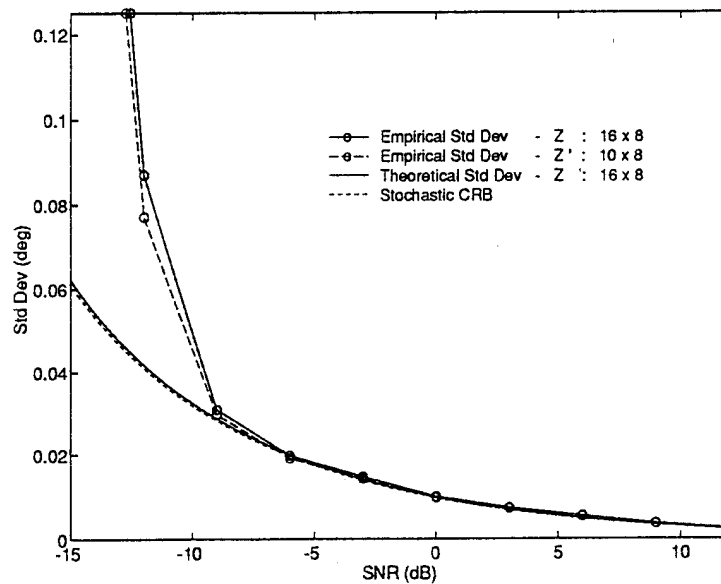


Figure 12: Experiment 2: Left Signal Standard Deviation vs. Source SNR

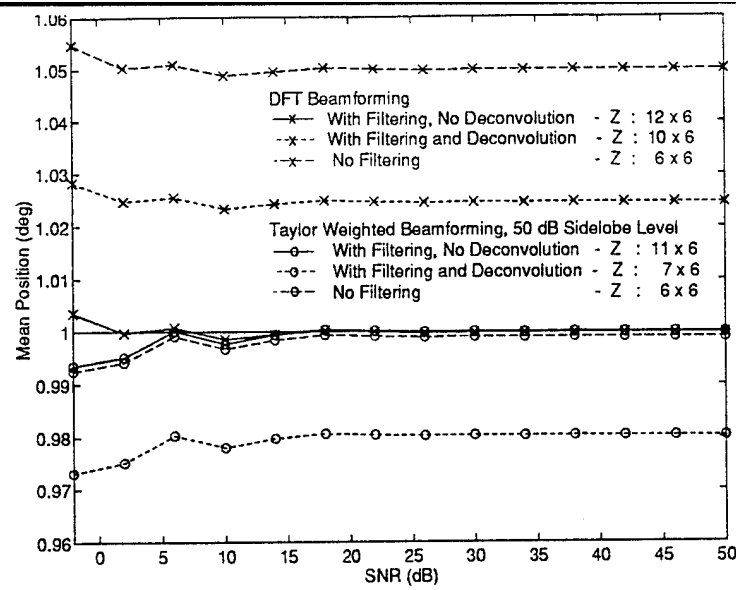


Figure 13: Experiment 3: Location Bias vs. Source SNR

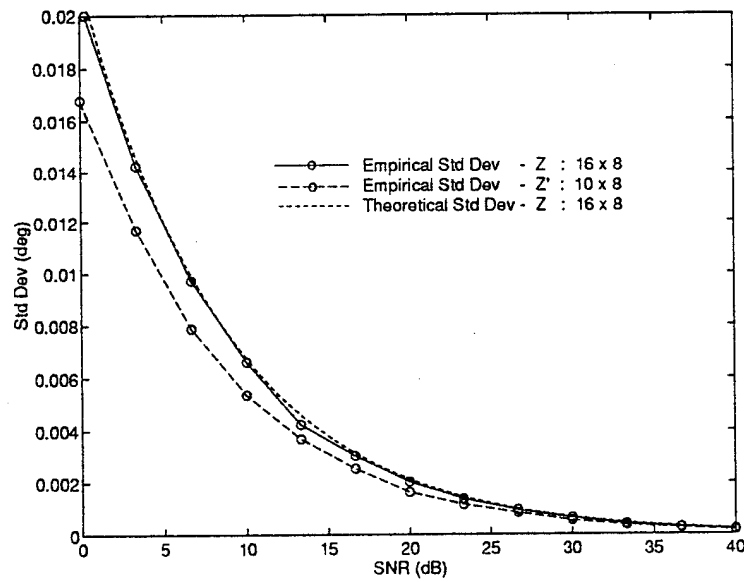


Figure 14: Experiment 4: Left Signal Standard Deviation vs. Source SNR

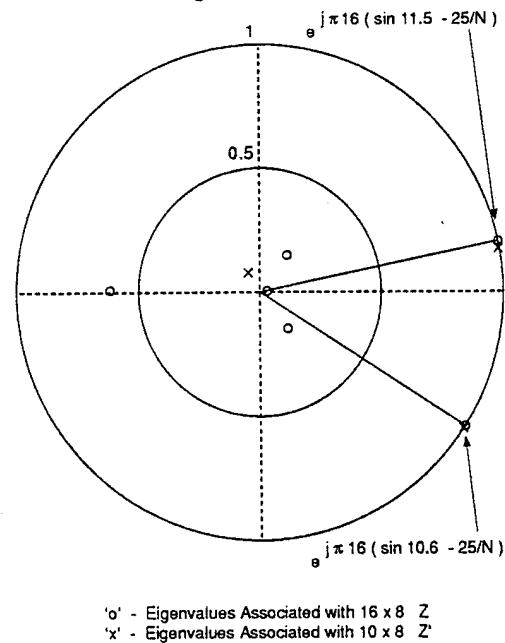


Figure 15: Experiment 4: Quiescent Locations of the ESPRIT Eigenvalues

3 Multidimensional Multirate DOA Estimation in Beamspace

The 1D multirate approach developed in the previous section is extended to the more general case of 2D angle estimation with a uniform rectangular array (URA) of sensors. Multidimensional multirate processing is employed to ultimately yield a small order polynomial in two variables. Again, due to the linearity of the 2D filtering and 2D decimation operations, the actual algorithm merely premultiplies each beam space noise eigenvector by a precomputed transformation matrix. To avoid the spectral search, despite the fact that the fundamental theorem of algebra does not hold in 2D, we propose taking the orthogonal complement of the resulting transformed noise eigenvectors and applying a novel version of ESPRIT facilitating closed-form 2D angle estimation. Simulations demonstrating the efficacy of the approach are presented along with theoretical performance analysis.

3.1 Introduction

3.2 Array Geometry

3.3 Beamforming

3.4 Eigenanalysis

3.5 TLS-ESPRIT

3.6 Bandlimiting the Response

3.7 Further Reductions in Complexity

3.8 Algorithm Summary

3.9 Performance Analysis

3.10 Computer Simulations

3.11 Conclusions

3.12 Appendix: Characterizing the Asymptotic Error

3.13 References

3.14 Figures

1 Introduction

The eigenstructure based Spectral Music Algorithm of Schmidt [1] has become the standard for estimating the Direction of Arrival (DOA) of narrowband plane waves impinging upon a sensor array. Unfortunately the required spectral search is a burdensome task for 1D arrays and computationally prohibitive for 2D arrays. Two well developed methods for reducing this complexity are beamforming techniques [5] and Esprit [4] [6]. Beam space methods reduce the complexity from the number of array sensor elements to the number of beams used to probe a given sector or subband. Furthermore, in the case of a uniform linear array (ULA), beam space techniques yield an implementation (Beam space Root-Music) that allows one to solve for the arrival angles by rooting a small order polynomial. Alternatively, Esprit places a minor restriction on the array geometry and then determines the arrival angles from the eigenvalues of a rotation matrix.

For maximum computational savings, a beam space formulation of Esprit has been desired, but previous attempts have resulted in restrictive requirements on the beamformer. Recently Zoltowski and Kautz [2] [3] developed a beam space formulation of Esprit for 1D ULA's that works with any type of front end beamformer. The new approach is based on the observation that beam space noise eigenvectors may be transformed to vectors in the element space noise subspace, which are bandpass and exhibit nulls at the location of inband sources. This facilitates multirate processing involving modulation to baseband, filtering, and decimation. From the linearity of these operations, the actual algorithm need only need premultiply each beam space noise eigenvector by a simple transformation matrix that is computed a priori. The resulting "telescoped" noise eigenvectors yield a small dimensional element space noise subspace which is used to obtain a small dimensional signal subspace where the Esprit algorithm can be applied.

With the combined advantages of beam space processing and Esprit, multidimensional DOA

estimation becomes computationally feasible. This paper extends the beam space approach to the more general case of 2D angle estimation with a uniform rectangular array (URA). Multidimensional multirate processing is employed to ultimately yield a small dimensional signal subspace. Again, due to the linearity of the 2D filtering and 2D decimation operations, a simple transformation matrix is computed apriori so that the actual algorithm need only premultiply each beam space noise eigenvector by this matrix.

Directly applying the 1D Esprit algorithm to the URA would require two separate applications of Esprit, one for each direction. This estimates the two direction angles independently and leads to the problem of how they can be paired. Alternatively, a novel version of Esprit is developed that estimates the two directions from a single eigenvalue eigenvector pair. Hence they are automatically coupled.

The paper is organized as follows. The array geometry and data model are described in Section 2 and the beamforming process is briefly reviewed in Section 3. The eigen characteristics of the system are developed in Section 4 and multirate processing techniques are applied to the eigenvectors in Section 5. The applicability of the Esprit algorithm is verified in Section 6. Section 7 addresses the issue of bandlimiting the beamformer response and Section 8 describes some further reductions in computational complexity. Finally in Section 9 the proposed 2D Multirate Esprit Algorithm is presented. A theoretical performance analysis is presented in Section 10 and computer simulations are examined in Section 11. A few concluding remarks are included as Section 12. The notation used in this paper indicates vectors by lower case bold letters and matrices by upper case bold letters. The Hermitian, conjugate transpose, will be denoted by a superscript H and the conjugate will be denoted by a superscript $*$.

2 Array Geometry

The array geometry considered in this paper is a rectangular array comprised of M elements in the $\hat{x}$ direction and N elements in the $\hat{y}$ direction uniformly spaced by $\Delta_x = \Delta_y = \lambda_o/2$ (see Figure 1). To specify the source directions, define $\hat{x}$, $\hat{y}$, and $\hat{z}$ to be unit vectors along the coordinate axes and

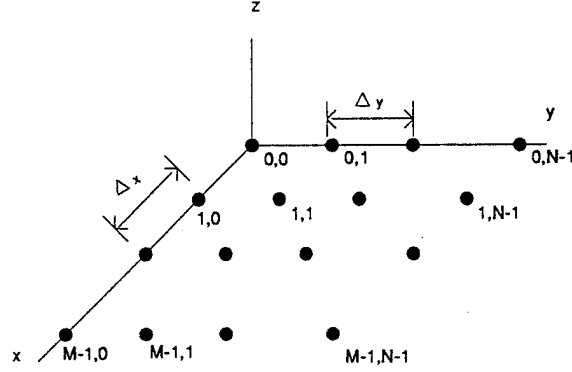


Figure 1: Array Geometry

$\alpha_x, \alpha_y, \alpha_z$ to be the angles between a vector and the respective coordinate axis. If $\hat{p}_i$ is a unit vector normal to a plane wave emanating from the i^{th} source, then $\hat{p}_i = \cos(\alpha_{x_i})\hat{x} + \cos(\alpha_{y_i})\hat{y} + \cos(\alpha_{z_i})\hat{z}$. These direction cosines are converted to azimuth, θ , and elevation, ϕ , angles as $\cos \alpha_{x_i} = \cos \theta_i \sin \phi_i$ and $\cos \alpha_{y_i} = \sin \theta_i \sin \phi_i$ (see Figure 2).

Define an arbitrary reference point to be $r \triangleq (x_r, y_r, 0) = (k_r \Delta_x, l_r \Delta_y, 0)$ and let $\vec{r}_{k,l}$ be a vector from r to the k, l^{th} sensor. Then $\vec{r}_{k,l} = (k - k_r) \Delta_x \hat{x} + (l - l_r) \Delta_y \hat{y}$. Assuming that the signals are narrowband with common center frequency ω_o , the response of the k, l^{th} sensor to the i^{th} source at time t can be written as

$$X_{k,l}^i(t) = s_i(t) e^{j\omega_o (\frac{1}{c} \vec{r}_{k,l} \cdot \hat{p}_i)} = s_i(t) e^{j\frac{2\pi}{\lambda_o} [(k - k_r) \Delta_x \cos \theta_i \sin \phi_i + (l - l_r) \Delta_y \sin \theta_i \sin \phi_i]}. \quad (1)$$

Because $\sin(\phi) = \sin(\pi - \phi)$, a signal with direction angles (θ, ϕ) and a signal with direction angles $(\theta, \pi - \phi)$ will have the same array response. This produces a directional ambiguity that is inherent to uniform rectangular arrays (URA). To see how this ambiguity manifests itself notice that if ϕ

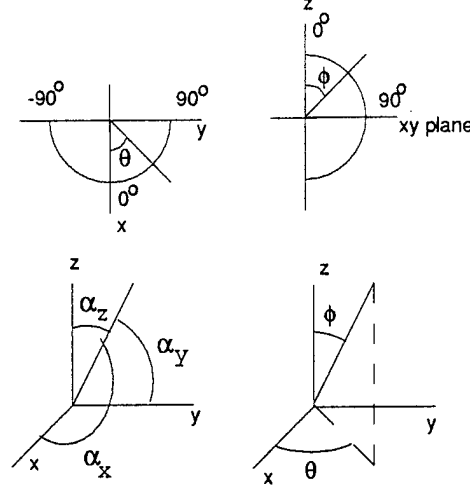


Figure 2: Angle Definitions

corresponds to a direction from above the array, $\phi \in [0, \frac{\pi}{2}]$, then $\pi - \phi$ corresponds to a direction from below the array, $\pi - \phi \in [\frac{\pi}{2}, \pi]$. Therefore, the direction of arrival (DOA) of a signal can only be resolved to two possibilities, one from above the array and one from below the array. This is a significant reduction from the directional ambiguity of a uniform linear array (ULA) which consists of a cone encircling the array. Since the sensor array will generally be mounted on a platform or the body of a plane, the ambiguity can be removed by assuming that all signals impinge from above the array. To facilitate this assumption, the angles are restricted to the ranges $\theta \in [-\pi, \pi]$ and $\phi \in [0, \frac{\pi}{2}]$, and the spatial frequency variables μ and ν are defined as

$$\begin{aligned}\mu &= \frac{2\pi}{\lambda_o} \Delta_x \cos \theta \sin \phi = \pi \cos \theta \sin \phi \in [-\pi, \pi] \\ \nu &= \frac{2\pi}{\lambda_o} \Delta_y \sin \theta \sin \phi = \pi \sin \theta \sin \phi \in [-\pi, \pi].\end{aligned}\tag{2}$$

The azimuth and elevation angles can be recovered from the spatial frequencies by noticing that $\theta = \arctan\left(\frac{\nu}{\mu}\right)$ and $\phi = \arcsin\left(\frac{1}{\pi}\sqrt{\mu^2 + \nu^2}\right)$. Where the full four quadrant inverse tangent is used. With these spatial frequency variables, equation (1) becomes

$$X_{k,l}^i(t) = s_i(t)e^{j[(k-k_r)\mu_i + (l-l_r)\nu_i]} = s_i(t)e^{-j(k_r\mu_i + l_r\nu_i)}e^{j(k\mu_i + l\nu_i)}.\tag{3}$$

For notational convenience $\mathbf{X}(m)$ will be used to denote the matrix obtained by sampling

the array at time t_m and $\mathbf{X}^i(m)$ will denote the array response due to the i^{th} signal at time t_m where m is the snapshot index. Using equation (3) $\mathbf{X}^i(m)$ is given by

$$\mathbf{X}^i(m) = s_i(t_m) e^{-j(k_r \mu_i + l_r \nu_i)} \begin{bmatrix} 1 & e^{j\nu_i} & \dots & e^{j(N-1)\nu_i} \\ e^{j\mu_i} & e^{j(\mu_i + \nu_i)} & \dots & e^{j[\mu_i + (N-1)\nu_i]} \\ \vdots & \vdots & \dots & \vdots \\ e^{j(M-1)\mu_i} & e^{j[(M-1)\mu_i + \nu_i]} & \dots & e^{j[(M-1)\mu_i + (N-1)\nu_i]} \end{bmatrix}. \quad (4)$$

The term $e^{-j(k_r \mu_i + l_r \nu_i)}$ is an arbitrary phase common to all sensors, that is determined by the reference point. Choosing the origin as reference point ($k_r = l_r = 0$), yields an element space array manifold, i.e. the array response viewed as a function of μ and ν , of the form $\mathbf{A}_{MN}(\mu, \nu) = \mathbf{a}_M(\mu) \mathbf{a}_N^T(\nu)$. Where $\mathbf{a}_M(\mu)$ and $\mathbf{a}_N(\nu)$ are defined to be the one dimensional uniform linear array (ULA) manifold vectors.

$$\begin{aligned} \mathbf{a}_M(\mu) &\triangleq [1, e^{j\mu}, \dots, e^{j(M-1)\mu}]^T \\ \mathbf{a}_N(\nu) &\triangleq [1, e^{j\nu}, \dots, e^{j(N-1)\nu}]^T \end{aligned} \quad (5)$$

The m^{th} snapshot of the array due to all d impinging signals can now be written as

$$\mathbf{X}(m) = \sum_{i=0}^{d-1} s_i(m) \mathbf{A}_{MN}(\mu_i, \nu_i) + \mathbf{N}(m) \quad (6)$$

Where $N_{k,l}(m)$ is measurement noise associated with the m^{th} snapshot of the k, l^{th} sensor, and $\mathbf{N}(m)$ is the $M \times N$ noise matrix.

It is useful to view the array response as an $MN \times 1$ vector as well as an $M \times N$ matrix. To facilitate conversions between these forms, consider the operator vec that maps an $M \times N$ matrix to an $MN \times 1$ vector by concatenating its columns, and the inverse operator, mat , that maps an $MN \times 1$ vector to an $M \times N$ matrix by using M consecutive elements of the vector for each column of the resulting matrix. If $\mathbf{Q}$ is an arbitrary matrix with columns denoted $\mathbf{q}_k$, then $vec(\mathbf{Q}) = [\mathbf{q}_1^T, \mathbf{q}_2^T, \dots, \mathbf{q}_N^T]^T$. If $\mathbf{q}$ is an arbitrary vector and $\mathbf{q}(k:l)$ denotes its k^{th} through l^{th}

Table 1: Properties

$$\begin{aligned}
\text{P1:} \quad & \mathbf{v}_N \otimes \mathbf{x}_M = \text{vec}(\mathbf{x}_M \mathbf{v}_N^T) \\
\text{P2:} \quad & \mathbf{x}_M \mathbf{v}_N^T = \text{mat}(\mathbf{v}_N \otimes \mathbf{x}_M) \\
\text{P3:} \quad & \text{vec}(\mathbf{A} \mathbf{D} \mathbf{B}) = (\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{D}) \\
\text{P4:} \quad & (\mathbf{A} \otimes \mathbf{B})^T = (\mathbf{A}^T \otimes \mathbf{B}^T) \\
\text{P5:} \quad & (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A} \mathbf{C}) \otimes (\mathbf{B} \mathbf{D})
\end{aligned}$$

elements, then $\text{mat}(\mathbf{q}) = [\mathbf{q}(0 : M - 1) \mid \dots \mid \mathbf{q}(MN - M : MN - 1)]$. The following example illustrates these operators:

$$\begin{aligned}
\text{vec} \left(\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \right) &= [1, 2, 3, 4, 5, 6]^T \\
\text{mat}([1, 2, 3, 4, 5, 6]^T) &= \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.
\end{aligned}$$

Some important properties of vec , mat , and the Kronecker product, $\otimes$, are listed in Table 1 (see also [9]). Most notably, Property 1 allows the array response to be written in vector form as

$$\mathbf{x}(m) \triangleq \text{vec}(\mathbf{X}(m)) = \sum_{i=0}^{d-1} s_i(m) \mathbf{a}_{MN}(\mu_i, \nu_i) + \mathbf{n}(m) = \mathcal{A}_{MNS}(m) + \mathbf{n}(m). \quad (7)$$

Where $\mathbf{s}(m)$ is the vector of signal amplitudes, $\mathbf{a}_{MN}(\mu, \nu)$ is the array manifold in vector form, and the columns of $\mathcal{A}_{MN}$ are the signal steering vectors, i.e., the array manifold evaluated at the spatial frequencies corresponding to the signal directions.

$$\mathbf{s}(m) \triangleq [s_0(m), \dots, s_{d-1}(m)]^T \quad (8)$$

$$\mathbf{a}_{MN}(\mu, \nu) \triangleq \text{vec}(\mathbf{A}_{MN}(\mu, \nu)) = \mathbf{a}_N(\nu) \otimes \mathbf{a}_M(\mu) \quad (9)$$

$$\mathcal{A}_{MN} \triangleq [\mathbf{a}_{MN}(\mu_0, \nu_0) \mid \dots \mid \mathbf{a}_{MN}(\mu_{d-1}, \nu_{d-1})]. \quad (10)$$

The element space signal subspace is defined to be the column-space of $\mathcal{A}_{MN}$, $\mathcal{S}_e \triangleq \mathcal{R}\{\mathcal{A}_{MN}\}$, and

the element space noise subspace is defined to be the orthogonal complement of $\mathcal{S}_e$. Since $\mathcal{A}_{MN}$ is $MN \times d$, $\mathcal{S}_e$ is a d dimensional subspace of MN dimensional space.

3 Beamforming

Now consider the class of separable two dimensional beamformers. Let $\mathbf{W}_\mu$ be an arbitrary $M \times M_b$ beamforming matrix (with $M_b < M$) for the μ spatial frequency. The k^{th} row of $\mathbf{W}_\mu^H$ denoted $\mathbf{w}_{\mu k}^H$ forms a beam for a specific frequency in the desired range. $\mathbf{W}_\nu$ is defined in a similar fashion for the ν spatial frequency. Using $\mathbf{w}_{\mu k}^H$ in conjunction with $\mathbf{w}_{\nu l}^*$ a beam is generated for a specific 2D frequency in the subband of interest. Therefore, the $M_b \times N_b$ beam space snapshot matrix is formed as $\mathbf{Y}(m) = \mathbf{W}_\mu^H \mathbf{X}(m) \mathbf{W}_\nu^*$, and the $M_b N_b \times 1$ beam space snapshot vector is given by $\mathbf{y} = \text{vec}(\mathbf{Y})$.

Using Property 3 from Table 1 these can be written as

$$\begin{aligned} \mathbf{Y}(m) &= \mathbf{W}_\mu^H \mathbf{X}(m) \mathbf{W}_\nu^* = \sum_{i=0}^{d-1} s_i(m) \mathbf{W}_\mu^H \mathbf{A}_{MN}(\mu_i, \nu_i) \mathbf{W}_\nu^* + \mathbf{W}_\mu^H \mathbf{N}(m) \mathbf{W}_\nu^* \\ \mathbf{y}(m) &= \text{vec}(\mathbf{Y}(m)) = [\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H] \mathcal{A}_{MNS}(m) + [\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H] \mathbf{n}(m). \end{aligned} \quad (11)$$

Therefore, the beam space array manifold is described by

$$\mathbf{B}(\mu, \nu) = \mathbf{W}_\mu^H \mathbf{A}_{MN}(\mu, \nu) \mathbf{W}_\nu^* \quad (12)$$

$$\mathbf{b}(\mu, \nu) = [\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H] \mathbf{a}_{MN}(\mu, \nu) = [\mathbf{W}_\nu^H \mathbf{a}_N(\nu)] \otimes [\mathbf{W}_\mu^H \mathbf{a}_M(\mu)] \quad (13)$$

$$\mathcal{B} \triangleq [\mathbf{b}(\mu_0, \nu_0) \mid \dots \mid \mathbf{b}(\mu_{d-1}, \nu_{d-1})]. \quad (14)$$

Finally the beam space signal subspace, $\mathcal{S}_b$, is defined to be the d dimensional subspace of $M_b N_b$ dimensional space that is spanned by the columns of $\mathcal{B}$, and the $M_b N_b - d$ dimensional beam space noise subspace is defined to be the orthogonal complement of $\mathcal{S}_b$.

At this time it should be noted that the beamformer need not be separable. This assumption was made because the separable nature of the array structure leads directly to separable beamform-

ers. In vector form, nonseparable beamformers modify equation (11) by replacing $\mathbf{W}_\nu \otimes \mathbf{W}_\mu$ with an arbitrary $MN \times M_b N_b$ matrix $\mathbf{W}_b$.

4 Eigenanalysis

Using equation (7) and assuming that the measurement noise is zero mean, uncorrelated between sensors, and has equal power σ^2 , the $MN \times MN$ element space autocorrelation matrix is

$$\mathbf{R}_x = \mathbb{E} \{ \mathbf{x}(n) \mathbf{x}(n)^H \} = \mathcal{A} \mathbf{R}_s \mathcal{A}^H + \sigma^2 \mathbf{I}. \quad (15)$$

It has been observed [4] that the eigenvectors of $\mathbf{R}_x$ corresponding to the d largest eigenvalues form a basis for $\mathcal{S}_e$ and the remaining eigenvectors of $\mathbf{R}_x$ form a basis for $\mathcal{S}_e^\perp$. Under the assumption of orthonormal beams, the $M_b N_b \times M_b N_b$ beam space autocorrelation matrix has the form

$$\mathbf{R}_y = [\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H] \mathcal{A} \mathbf{R}_s \mathcal{A}^H [\mathbf{W}_\nu \otimes \mathbf{W}_\mu] + \sigma^2 \mathbf{I} = \mathcal{B} \mathbf{R}_s \mathcal{B}^H + \sigma^2 \mathbf{I}. \quad (16)$$

The eigenvectors of $\mathbf{R}_y$ are also divided into two sets, the “beam space signal eigenvectors” $\{\mathbf{f}_i : i = 0, \dots, d-1\}$ that form a basis for $\mathcal{S}_b$, and the “beam space noise eigenvectors” $\{\mathbf{f}_j : j = d, \dots, M_b N_b - 1\}$ that form a basis for $\mathcal{S}_b^\perp$.

If the number of signals, d , is unknown it can be estimated at this time using the AIC or MDL methods of [10]. Therefore, in all further developments, it will be assumed that d is known.

Since the “beam space noise eigenvectors” lie in $\mathcal{S}_b^\perp$, they are orthogonal to the beam space signal steering vectors, i.e., $\mathbf{b}^H(\mu_i, \nu_i) \mathbf{f}_j = 0$ for all $i = 0, \dots, d-1$ and $j = d, \dots, M_b N_b - 1$. Recalling equation (13) yields the following important result.

$$0 = \mathbf{b}^H(\mu_i, \nu_i) \mathbf{f}_j = [(\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H) \mathbf{a}_{MN}(\mu_i, \nu_i)]^H \mathbf{f}_j = \mathbf{a}_{MN}^H(\mu_i, \nu_i) [(\mathbf{W}_\nu \otimes \mathbf{W}_\mu) \mathbf{f}_j] \quad (17)$$

This shows that the matrix $\mathbf{W}_\nu \otimes \mathbf{W}_\mu$ maps the beam space noise eigenvectors, $\mathbf{f}_j$, to the element space noise subspace, $\mathcal{S}_e^\perp$ [3]. However, there are only $M_b N_b - d$ beam space noise eigenvectors, so

this mapping does not yield a complete basis for $\mathcal{S}_e^\perp$.

The general Music algorithm exploits the orthogonality between the beam space noise eigenvectors and the signal steering vectors, by forming the beam space Music null spectrum $S_B(\mu, \nu)$.

$$S_B(\mu, \nu) = \sum_{j=d}^{M_b N_b - 1} \left| \mathbf{b}^H(\mu, \nu) \mathbf{f}_j \right|^2 = \sum_{j=d}^{M_b N_b - 1} \left| \mathbf{a}_M^H(\mu) \mathbf{W}_\mu \mathbf{F}_j \mathbf{W}_\nu^T \mathbf{a}_N^*(\nu) \right|^2 \quad (18)$$

Where $\mathbf{F}_j$ is defined to be the beam space noise eigenvector written in matrix form ($\mathbf{F}_j \triangleq \text{mat}(\mathbf{f}_j)$). Signal directions are then estimated from values of μ and ν corresponding to nulls in $S_B(\mu, \nu)$. Two well established problems with this method [4] are that the array manifold must be known and stored, and the search over a two dimensional space can be computationally prohibitive.

5 Multirate Processing of Beam Space Noise Eigenvectors

In an effort to circumvent these problems, notice that the crux of the beam space Music null spectrum are the $M \times N$ telescoped [3] beam space noise eigenvectors $\mathbf{G}_j \triangleq \mathbf{W}_\mu \mathbf{F}_j \mathbf{W}_\nu^T$ ($j = d, \dots, M_b N_b - 1$) and $G_j(\mu, \nu) \triangleq \mathbf{a}_M^H(\mu) \mathbf{G}_j \mathbf{a}_N^*(\nu)$ is the two dimensional Discrete Space Fourier Transform of $\mathbf{G}_j$. Letting $F_j(k, l)$ denote the k, l^{th} element of $\mathbf{F}_j$ yields

$$G_j(\mu, \nu) = \mathbf{a}_M^H(\mu) \mathbf{W}_\mu \mathbf{F}_j \mathbf{W}_\nu^T \mathbf{a}_N^*(\nu) = \sum_{k=0}^{M_b-1} \sum_{l=0}^{N_b-1} F_j(k, l) \left[\mathbf{a}_M^H(\mu) \mathbf{w}_{\mu_k} \mathbf{w}_{\nu_l}^T \mathbf{a}_N^*(\nu) \right]. \quad (19)$$

Since $\mathbf{w}_{\mu_k}$ and $\mathbf{w}_{\nu_l}$ form a beam in the desired subband, $\mathbf{a}_M^H(\mu) \mathbf{w}_{\mu_k} \mathbf{w}_{\nu_l}^T \mathbf{a}_N^*(\nu)$ is a bandpass function of μ and ν for all k, l . Consequently $G_j(\mu, \nu)$ is a bandpass function of μ and ν . Without loss of generality, assume that the $M_b N_b$ beams encompass the spatial subband defined by $-\pi \left(\frac{M_b}{M} \right) \leq \mu \leq \pi \left(\frac{M_b}{M} \right)$ and $-\pi \left(\frac{N_b}{N} \right) \leq \nu \leq \pi \left(\frac{N_b}{N} \right)$ and have sufficiently low out of band sidelobes. Then the beamformer response is negligible outside this subband, i.e., $G_j(\mu, \nu) \approx 0$ for $\frac{M_b}{M} \pi < |\mu| < \pi$ and $\frac{N_b}{N} \pi < |\nu| < \pi$, so $\mathbf{G}_j$ can be decimated by $d_x = \frac{M}{M_b}$ and $d_y = \frac{N}{N_b}$ without incurring a significant amount of aliasing. (Note: $\mathbf{G}_j$ can always be modulated to baseband and filtered to make this assumption valid. See section 7).

The decimation process can be modeled mathematically as premultiplying $\mathbf{G}_j$ by $\mathbf{D}_x$ and postmultiplying by $\mathbf{D}_y^T$ where $\mathbf{D}_x$ and $\mathbf{D}_y$ are the $M_b \times M$ and $N_b \times N$ decimation matrices. For example if $M_b = 2$ and $M = 6$

$$\mathbf{D}_x = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Therefore the decimated telescoped beam space noise eigenvectors are given by

$$\begin{aligned} \mathbf{H}_j &= \mathbf{D}_x \mathbf{G}_j \mathbf{D}_y^T = (\mathbf{D}_x \mathbf{W}_\mu) \mathbf{F}_j (\mathbf{D}_y \mathbf{W}_\nu)^T & (M_b \times N_b) \\ \mathbf{h}_j &= \text{vec}(\mathbf{H}_j) = [(\mathbf{D}_y \mathbf{W}_\nu) \otimes (\mathbf{D}_x \mathbf{W}_\mu)] \mathbf{f}_j & (M_b N_b \times 1) \end{aligned} \quad (20)$$

for $j = d, \dots, M_b N_b - 1$.

It is important to note that since decimation is a linear operation it can be performed apriori on the telescoping matrices. Furthermore, fractional sampling rate alterations can be effected by replacing $\mathbf{D}_x$ with $\mathbf{D}_x \mathbf{F}_x \mathbf{I}_x$, where $\mathbf{I}_x$ represents an interpolation matrix and $\mathbf{F}_x$ represents filtering. The space spanned by the decimated telescoped beam space noise eigenvectors, $\mathbf{h}_j$, will be referred to as $\mathcal{S}_d^\perp$.

Since $\mathbf{H}_j$ is $M_b \times N_b$, it has a 2D-DSFT given by $H_j(\mu, \nu) = \mathbf{a}_{M_b}^H(\mu) \mathbf{H}_j \mathbf{a}_{N_b}^*(\nu)$ where $\mathbf{a}_{M_b}(\mu)$ and $\mathbf{a}_{N_b}(\nu)$ are defined analogously to equation (5). By standard Multirate analysis [13] the relationship between $H_j(\mu, \nu)$ and $G_j(\mu, \nu)$ is

$$H_j(\mu, \nu) = \frac{1}{d_x d_y} \sum_{p=0}^{d_x-1} \sum_{q=0}^{d_y-1} G_j \left(\frac{\mu - 2\pi p}{d_x}, \frac{\nu - 2\pi q}{d_y} \right). \quad (21)$$

Since the beamformer response is negligible outside the subband, $H_j(\mu, \nu) \approx \frac{1}{d_x d_y} G_j \left(\frac{\mu}{d_x}, \frac{\nu}{d_y} \right)$ for all $-\pi \leq \mu, \nu \leq \pi$ and therefore

$$H_j(d_x \mu_i, d_y \nu_i) = \frac{1}{d_x d_y} G_j(\mu_i, \nu_i) = \frac{1}{d_x d_y} \mathbf{b}^H(\mu_i, \nu_i) \mathbf{f}_j = 0. \quad (22)$$

This shows that the decimation process preserves the in band source nulls and increases their separation by a factor of d_x and d_y . Hence, the beam space Music null spectrum (18) could be

reformulated as

$$S_B(\mu, \nu) = \sum_{j=d}^{M_b N_b - 1} |H_j(\mu, \nu)|^2 = \sum_{j=d}^{M_b N_b - 1} |\mathbf{a}_{M_b}^H(\mu) \mathbf{H}_j \mathbf{a}_{N_b}^*(\nu)|^2 \quad (23)$$

thereby reducing the computational intensity of each evaluation of $S_B(\mu, \nu)$.

In the one dimensional case, this search can be removed by defining $z = e^{j\mu}$ and writing S_B as a polynomial in z . Signal directions are then obtained from the roots of S_B . This procedure, referred to as Root-Music [5], has always been theoretically possible for the two dimensional case, but the lack of 2D rooting algorithms has precluded its use in practice. However the efficient 2D rooting algorithm recently proposed by *someone* [15] has made 2D Root-Music a viable option.

It is well known that the Esprit algorithm [4] offers another alternative to the spectral search of Music. In an effort to apply Esprit, notice that

$$H_j(d_x \mu_i, d_y \nu_i) = \mathbf{a}_{M_b}^H(d_x \mu_i) \mathbf{H}_j \mathbf{a}_{N_b}^*(d_y \nu_i) = \mathbf{a}_{M_b N_b}^H(d_x \mu_i, d_y \nu_i) \mathbf{h}_j. \quad (24)$$

Comparing equations (22) and (24), it is seen that

$$\mathbf{a}_{M_b N_b}^H(d_x \mu_i, d_y \nu_i) \mathbf{h}_j = 0 \quad \forall \quad i = 0, \dots, d-1 \text{ and } j = d, \dots, M_b N_b - 1. \quad (25)$$

This shows that the decimated telescoped noise eigenvectors form a complete basis for a lower dimensional element space noise subspace $\mathcal{S}_d^\perp$. Therefore, the orthogonal complement, $\mathcal{S}_d$, is a lower dimensional element space signal subspace. This space will be referred to as the decimated signal subspace, even though it is not obtained by decimating the signal subspace.

Before showing that $\mathcal{S}_d$ has the Esprit structure, it is useful to summarize the preceding results. The original array response resided in the element space signal subspace which is a d dimensional subspace of MN dimensional space defined as $\mathcal{S}_e = \text{span} \{ \mathbf{a}_{MN}(\mu_i, \nu_i) \mid i = 0, \dots, d-1 \}$. Due to the array geometry, this space has the Esprit structure, however, the array response is an $M \times N$ matrix (or $MN \times 1$ vector) which can make computations unwieldy. Since the signals are

known to be in a certain subband, a beamformer is applied that imposes a bandpass characteristic on the array response, which can then be decimated without incurring aliasing. This reduces the signal space to a d dimensional subspace of an $M_b N_b$ dimensional space and yields matrices that are only $M_b \times N_b$. However in the beamforming process the Esprit structure is lost. This structure can be restored by decimating and telescoping in the noise subspace, then converting back to the orthogonal complement (see Figure 3). It is important to notice that the element space signal eigen-

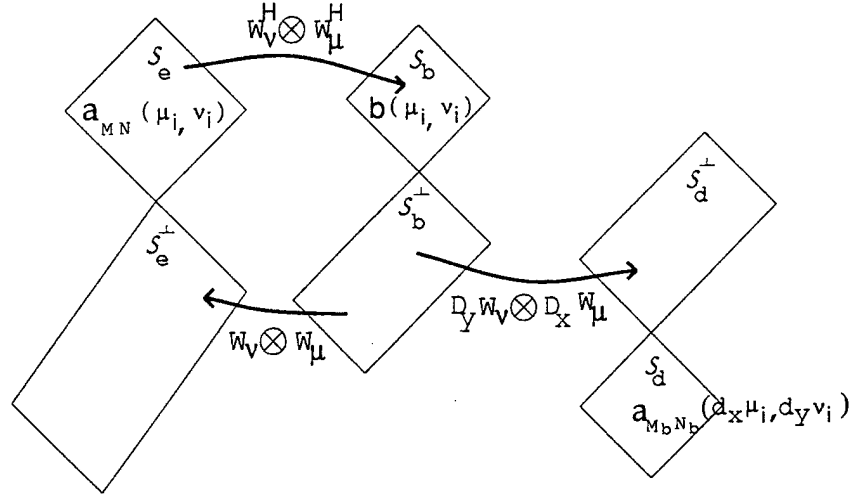


Figure 3: Subspace Relations

vectors which are obtained directly from the element space correlation matrix cannot be decimated because they are not bandlimited. The beam space signal eigenvectors are not telescoped because this does not yield vectors in the element space signal subspace. Hence they will not have the Esprit structure. However telescoping and decimating the beam space noise eigenvectors yields a space that is the orthogonal complement to a smaller dimensional element space signal space.

6 TLS-Esprit

It remains to show that S_d does indeed have the Esprit structure. The Esprit algorithm requires an array formed by “sensor doublets” that are separated by a constant displacement vector [4]. This

can be accomplished by viewing the $M_b \times N_b$ rectangular array as two overlapping $M_b \times (N_b - 1)$ subarrays with $M_b(N_b - 2)$ common elements (see Figure 4a). The resulting subarray manifolds are

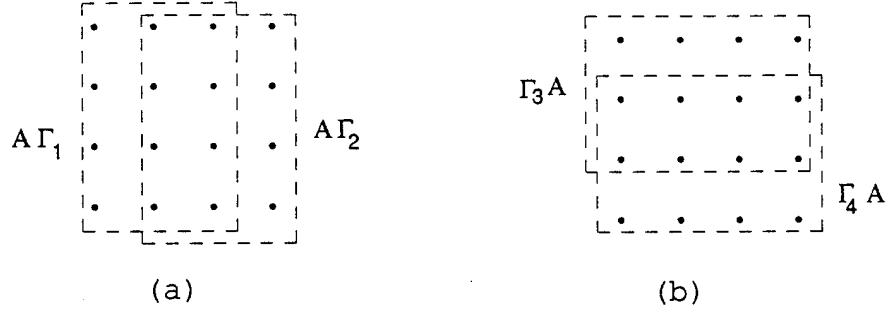


Figure 4: Array Partitioning

given by the first and last $N_b - 1$ columns of $\mathbf{A}_{M_b N_b}(\mu, \nu)$. Mathematically this can be modeled as $\mathbf{A}_{M_b N_b}(\mu, \nu)\mathbf{\Gamma}_1$ and $\mathbf{A}_{M_b N_b}(\mu, \nu)\mathbf{\Gamma}_2$, where $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ are the first and last $N_b - 1$ columns of the $N_b \times N_b$ identity matrix. For example, if $N_b = 3$, $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ are as follows:

$$\mathbf{\Gamma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \mathbf{\Gamma}_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The vector form of the subarray manifolds are obtained by premultiplying by the $M_b \times M_b$ identity matrix and applying Property 3 of Table 1. It is easily verified that the two subarray manifolds, in matrix and vector form, are related as follows:

$$\begin{aligned} \mathbf{A}_{M_b N_b}(\mu, \nu)\mathbf{\Gamma}_2 &= e^{j\nu} \mathbf{A}_{M_b N_b}(\mu, \nu)\mathbf{\Gamma}_1 \\ \mathbf{J}_2 \mathbf{a}_{M_b N_b}(\mu, \nu) &= e^{j\nu} \mathbf{J}_1 \mathbf{a}_{M_b N_b}(\mu, \nu). \end{aligned} \tag{26}$$

Where $\mathbf{J}_1 \triangleq \mathbf{\Gamma}_1^T \otimes \mathbf{I}_{M_b}$ and $\mathbf{J}_2 \triangleq \mathbf{\Gamma}_2^T \otimes \mathbf{I}_{M_b}$. Therefore, the signal steering vectors for the subarrays are related by $\mathbf{J}_2 \mathbf{a}_{M_b N_b}(d_x \mu_i, d_y \nu_i) = e^{jd_y \nu_i} \mathbf{J}_1 \mathbf{a}_{M_b N_b}(d_x \mu_i, d_y \nu_i)$ for all $i = 0, \dots, d - 1$ and consequently

$$\mathbf{J}_2 \mathcal{A}_{M_b N_b} = \mathbf{J}_1 \mathcal{A}_{M_b N_b} \Upsilon_\nu \tag{27}$$

where $\Upsilon_\nu = \text{diag}\{e^{jd_y\nu_0}, \dots, e^{jd_y\nu_{d-1}}\}$. This can be generalized to an arbitrary basis, $\mathbf{K} = \mathcal{A}_{M_b N_b} \mathbf{T}$, for $\mathcal{S}_d$ by postmultiplying both sides by any $d \times d$ nonsingular matrix $\mathbf{T}$ and premultiplying Υ by $\mathbf{I}_d = \mathbf{T}\mathbf{T}^{-1}$.

$$\mathbf{J}_2 \mathcal{A}_{M_b N_b} \mathbf{T} = \mathbf{J}_1 \mathcal{A}_{M_b N_b} \mathbf{T} \mathbf{T}^{-1} \Upsilon \mathbf{T} \quad (28)$$

$$\mathbf{J}_2 \mathbf{K} = \mathbf{J}_1 \mathbf{K} \Psi. \quad (29)$$

Where $\Psi = \mathbf{T}^{-1} \Upsilon \mathbf{T}$. This relationship is the basis for the TLS-Esprit algorithm [4]. It shows that the ν spatial frequencies can be estimated from the eigenvalues of the matrix that rotates the first $M_b(N_b - 1)$ rows of $\mathbf{K}$ into the last $M_b(N_b - 1)$ rows of $\mathbf{K}$.

Alternatively, the array can be viewed as two overlapping $(M_b - 1) \times N_b$ subarrays with $(M_b - 2)N_b$ common elements (see Figure 4b). This yields subarray manifolds that are the first and last $M_b - 1$ rows of $\mathbf{A}_{M_b N_b}(\mu, \nu)$ and modelled by $\Gamma_3 \mathbf{A}_{M_b N_b}(\mu, \nu)$ and $\Gamma_4 \mathbf{A}_{M_b N_b}(\mu, \nu)$. In this case, Γ_3 and Γ_4 are the first and last $M_b - 1$ rows of the $M_b \times M_b$ identity matrix and

$$\mathbf{J}_4 \mathcal{A}_{M_b N_b} = \mathbf{J}_3 \mathcal{A}_{M_b N_b} \Upsilon_\mu \quad (30)$$

where $\Upsilon_\mu = \text{diag}\{e^{jd_x\mu_0}, \dots, e^{jd_x\mu_{d-1}}\}$, $\mathbf{J}_3 \triangleq \mathbf{I}_{N_b} \otimes \Gamma_3$, and $\mathbf{J}_4 \triangleq \mathbf{I}_{N_b} \otimes \Gamma_4$.

Therefore, if the array is divided in a row-wise fashion, the μ spatial frequencies can be estimated from the eigenvalues of the matrix that rotates $\mathbf{J}_3 \mathbf{K}$ into $\mathbf{J}_4 \mathbf{K}$. However, if the μ and ν frequencies are obtained independently by applying Esprit to $\mathbf{K}$ in a row-wise and column-wise fashion, there is no apparent way to pair the frequency components corresponding to a specific signal.

In an effort to circumvent this problem, notice that as long as no two signals have the same μ and ν frequencies, $\mathbf{J}_1 \mathbf{K}$ and $\mathbf{J}_2 \mathbf{K}$ are rank d . Therefore, Ψ always exists and has a full set of eigenvalues and linearly independent eigenvectors. Consider performing an eigenvalue decomposition of Ψ to obtain $\Psi = \mathbf{E} \mathbf{\Lambda} \mathbf{E}^{-1}$.

If ν_i is a distinct frequency, then γ_i is a distinct eigenvalue of Ψ and the associated right eigenvector is unique (to within a scalar multiple). Therefore, $\mathbf{e}_i$ is the i^{th} column of $\mathbf{T}^{-1}$ and the i^{th} signal steering vector can be obtained as

$$\mathbf{K}\mathbf{e}_i = \mathcal{A}_{M_b N_b} \mathbf{T}\mathbf{e}_i = \alpha_i \mathbf{a}_{M_b N_b}(d_x \mu_i, d_y \nu_i). \quad (31)$$

To estimate the μ frequency from the steering vector recall equation (30), let $\mathbf{l}_i = \mathbf{K}\mathbf{e}_i$, and notice that $(\mathbf{J}_3 \mathbf{l}_i)^H (\mathbf{J}_4 \mathbf{l}_i) = (M_b - 1) N_b e^{j d_x \mu_i}$. Therefore, define ρ_i as

$$\rho_i \triangleq \frac{1}{(M_b - 1) N_b} (\mathbf{J}_3 \mathbf{l}_i)^H (\mathbf{J}_4 \mathbf{l}_i) = \mathbf{l}_i^H \frac{[\mathbf{I}_{N_b} \otimes (\mathbf{\Gamma}_3^T \mathbf{\Gamma}_4)]}{(M_b - 1) N_b} \mathbf{l}_i = \mathbf{l}_i^H \mathbf{P} \mathbf{l}_i = e^{j d_x \mu_i}. \quad (32)$$

So μ_i is obtained as $\mu_i = \frac{1}{d_x} \arg(\rho_i)$. An important observation here is that the μ and ν frequencies for a signal are estimated from an eigenvalue-eigenvector pair and as such are automatically coupled.

Now consider the case where ν_i is not a distinct frequency, say $\nu_0 = \dots = \nu_{p-1}$, then Ψ has an eigenvalue of multiplicity p and the associated eigenvectors $\{\mathbf{e}_0, \dots, \mathbf{e}_{p-1}\}$, are not unique. Therefore, $\mathbf{e}_i$ is not the i^{th} column of $\mathbf{T}^{-1}$ and

$$\mathbf{K}\mathbf{e}_i = [\mathbf{a}_{M_b N_b}(d_x \mu_0, d_y \nu_0) \mid \dots \mid \mathbf{a}_{M_b N_b}(d_x \mu_{p-1}, d_y \nu_{p-1})] \mathbf{c}_i \quad \forall i = 0, \dots, p-1 \quad (33)$$

for some arbitrary $p \times 1$ vector $\mathbf{c}_i$. In this case the eigenvector will not directly yield the μ frequency.

However, the matrix

$$\mathbf{K} [\mathbf{e}_0 \mid \dots \mid \mathbf{e}_{p-1}] = [\mathbf{a}_{M_b N_b}(d_x \mu_0, d_y \nu_0) \mid \dots \mid \mathbf{a}_{M_b N_b}(d_x \mu_{p-1}, d_y \nu_{p-1})] \mathbf{C} \quad (34)$$

has the Esprit structure. Therefore, applying Esprit in a row-wise fashion will yield the μ frequencies. Coupling the frequencies is not an issue because all of the corresponding ν frequencies are identical and already known.

Before proceeding, several aspects of the above development need to be emphasized. First, $\mathbf{K}$, the arbitrary basis for $\mathcal{S}_d$, is obtained as the orthogonal complement of the decimated telescoped noise eigenvectors $\mathbf{h}_j$. Second, the matrix products $\mathbf{J}_1 \mathbf{K}$ and $\mathbf{J}_2 \mathbf{K}$ are the first and last $M_b(N_b - 1)$

rows of $\mathbf{K}$, and can be effected without actually performing any matrix multiplies. Third, the matrix Ψ that rotates $\mathbf{J}_1\mathbf{K}$ into $\mathbf{J}_2\mathbf{K}$ is estimated by applying the Total Least Squares method of Golub and Van Loan [11] [12] to equation (29) (as in [4]). Finally, equation (29) has a unique solution for Ψ provided that the number of rows exceeds (or equals) the number of columns in $\mathbf{J}_1\mathbf{K}$. Since $\mathbf{J}_1\mathbf{K}$ is $M_b(N_b - 1) \times d$, TLS-Esprit can determine up to $M_b(N_b - 1)$ signal directions. It is also worth mentioning that since S_d has the Esprit structure any type of Esprit algorithm, such as PRO-Esprit [6], can be applied to $\mathbf{K}$.

7 Bandlimiting the Response

Thus far it has been assumed that beamformer employed is comprised of $M_b N_b$ beams that encompass the subband defined by $-\pi \left(\frac{M_b}{M} \right) \leq \mu \leq \pi \left(\frac{M_b}{M} \right)$ and $-\pi \left(\frac{N_b}{N} \right) \leq \nu \leq \pi \left(\frac{N_b}{N} \right)$. If the beams are insufficiently bandlimited or not centered at broadside, the beam space noise eigenvectors can be modulated to baseband and filtered prior to decimation to make the assumption valid. It is important to realize that the filtering process increases the length of the eigenvectors. In the one dimensional case Kautz [2] showed that a decimated version of the filter can be deconvolved from the decimated telescoped eigenvectors to remove most of this extra dimensionality. Furthermore since filtering and deconvolution are linear operations they can also be performed apriori on the telescoping matrix. For the 1D case this yields telescoped eigenvectors that are $(M_b + 1) \times 1$ and a resulting decimated signal subspace that is a $d + 1$ dimensional subspace of $M_b + 1$ dimensional space.

This extra dimensionality does not cause any problems for 2D Music. In fact, if the beamformer employed has the common out of band nulls property then filter nulls can be positioned to coincide with out of band peaks thereby effectively eliminating aliasing [5]. For 1D Esprit, this

extra dimensionality produces a Ψ matrix that is $(d + 1) \times (d + 1)$ and has an eigenvalue that is not related to a signal direction. Kautz argued that this extraneous eigenvalue is far removed from the unit circle, so it is easily identified and ignored. In the case of a rectangular array, the filtering is two dimensional, so after deconvolution the resulting eigenvectors are $(M_b + 1) \times (N_b + 1)$. This yields a $d + M_b + N_b + 1$ dimensional decimated signal subspace and Ψ has $M_b + N_b + 1$ extra eigenvalues. Kautz's argument that these additional eigenvalues are far removed from the unit circle is still valid, but now the eigenvalue decomposition is performed on a matrix that is $(d + M_b + N_b + 1) \times (d + M_b + N_b + 1)$ instead of $d \times d$. This is a nonnegligible increase in complexity that can be easily circumvented by improving the front end beamformer.

An obvious choice for the beamformer is $(M_b - 2) \times (N_b - 2)$ Hamming weighted orthonormal DFT beams centered at $\mu = -\pi(M_b - 3)/M \dots \pi(M_b - 3)/M$ and $\nu = -\pi(N_b - 3)/N \dots \pi(N_b - 3)/N$ because they have low sidelobes and common out of band nulls [5]. However this yields an even larger increase in complexity because the eigenvectors have length $M_b N_b$ but there are only $(M_b - 2)(N_b - 2)$ of them. Therefore S_d is a $d + 2M_b + 2N_b + 4$ dimensional subspace and Ψ has $2M_b + 2N_b + 4$ extra eigenvalues. So the beamformer employed must have $M_b \times N_b$ beams.

In section 11 several types of beamformers will be investigated to determine which ones yield the best performance. For the moment, it is sufficient to point out a few properties that need to be considered when choosing the beamformer. First, for the Esprit formulation no filtering is performed, so common out of band nulls are unnecessary. Second, larger main lobes are required to reduce the sidelobe ripple, but this yields beams that may extend outside the desired subband (see Figure 5). This is not a problem, since aliasing caused by these wider main lobes will only effect signals at the band edge. It is a well established fact that performance decays at the band edge even without aliasing, so subbands should be overlapped. Third, orthogonal beams are required for equation (16) to be valid, but this increases the ripple, and consequently the estimation error due

to aliasing. However, nonorthogonal beams introduce error because $\mathbf{R}_y = \mathcal{B}\mathbf{R}_s\mathcal{B}^H + \sigma^2\mathbf{W}_b^H\mathbf{W}_b$. One way to reduce the error due to nonorthogonal beams is to “clean” the matrix $\mathbf{R}_y$. This is mentioned in [6] for use in cases where the noise is not spatially white. Notice that nonorthogonal beams only cause a problem at low SNR.

8 Further Reductions in Complexity

In this section several remarkable computational savings that have been developed for 1D are extended to 2D. Since these are direct extensions of the 1D case and are given a thorough treatment elsewhere, the details will be omitted.

8.1 Real Covariance Processing

It has been observed [5] that for the 1D ULA, placing the reference point in the center of the array and making appropriate restrictions on the beamformer enables one to replace the EVD of $\mathbf{R}_y$ with the EVD of $\mathcal{R}\{y\}$. This effects a considerable reduction in complexity and is readily extended to the uniform rectangular array.

With the reference point in the center of the array, $r = (\frac{M-1}{2}\Delta_x, \frac{N-1}{2}\Delta_y, 0)$, the array manifold vectors have the form $\mathbf{A}_{MN}(\mu, \nu) = \mathbf{a}_M(\mu)\mathbf{a}_N^T(\nu)$ where

$$\begin{aligned} \mathbf{a}_M(\mu) &= \left[e^{-j(\frac{M-1}{2})\mu}, \dots, e^{j(\frac{M-1}{2})\mu} \right]^T \\ \mathbf{a}_N(\nu) &= \left[e^{-j(\frac{N-1}{2})\nu}, \dots, e^{j(\frac{N-1}{2})\nu} \right]^T. \end{aligned} \quad (35)$$

These steering vectors have the following conjugate centrosymmetric property:

$$\tilde{\mathbf{I}}_{MN}\mathbf{a}_{MN}(\mu, \nu) = (\tilde{\mathbf{I}}_N \otimes \tilde{\mathbf{I}}_M)(\mathbf{a}_N(\nu) \otimes \mathbf{a}_M(\mu)) = \mathbf{a}_N^*(\nu) \otimes \mathbf{a}_M^*(\mu) = \mathbf{a}_{MN}^*(\mu, \nu). \quad (36)$$

Where $\tilde{\mathbf{I}}_M$ is the $M \times M$ reverse permutation matrix, that “flips” the $M \times 1$ column vector. Since $\tilde{\mathbf{I}}_M$ is its own inverse and $\tilde{\mathbf{I}}_{MN} = \tilde{\mathbf{I}}_N \otimes \tilde{\mathbf{I}}_M$, applying a conjugate centrosymmetric beamformer yields

real valued beam space steering vectors.

$$\mathcal{B} = [\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H] \mathcal{A} = [\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H] \tilde{\mathbf{I}}_{MN} \tilde{\mathbf{I}}_{MN} \mathcal{A} = [\mathbf{W}_\nu^T \otimes \mathbf{W}_\mu^T] \mathcal{A}^* = \mathcal{B}^* \quad (37)$$

Using this in equation (16) and taking the real part of the beam space correlation matrix yields

$$\mathcal{R}e\{\mathbf{R}_y\} = \mathcal{R}e\left\{[\mathbf{W}_\nu^H \otimes \mathbf{W}_\mu^H] \mathcal{A} \mathbf{R}_s \mathcal{A}^H [\mathbf{W}_\nu \otimes \mathbf{W}_\mu] + \sigma^2 \mathbf{I}\right\} = \mathcal{B} \mathcal{R}e\{\mathbf{R}_s\} \mathcal{B}^T + \sigma^2 \mathbf{I}. \quad (38)$$

Therefore, the real part of the beam space correlation matrix has the desired eigen structure and the TLS-Esprit algorithm can be applied to the real part, instead of the “full blown”, correlation matrix.

8.2 Orthogonal Complement

Recall that columns of $\mathbf{K}$ form a basis for $\mathcal{S}_d$ and have thus far been obtained as the orthogonal complement of $\mathbf{H} = \mathbf{W}_t \mathbf{F}_n$. This requires a computationally intensive SVD on $\mathbf{W}_t \mathbf{F}_n$ to find $\mathbf{K}$. Kautz [2] noticed that an alternate basis for $\mathcal{S}_d$ can be obtained by applying a simple linear transformation to $\mathbf{F}_s$. To show this, let $\mathbf{Z} = \mathbf{W}_t (\mathbf{W}_t^H \mathbf{W}_t)^{-1}$ and notice that

$$(\mathbf{Z} \mathbf{f}_i)^H \mathbf{h}_j = \left[\mathbf{W}_t (\mathbf{W}_t^H \mathbf{W}_t)^{-1} \mathbf{f}_i \right]^H \mathbf{W}_t \mathbf{f}_j = \mathbf{f}_i^H \mathbf{f}_j = 0. \quad (39)$$

Hence the matrix $\mathbf{Z}$ maps a beam space signal eigenvector to element space (but not to its element space counterpart). Therefore $\mathbf{K}$ can be determined as

$$\mathbf{K} = \mathbf{Z} \mathbf{F}_s = [\mathbf{Z} \mathbf{f}_0 \mid \dots \mid \mathbf{Z} \mathbf{f}_{d-1}]. \quad (40)$$

In the event that filtering is employed, this transformation yields an insufficient basis for $\mathcal{S}_d$. However the remaining basis vectors can be obtained by precomputing the orthogonal complement of $\mathbf{W}_t$.

9 Algorithm

1. Precompute the beamforming matrix, $\mathbf{W}_b$, and $\mathbf{Z} = \mathbf{W}_t (\mathbf{W}_t^H \mathbf{W}_t)^{-1}$.
2. Store P snapshots of the array as the columns of $\mathbf{X}$, and form $\mathbf{Y} = \mathbf{W}_b^H \mathbf{X}$.
3. Compute the EVD of the real part of the beam space correlation matrix and form $\mathbf{K}$.

$$\mathcal{Re}\{\mathbf{R}_y\} = \frac{1}{P} \mathcal{Re}\{\mathbf{Y}\mathbf{Y}^H\} = \sum_{i=0}^{M_b N_b - 1} \lambda_i \mathbf{f}_i \mathbf{f}_i^H \quad \mathbf{K} = \mathbf{Z} [\mathbf{f}_0 \mid \dots \mid \mathbf{f}_{d-1}]$$

4. Form $\mathbf{K}_{12} = [\mathbf{K}_1 \mid \mathbf{K}_2]$, where $\mathbf{K}_1$ and $\mathbf{K}_2$ are the first and last $M_b(N_b - 1)$ rows of $\mathbf{K}$, and compute the EVD of $\mathbf{K}_{12}^H \mathbf{K}_{12} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^{-1}$.
5. Partition $\mathbf{Q}$ into $d \times d$ blocks and estimate $\mathbf{\Psi} = -\mathbf{Q}_{12} (\mathbf{Q}_{22})^{-1}$.
6. Compute the EVD of $\mathbf{\Psi}$ to obtain $\mathbf{\Upsilon}$ and estimate the ν frequencies.

$$\mathbf{\Psi} = \mathbf{E} \mathbf{\Upsilon} \mathbf{E}^{-1} \quad \mathbf{\Upsilon} = \text{diag}\{\gamma_0, \dots, \gamma_{d-1}\} \quad \nu_i = \frac{1}{d_y} \arg \gamma_i$$

7. For distinct ν_i estimate the μ frequencies as $\mu_i = \frac{1}{d_x} \arg \rho_i$. Where $\rho_i = \mathbf{l}_i^H \mathbf{P} \mathbf{l}_i$, $\mathbf{l}_i = \mathbf{K} \mathbf{e}_i$, and

$$\mathbf{P} = \frac{1}{(M_b - 1)N_b} [\mathbf{I}_N \otimes (\mathbf{\Gamma}_3^T \mathbf{\Gamma}_4)].$$

8. For repeated ν_i form $\mathbf{L} = \mathbf{K} [\mathbf{e}_1 \mid \dots \mid \mathbf{e}_{p-1}]$ and estimate the μ frequencies as $\mu_i = \frac{1}{d_x} \arg \rho_i$.

Where ρ_i are the eigenvalues of the matrix that rotates the $\mathbf{J}_3 \mathbf{L}$ into $\mathbf{J}_4 \mathbf{L}$ (steps 4 - 6).

10 Performance Analysis

A large portion of the 2D performance analysis is identical to the 1D analysis performed by Kautz [7] and Rao and Hari [8], so their work will be followed as much as possible. The bulk of the error analysis is included as Appendix A and the major results are presented in this section. To maintain

a consistent notation, estimated quantities will be denoted with a “hat”, and the error between estimated and actual quantities will be denoted with a “ Δ ” (e.g. $\Delta \mathbf{f}_i = \hat{\mathbf{f}}_i - \mathbf{f}_i$).

The primary source of error in the proposed algorithm results from the finite snapshot approximation of the beam space correlation matrix $\mathbf{R}_y$. Let $\mathbf{R}_y = \mathbf{R}_I + j\mathbf{R}_Q$ and recall that $\mathbf{R}_I$ is real, symmetric, and positive definite so its eigenvalues, λ , are real and the associated eigenvectors, $\mathbf{f}$, can be chosen to be real. Kautz has shown [7] that the error in the signal eigenvectors is asymptotically zero mean with with covariance

$$\mathbb{E} \{ \Delta \mathbf{f}_k \Delta \mathbf{f}_l^T \} = \frac{1}{P} \sum_{\substack{m=0 \\ m \neq k}}^{M_b N_b - 1} \sum_{\substack{n=0 \\ n \neq l}}^{M_b N_b - 1} \frac{\Gamma_{mnlk}}{(\lambda_k - \lambda_m)(\lambda_l - \lambda_n)} \mathbf{f}_m \mathbf{f}_n^T \quad k, l = 0, \dots, d-1 \quad (41)$$

$$\begin{aligned} \Gamma_{mnlk} = & \frac{1}{2} \left[\lambda_k \lambda_l \delta_{ml} \delta_{nk} + \lambda_k \lambda_m \delta_{mn} \delta_{kl} + (\mathbf{f}_m^T \mathbf{R}_Q \mathbf{f}_l) (\mathbf{f}_k^T \mathbf{R}_Q \mathbf{f}_n) (1 - \delta_{ml}) (1 - \delta_{kn}) \right. \\ & \left. + (\mathbf{f}_m^T \mathbf{R}_Q \mathbf{f}_n) (\mathbf{f}_k^T \mathbf{R}_Q \mathbf{f}_l) (1 - \delta_{mn}) (1 - \delta_{kl}) \right]. \end{aligned} \quad (42)$$

It should be noted that the multiplicative factors of the form $(1 - \delta_{\bullet\bullet})$ can be removed. They are only included to emphasize the fact that $\mathbf{f}_k^T \mathbf{R}_Q \mathbf{f}_k = 0$.

This error propagates through to γ and $\mathbf{e}$, the eigenvalues and right eigenvectors of Ψ , as

$$\mathbb{E} \{ |\Delta \gamma_i|^2 \} = \boldsymbol{\alpha}_{ii}^H \left[\sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_i^*(l) \mathbb{E} \{ \Delta \mathbf{f}_k \Delta \mathbf{f}_l^T \} \right] \boldsymbol{\alpha}_{ii} \quad (43)$$

$$(\gamma_i^*)^2 \mathbb{E} \{ (\Delta \gamma_i)^2 \} = \boldsymbol{\alpha}_{ii}^H \left[\sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_i(l) \mathbb{E} \{ \Delta \mathbf{f}_k \Delta \mathbf{f}_l^T \} \right] \boldsymbol{\alpha}_{ii}^* \quad (44)$$

$$\mathbb{E} \{ \Delta \mathbf{e}_i \Delta \mathbf{e}_j^H \} = \sum_{\substack{m=0 \\ m \neq i}}^{d-1} \sum_{\substack{n=0 \\ n \neq j}}^{d-1} \left\{ \frac{(\gamma_i \gamma_j^*)}{(\gamma_i - \gamma_m)(\gamma_j - \gamma_n)^*} \left[\boldsymbol{\alpha}_{im}^H \left(\sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_j^*(l) \mathbb{E} \{ \Delta \mathbf{f}_k \Delta \mathbf{f}_l^T \} \right) \boldsymbol{\alpha}_{jn} \right] \mathbf{e}_m \mathbf{e}_n^H \right\} \quad (45)$$

$$\mathbb{E} \{ \Delta \mathbf{e}_i \Delta \mathbf{e}_j^T \} = \sum_{\substack{m=0 \\ m \neq i}}^{d-1} \sum_{\substack{n=0 \\ n \neq j}}^{d-1} \left\{ \frac{(\gamma_i \gamma_j)}{(\gamma_i - \gamma_m)(\gamma_j - \gamma_n)} \left[\boldsymbol{\alpha}_{im}^H \left(\sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_j(l) \mathbb{E} \{ \Delta \mathbf{f}_k \Delta \mathbf{f}_l^T \} \right) \boldsymbol{\alpha}_{jn}^* \right] \mathbf{e}_m \mathbf{e}_n^T \right\} \quad (46)$$

$$\boldsymbol{\alpha}_{ij} = \mathbf{Z}^H (\mathbf{J}_1 - \gamma_i^* \mathbf{J}_2)^H (\mathbf{K}_1^+)^H \mathbf{q}_j \quad (47)$$

The relationship between $\Delta \mathbf{e}_i$ and $\Delta \rho_i$ is given by

$$\mathbb{E} \{ |\Delta \rho_i|^2 \} = \mathbf{e}_i^H \mathbf{K}^H \mathbf{P} \mathbf{K} \mathbb{E} \{ \Delta \mathbf{e}_i \Delta \mathbf{e}_i^H \} \mathbf{K}^H \mathbf{P}^T \mathbf{K} \mathbf{e}_i + \mathbf{e}_i^H \mathbf{K}^H \mathbf{P}^T \mathbf{K} \mathbb{E} \{ \Delta \mathbf{e}_i \Delta \mathbf{e}_i^T \} \mathbf{K}^H \mathbf{P} \mathbf{K} \mathbf{e}_i$$

$$+2 \operatorname{Re} \left\{ \mathbf{e}_i^H \mathbf{K}^H \mathbf{P} \mathbf{K} \mathbf{E} \left\{ \Delta \mathbf{e}_i \Delta \mathbf{e}_i^T \right\} \mathbf{K}^T \mathbf{P} \mathbf{K} \mathbf{e}_i \right\} \quad (48)$$

$$\begin{aligned} \mathbf{E} \left\{ (\Delta \rho_i)^2 \right\} &= \mathbf{e}_i^H \mathbf{K}^H \mathbf{P} \mathbf{K} \mathbf{E} \left\{ \Delta \mathbf{e}_i \Delta \mathbf{e}_i^T \right\} \mathbf{K}^T \mathbf{P}^T \mathbf{K}^* \mathbf{e}_i + \mathbf{e}_i^T \mathbf{K}^T \mathbf{P}^T \mathbf{K}^* \mathbf{E} \left\{ \Delta \mathbf{e}_i \Delta \mathbf{e}_i^T \right\}^* \mathbf{K}^H \mathbf{P} \mathbf{K} \mathbf{e}_i \\ &\quad + 2 \operatorname{Re} \left\{ \mathbf{e}_i^T \mathbf{K}^T \mathbf{P} \mathbf{K} \mathbf{E} \left\{ \Delta \mathbf{e}_i \Delta \mathbf{e}_j^H \right\} \mathbf{K}^H \mathbf{P} \mathbf{K} \mathbf{e}_i \right\} \end{aligned} \quad (49)$$

$$\mathbf{P} = \frac{1}{(M_b - 1)N_b} \mathbf{I}_{N_b} \otimes (\mathbf{\Gamma}_3^T \mathbf{\Gamma}_4). \quad (50)$$

It remains to show how $\Delta \rho_i$ and $\Delta \gamma_i$ relate to $\Delta \mu_i$ and $\Delta \nu_i$. This was done by Rao and Hari [8].

$$\mathbf{E} \left\{ (\Delta \nu_i)^2 \right\} = \frac{\mathbf{E} \left\{ |\Delta \gamma_i|^2 \right\} - \operatorname{Re} \left\{ (\gamma_i^*)^2 \mathbf{E} \left\{ (\Delta \gamma_i)^2 \right\} \right\}}{2(d_y)^2} \quad (51)$$

$$\mathbf{E} \left\{ (\Delta \mu_i)^2 \right\} = \frac{\mathbf{E} \left\{ |\Delta \rho_i|^2 \right\} - \operatorname{Re} \left\{ (\rho_i^*)^2 \mathbf{E} \left\{ (\Delta \rho_i)^2 \right\} \right\}}{2(d_x)^2}. \quad (52)$$

Combining equations 41 through 52 produces the desired asymptotic error characteristics of the signal frequency estimates. However, This does not yield any insightful information. For the case of d uncorrelated sources with equal signal power σ_s , and noise power σ_n , the resulting asymptotic error reduces to

11 Simulations

Various computer simulations were performed to verify the efficacy of the proposed 2D Esprit algorithm. Unless stated otherwise, all experiments simulate a 32×32 array with half wavelength spacing and 3 equal power sources, σ_s^2 . The beam space correlation matrix is estimated from 32 snapshots of the array and 200 trials are executed for each particular point of interest. The front end beamformer consists of 64 beams centered at broadside, so the subband being probed is $-\frac{\pi}{4} \leq (\mu, \nu) \leq \frac{\pi}{4}$, and the maximal decimation rate of $d_x = d_y = 4$ is used. To investigate the effects mentioned in section 7, three separate types of beams are simulated, DFT beams, Hamming beams, and Ortonormal Hamming beams (see Figure 5 for 1D plots). Finally, it is important to note that the error criterion used to evaluate the estimator is the average RMS error between the

actual signal frequencies and their estimates, $\overline{\text{rms}} = \frac{1}{d} \sum_{i=0}^{d-1} \sqrt{(\mu_i - \hat{\mu}_i)^2 + (\nu_i - \hat{\nu}_i)^2}$, and SNR refers to the per signal per element signal to noise ratio, $\text{SNR} = 10 \log \frac{\sigma_s^2}{\sigma_n^2}$.

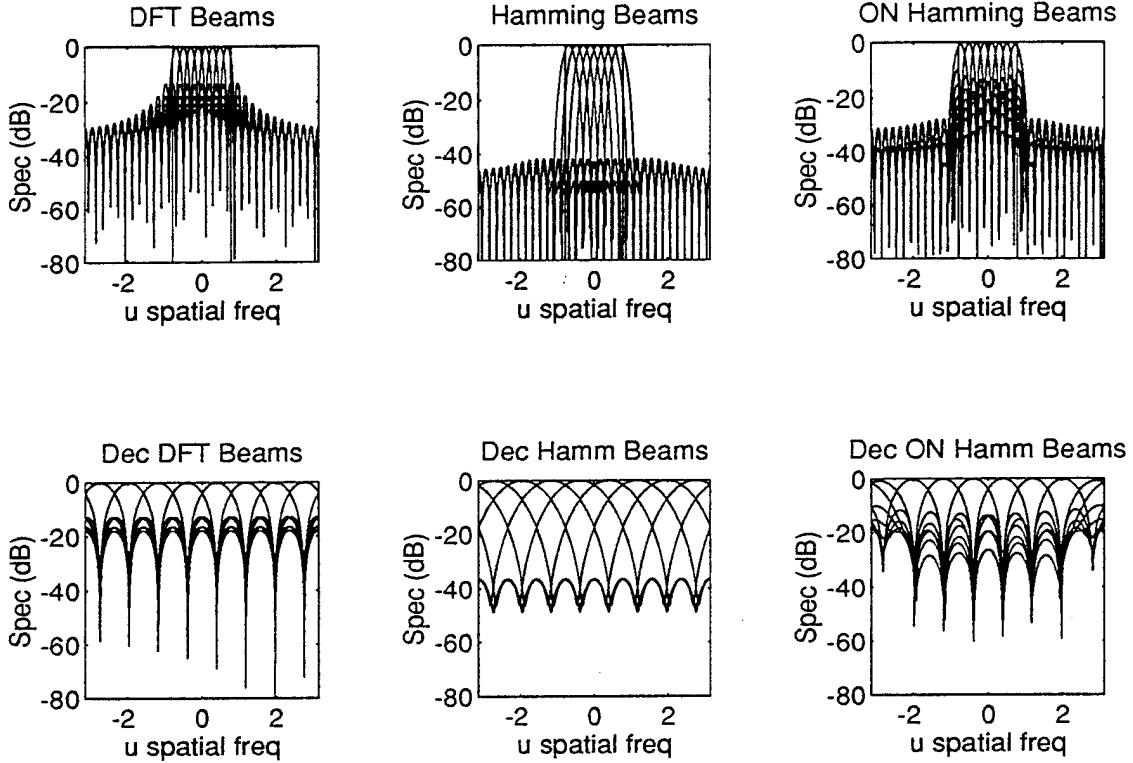


Figure 5: Beams

Before simulating Esprit, it is worthwhile to simulate 2D Multirate Spectral Music (Eqn. 23). The three signals simulated had 0 dB SNR and spatial frequencies $(\alpha, -\alpha)$, $(-\alpha, -\alpha)$, and $(-\alpha, \alpha)$, where $\alpha = \frac{2\pi}{32} = 0.1963$. Since the main lobe width of the Hamming beams is $\frac{4\pi}{32}$, these signals are said to have 100% beam width separation (see Figure 6). These plots show that 2D Multirate processing does indeed work, and the resulting spectral nulls are moved to $(\pm d_x \alpha, \pm d_y \alpha) = (\pm 0.7852, \pm 0.7852)$. For comparison purposes, Esprit was simulated with the same parameters and the results displayed as scatter plots (see Figure 7). Notice that the scatter plot verifies the automatic coupling properties of the proposed Esprit algorithm, but shows that DFT beams have a slight bias.

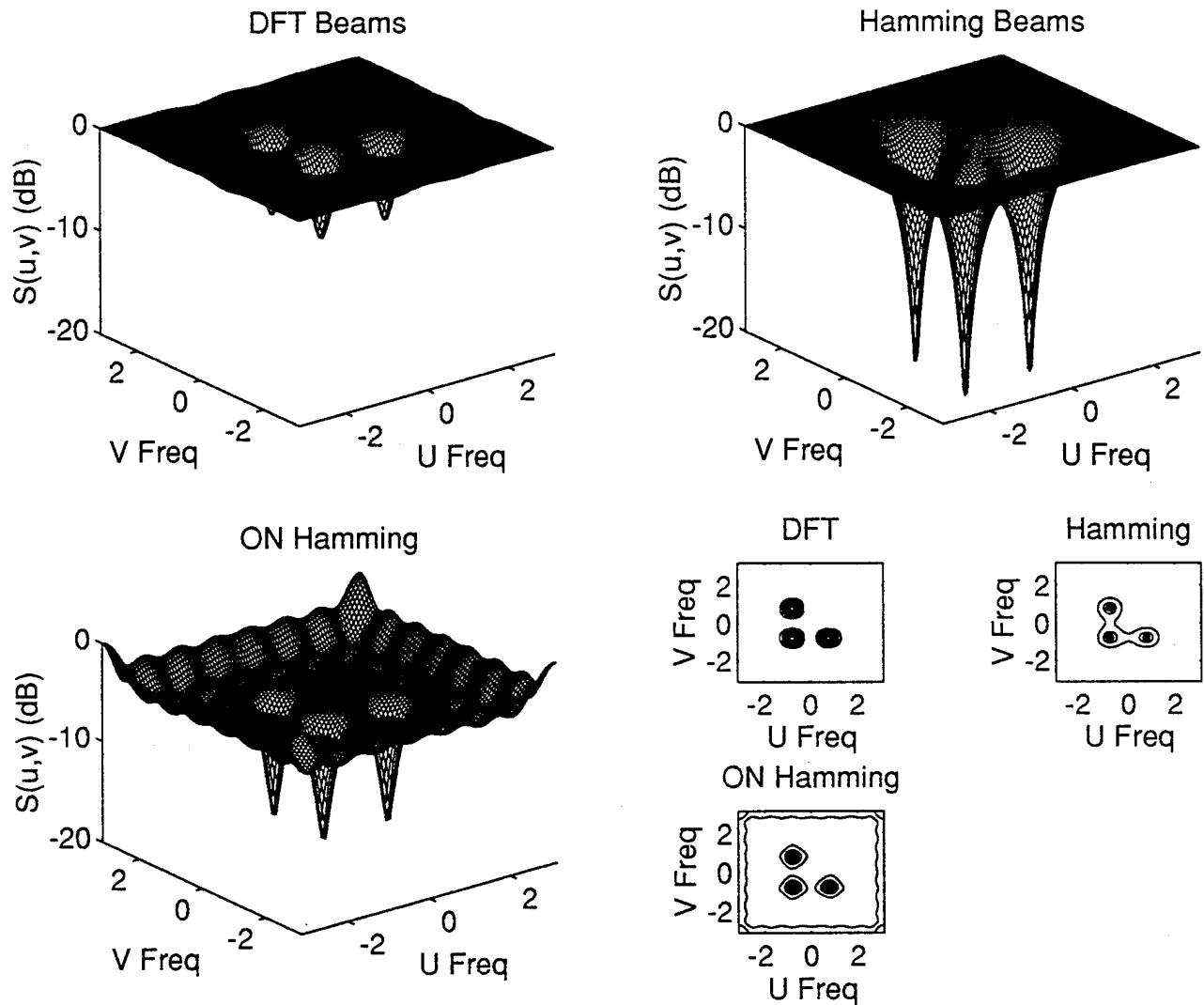


Figure 6: Beam Space Music Null Spectrum for 3 sources with 0dB SNR and 32 Array Snapshots

In section 7 it was stated that the orthogonality of the beamformer is only a consideration at low SNR, and at high SNR the dominant factor is the height of the sidelobes. Figure 6 illustrates this point. Hamming weighted beams have the lowest sidelobes and yield the deepest nulls. To further investigate the SNR dependence of the estimator, the same three signals were simulated and the SNR was varied from -30dB to 0dB (see Figure 8 first row). This figure shows that for SNR values below -10dB measurement noise dominates so DFT beams perform better, but for SNR values above -10dB aliasing due to high sidelobess dominates so Hamming beams perform better.

In an effort to investigate the performance of 2D Esprit for closely spaced sources, this

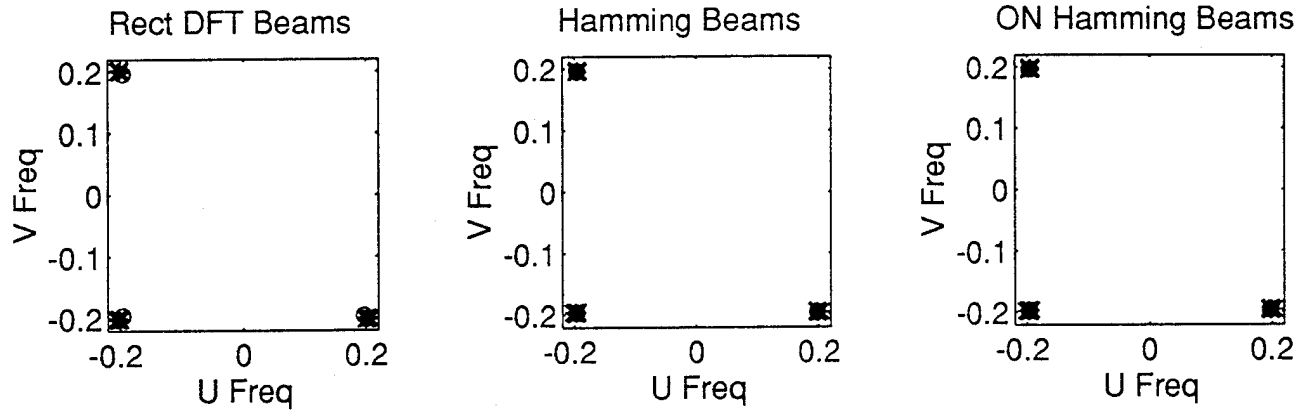


Figure 7: Scatter Plots: Esprit for 3 sources with 0dB SNR and 32 Snapshots

simulation was repeated with the signal separation reduced to 50% of the beamwidth ($\alpha = \frac{\pi}{M} = 0.0982$) (see Figure 8 second row). The performance actually improved. To understand why this happened, it is necessary to investigate the beamformer performance with respect to signal location.

It is a well established fact that the performance of 1D beamformers decay near the band edges. To see how the 2D beamformer performs, one signal was simulated and its position was varied from the center of the band to the band edge along the μ axis. This was repeated, varying the signal along the $\mu = \nu$ diagonal (see Figure 9). The performance does indeed decay, hence subbands should be overlapped.

Lastly the number of array snapshots was varied. The first row of Figure 10 depicts the original 3 signals with 0 dB SNR, and the second row depicts the same three signals with -20 dB SNR.

12 Conclusion

The proposed 2D Multirate Esprit Algorithm has been shown to work well. Let's submit this paper!

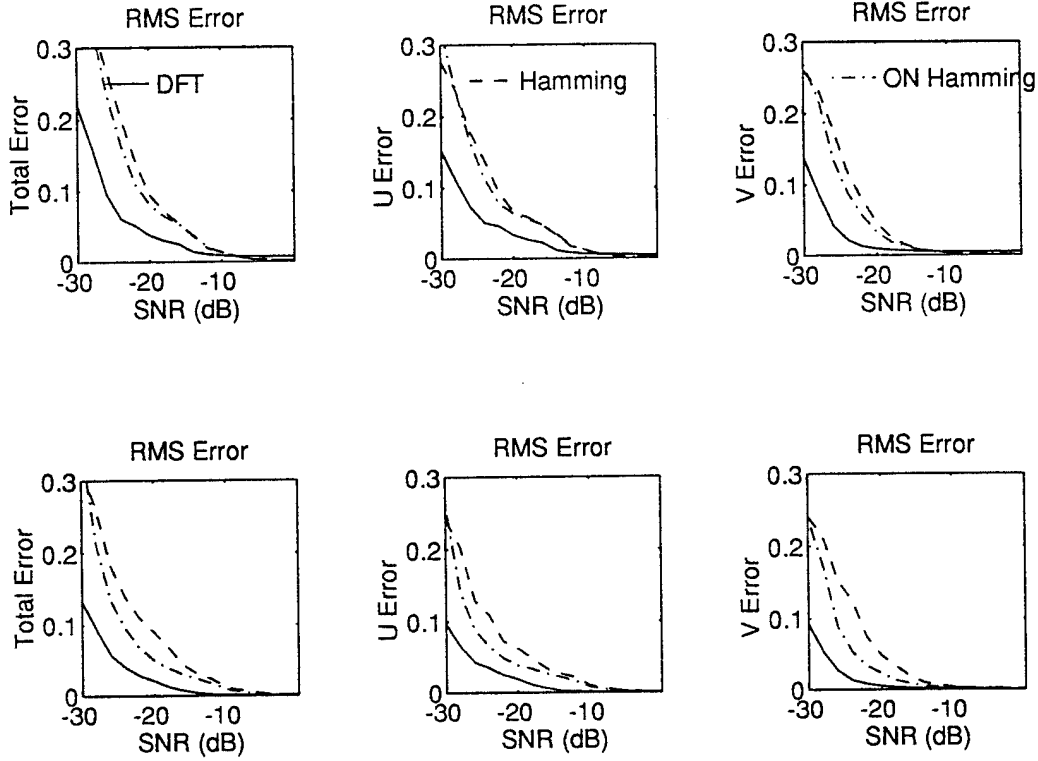


Figure 8: RMS Error vs SNR: Esprit for 3 sources, 0dB SNR, 32 Snapshots and 200 trials per SNR value. Row 1 has 100% BW separation, and Row 2 has 50% BW separation.

A Characterizing the Asymptotic Error

A detailed characterization of the asymptotic error in the eigenvectors of the real part of the beam space correlation matrix can be found in [7], so the current development will begin by showing how $\Delta \mathbf{f}$ effects the estimate of Ψ . Recall that $\mathbf{K} = [\mathbf{Z}\mathbf{f}_0 \mid \dots \mid \mathbf{Z}\mathbf{f}_{d-1}] = \mathbf{Z}\mathbf{F}_s$, and $\mathbf{K}$ is divided into $\mathbf{K}_1 = \mathbf{J}_1\mathbf{K}$ and $\mathbf{K}_2 = \mathbf{J}_2\mathbf{K}$. The error in these matrices is given by

$$\Delta \mathbf{K}_i = \hat{\mathbf{K}}_i - \mathbf{K}_i = \mathbf{J}_i \hat{\mathbf{K}} - \mathbf{J}_i \mathbf{K} = \mathbf{J}_i \mathbf{Z} \hat{\mathbf{F}}_s - \mathbf{J}_i \mathbf{Z} \mathbf{F}_s = \mathbf{J}_i \mathbf{Z} \Delta \mathbf{F}_s \quad i = 1, 2. \quad (53)$$

Using $\hat{\mathbf{K}}_1$ and $\hat{\mathbf{K}}_2$, $\hat{\Psi}$ is determined as the solution to $\hat{\mathbf{K}}_2 = \hat{\mathbf{K}}_1 \hat{\Psi}$. Therefore,

$$\mathbf{K}_2 + \Delta \mathbf{K}_2 = (\mathbf{K}_1 + \Delta \mathbf{K}_1)(\Psi + \Delta \Psi)$$

$$\mathbf{K}_2 + \Delta \mathbf{K}_2 = \mathbf{K}_1 \Psi + \mathbf{K}_1 \Delta \Psi + \Delta \mathbf{K}_1 \Psi + \Delta \mathbf{K}_1 \Delta \Psi$$

$$\Delta \mathbf{K}_2 \approx \mathbf{K}_1 \Delta \Psi + \Delta \mathbf{K}_1 \Psi$$

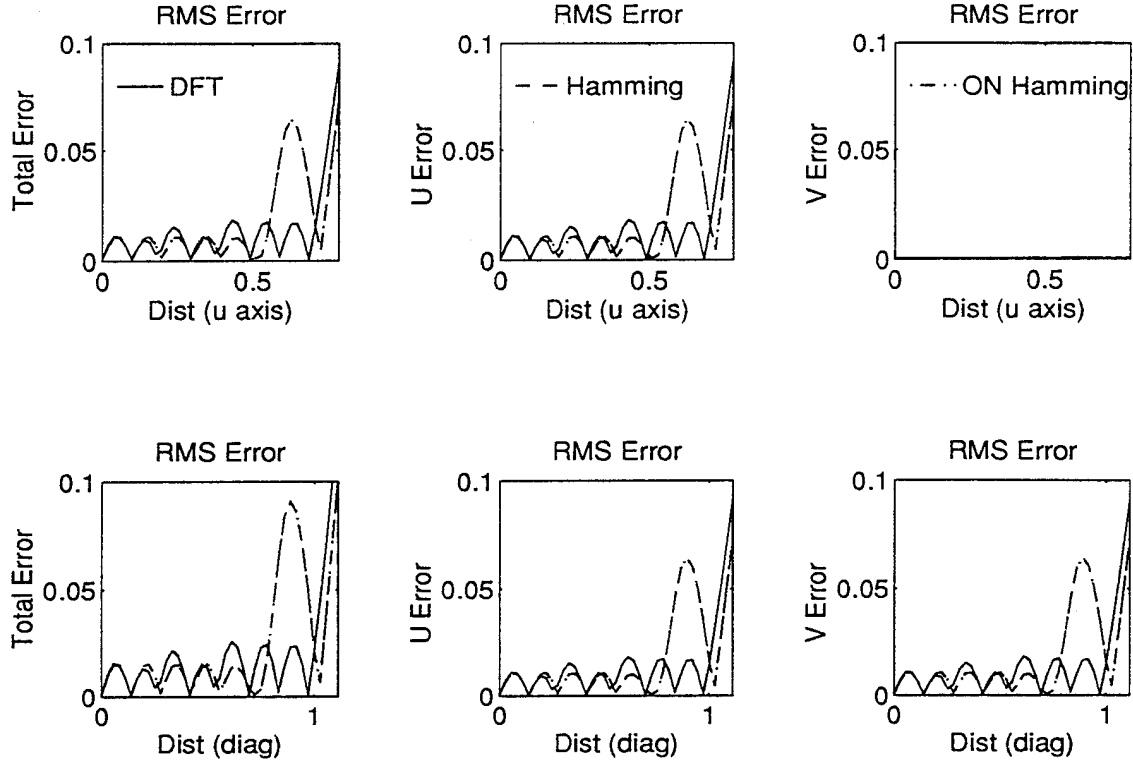


Figure 9: RMS Error vs Position: Esprit for 1 source, 0dB SNR, 32 Snapshots and 200 trials per Frequency. Row 1 varies the source frequency along the μ axis and Row 2 varies it along the $\mu = \nu$ diagonal.

$$\mathbf{K}_1 \Delta \Psi = \Delta \mathbf{K}_2 - \Delta \mathbf{K}_1 \Psi.$$

Hence $\Delta \Psi$ is given by

$$\Delta \Psi = \mathbf{K}_1^+ \Delta \mathbf{K}_2 - \mathbf{K}_1^+ \Delta \mathbf{K}_1 \Psi = \mathbf{K}_1^+ \mathbf{J}_2 \mathbf{Z} \Delta \mathbf{F}_s - \mathbf{K}_1^+ \mathbf{J}_1 \mathbf{Z} \Delta \mathbf{F}_s \Psi. \quad (54)$$

Where $\mathbf{K}_1^+ = (\mathbf{K}_1^H \mathbf{K}_1)^{-1} \mathbf{K}_1^H$ is the pseudo-inverse of $\mathbf{K}_1$. Rao and Hari [8] have shown that this error is valid for both Least Squares Esprit and Total Least Squares Esprit.

To find expressions for the error in the eigenvalues and eigenvectors of Ψ , recall [14] that if a matrix has the form $\mathbf{A} + \epsilon \mathbf{B}$ with eigenvalues $\gamma_i(\epsilon)$, right eigenvectors $\mathbf{e}_i(\epsilon)$ and left eigenvectors $\mathbf{q}_i(\epsilon)$, then

$$\gamma_i(\epsilon) = \gamma_i + \epsilon \frac{\mathbf{q}_i^H \mathbf{B} \mathbf{e}_i}{\mathbf{q}_i^H \mathbf{e}_i} \quad \text{and} \quad \mathbf{e}_i(\epsilon) = \mathbf{e}_i + \epsilon \sum_{\substack{j=1 \\ j \neq i}}^n \frac{(\mathbf{q}_j^H \mathbf{B} \mathbf{e}_i) \mathbf{e}_j}{(\gamma_i - \gamma_j) (\mathbf{q}_i^H \mathbf{e}_i)}.$$

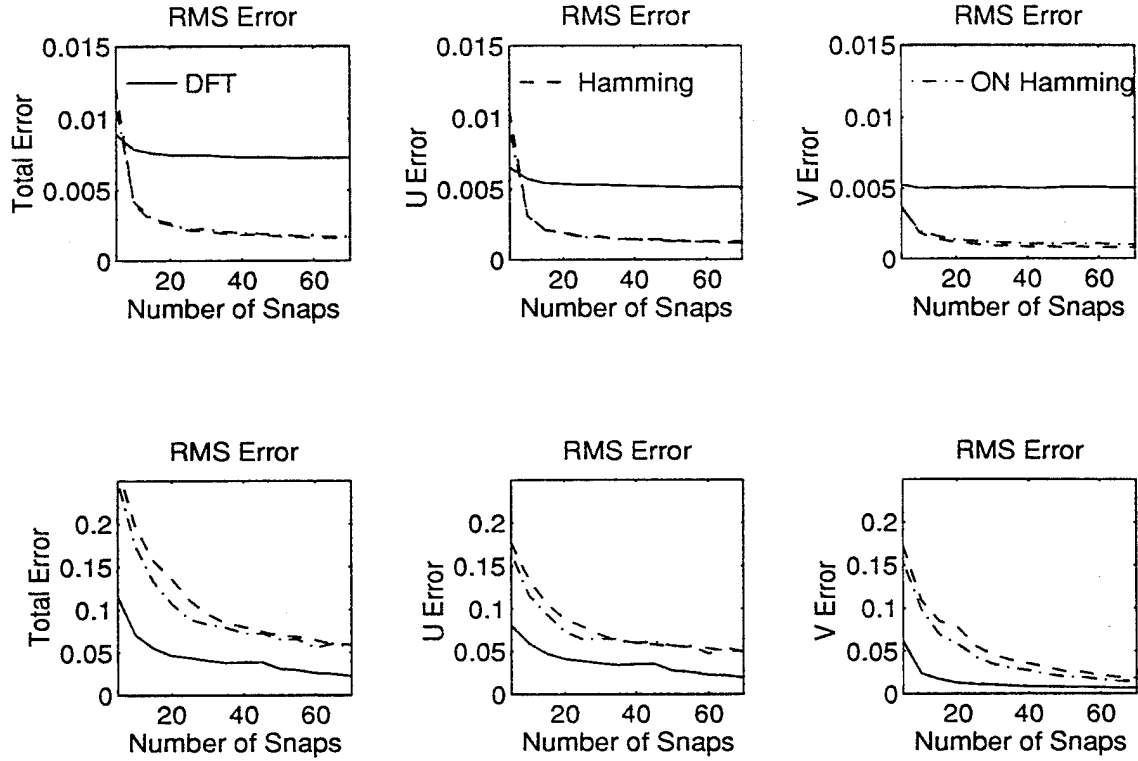


Figure 10: Performance vs Snapshots: Esprit with 3 sources, 32 Snapshots and 200 trials per Frequency. Row 1 has 0dB SNR and row 2 has -20dB SNR

In this case $\mathbf{A} = \mathbf{\Psi}$, $\mathbf{B} = \Delta\mathbf{\Psi}$, and $\mathbf{q}_i^H \mathbf{e}_i = 1$, so the eignvalue error is given by

$$\Delta\gamma_i = \mathbf{q}_i^H [\mathbf{K}_1^+ \Delta\mathbf{K}_2 - \mathbf{K}_1^+ \Delta\mathbf{K}_1 \mathbf{\Psi}] \mathbf{e}_i = -\gamma_i \mathbf{q}_i^H \mathbf{K}_1^+ (\mathbf{J}_1 - \gamma_i^* \mathbf{J}_2) \Delta\mathbf{K} \mathbf{e}_i = -\gamma_i \alpha_{ii}^H \Delta\mathbf{F}_s \mathbf{e}_i. \quad (55)$$

Where α_{ij} has been defined as $\alpha_{ij} = \mathbf{Z}^H (\mathbf{J}_1 - \gamma_i^* \mathbf{J}_2)^H (\mathbf{K}_1^+)^H \mathbf{q}_j$. From (55), it is easily verified that

$$|\Delta\gamma_i|^2 = \alpha_{ii}^H [\Delta\mathbf{F}_s \mathbf{e}_i \mathbf{e}_i^H \Delta\mathbf{F}_s^T] \alpha_{ii} = \alpha_{ii}^H \sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_i^*(l) \Delta f_k \Delta f_l^T \alpha_{ii} \quad (56)$$

$$(\Delta\gamma_i)^2 = (\gamma_i)^2 \alpha_{ii}^H [\Delta\mathbf{F}_s \mathbf{e}_i \mathbf{e}_i^T \Delta\mathbf{F}_s^T] \alpha_{ii}^* = (\gamma_i)^2 \alpha_{ii}^H \sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_i(l) \Delta f_k \Delta f_l^T \alpha_{ii}^*. \quad (57)$$

Now consider the right eigenvector.

$$\Delta\mathbf{e}_i = \sum_{\substack{j=0 \\ j \neq i}}^{d-1} \frac{(\mathbf{q}_j^H \Delta\mathbf{\Psi}_i \mathbf{e}_i) \mathbf{e}_j}{(\gamma_i - \gamma_j)} = \sum_{\substack{j=0 \\ j \neq i}}^{d-1} \left\{ \frac{-\gamma_i \alpha_{ij}^H \Delta\mathbf{F}_s \mathbf{e}_i}{(\gamma_i - \gamma_j)} \mathbf{e}_j \right\} \quad (58)$$

Therefore

$$\Delta \mathbf{e}_i \Delta \mathbf{e}_j^H = \sum_{m=0}^{d-1} \sum_{\substack{n=0 \\ m \neq i, n \neq j}}^{d-1} \left\{ \frac{(\gamma_i \gamma_j^*)}{(\gamma_i - \gamma_m)(\gamma_j - \gamma_n)^*} \left[\alpha_{im}^H \left(\sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_j^*(l) \Delta \mathbf{f}_k \Delta \mathbf{f}_l^H \right) \alpha_{jn} \right] \mathbf{e}_m \mathbf{e}_n^H \right\} \quad (59)$$

$$\Delta \mathbf{e}_i \Delta \mathbf{e}_j^T = \sum_{m=0}^{d-1} \sum_{\substack{n=0 \\ m \neq i, n \neq j}}^{d-1} \left\{ \frac{(\gamma_i \gamma_j)}{(\gamma_i - \gamma_m)(\gamma_j - \gamma_n)} \left[\alpha_{im}^H \left(\sum_{k=0}^{d-1} \sum_{l=0}^{d-1} e_i(k) e_j(l) \Delta \mathbf{f}_k \Delta \mathbf{f}_l^T \right) \alpha_{jn}^* \right] \mathbf{e}_m \mathbf{e}_n^T \right\} \quad (60)$$

Taking the expectation of equations (56, 57, 59, 60) yields equations (43 - 46). With the error in the eigenvectors of Ψ characterized, the error in ρ_i can now be determined.

$$\hat{\rho}_i = \hat{\mathbf{l}}_i^H \mathbf{P} \hat{\mathbf{l}}_i$$

$$\rho_i + \Delta \rho_i = (\mathbf{l}_i + \Delta \mathbf{l}_i)^H \mathbf{P} (\mathbf{l}_i + \Delta \mathbf{l}_i) = \mathbf{l}_i^H \mathbf{P} \mathbf{l}_i + \mathbf{l}_i^H \mathbf{P} \Delta \mathbf{l}_i + \Delta \mathbf{l}_i^H \mathbf{P} \mathbf{l}_i + \Delta \mathbf{l}_i^H \mathbf{P} \Delta \mathbf{l}_i$$

$$\Delta \rho_i \approx \mathbf{l}_i^H \mathbf{P} \Delta \mathbf{l}_i + \Delta \mathbf{l}_i^H \mathbf{P} \mathbf{l}_i = \mathbf{e}_i^H \mathbf{K}^H \mathbf{P} \mathbf{K} \Delta \mathbf{e}_i + \Delta \mathbf{e}_i^H \mathbf{K}^H \mathbf{P} \mathbf{K} \mathbf{e}_i \quad (61)$$

Equations (48 - 49) come directly from (61).

It remains to show how $\Delta \gamma_i$ and $\Delta \rho_i$ effect $\Delta \nu_i$ and $\Delta \mu_i$. Recall that ideally $\gamma_i = e^{j d_y \nu_i}$, but due to errors $\hat{\gamma}_i = \hat{r}_i e^{j d_y \hat{\nu}_i} = \hat{r}_i e^{j d_y (\nu_i + \Delta \nu_i)}$. Consider $\gamma = r e^{j d_y \nu}$ and notice that

$$\partial_\gamma = \partial_r e^{j d_y \nu} + j d_y r e^{j d_y \nu} \partial_\nu = e^{j d_y \nu} \partial_r + j d_y \gamma \partial_\nu$$

$$\partial_{\gamma_i} = e^{j d_y \nu_i} \partial_{r_i} + j d_y \gamma_i \partial_{\nu_i} = \gamma_i (\partial_{r_i} + j d_y \partial_{\nu_i})$$

Therefore

$$|\partial_{\gamma_i}|^2 = (\partial_{r_i})^2 + (d_y \partial_{\nu_i})^2$$

$$(\partial_{\gamma_i})^2 = (\gamma_i)^2 [(\partial_{r_i})^2 - (d_y \partial_{\nu_i})^2 + j (2 d_y \partial_{r_i} \partial_{\nu_i})]$$

$$\mathcal{Re} \{ (\gamma_i^*)^2 (\partial_{\gamma_i})^2 \} = [(\partial_{r_i})^2 - (d_y \partial_{\nu_i})^2]$$

and consequently

$$(\Delta \nu_i)^2 = \frac{|\Delta \gamma_i|^2 - \mathcal{Re} \{ (\gamma_i^*)^2 (\Delta \gamma_i)^2 \}}{2 (d_y)^2} \quad (62)$$

This results differs slightly from that obtained by Rao and Hari [8]. They were concerned with the direction angle θ not the frequency $\nu = \sin \theta$. Therefore, they had a $(\cos \theta)^2$ in the denominator due to the fact that $\partial_\nu = \cos \theta \partial_\theta$.

References

- [1] R. O. Schmidt, "A Signal Subspace Approach to Multiple Emitter Location and Spectral estimation" Ph.D. dissertation, Stanford University, Stanford, CA, 1981
- [2] G. M. Kautz and M. D. Zoltowski, "Beamspace DOA Estimation Featuring Multirate Eigenvector processing," *Submitted to IEEE Trans, on Signal Processing*, May. 1994.
- [3] M. D. Zoltowski, J. V. Krogmeier, and G. M. Kautz, "Novel Multirate Processing of Beamspace Noise Eigenvectors," *Submitted to IEEE Signal Processing Letters*, Jan. 1994.
- [4] R. Roy and T. Kailath, "ESPRIT-Estimation of Signal Parameters Via Rotational Invariance Techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984-995, July 1989.
- [5] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Beamspace Root-MUSIC," *IEEE Trans. on Signal Processing*, vol. 41, no. 1, pp. 344-364, Jan. 1993.
- [6] M. D. Zoltowski and D. Stavrinos, "Sensor Array Signal Processing Via a Procrustes Rotations Based Eigenanalysis of the ESPRIT Data Pencil," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 6, pp. 832-861, June 1989.
- [7] G. M. Kautz and M. D. Zoltowski, "Performance Analysis of MUSIC Employing Conjugate Symmetric Beamformers," *Submitted to IEEE Trans, on Signal Processing*, May. 1993.
- [8] B. D. Rao and K. V. S. Hari, "Performance Analysis of ESPRIT and TAM in Determining the Direction of Arrival of Plane Waves in Noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 12, pp. 1990-1995, Dec. 1989.

- [9] J. W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Trans. on Circuits and Systems*, vol. cas-25, no. 9, pp. 772-781, Sept. 1978.
- [10] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 2, pp. 387-392, April 1985.
- [11] G. H. Golub and C. F. Van Loan, "An Analysis of the Total Least Squares Problem," *SIAM J. Numerical Anal.*, vol. 17, no. 6, pp. 883-893, Dec 1980.
- [12] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore. MD: Johns Hopkins University Press. 1984
- [13] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, New Jersey: Prentice Hall, 1993.
- [14] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. Clarendon Press Oxford 1965
- [15] A. Nonymous, *Zoltowski mentioned that someone at MIT has developed a 2-D rooting algorithm, but it has not been published and we have not seen it.*

4 Real-Time Frequency And 2-D Angle Estimation With Sub-Nyquist Spatio-Temporal Sampling

An algorithm has been developed for real-time estimation of the frequency and azimuth and elevation angles of each signal incident upon 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 .5 ns, 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 unambiguous 2-D angle estimation 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 automatically paired with corresponding frequency estimates despite aliasing. Simulations are presented demonstrating the capabilities of the algorithm.

4.1 Introduction

4.2 Spatio-Temporal Sampling and Data Model

4.3 ESPRIT Based Frequency Estimation With Temporal Undersampling

4.4 2-D Angle Estimation With Spatial Undersampling Via PRO-ESPRIT and Integer Search Formulation

4.4.1 Estimation of the Array Manifold for Each Source

4.4.2 Prescription for Nonuniform Element Spacing Facilitating Nonambiguous Angle Estimation

4.4.3 Integer Search Algorithm for Direction Cosine Estimation

4.5 Simulation Examples

4.6 Final Comments

4.7 References

4.8 Computation of Cramer Rao Lower Bound for Frequency and 2D Angle Estimation

4.9 Figures

1 Introduction

The problem under investigation is that of real-time estimation of the frequency and azimuth and elevation angles of each signal incident upon 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 Figure 1 [1], the receiver output, after conversion to baseband, is sampled at a rate of 250 MHz, 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 Figure 3.

Note that the aliasing function plotted in Figure 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, a real-valued signal, is sampled and input to the system. Note, 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 2.

The procedure for frequency estimation with Sub-Nyquist temporal sampling developed within may be easily adapted for narrowband 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.

In order to estimate the baseband frequency of each signal despite aliasing, the baseband 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 Figure 1. The time-delay, τ , is chosen 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 Figure 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 baseband 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, and (iii) apply ESPRIT to a small set of DFT values around each peak.

In Section 3, 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 computationally 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 Figure 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 4, 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 mutual 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 Figure 2. The interelement spacings along either axis is much greater than a half-wavelength, particularly at 18 GHz. In Section 4, we develop (i) a prescription for interelement spacings for nonambiguous angle estimation and (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 4 assumes a small number of antenna elements due to cost and space limitations on the antenna platform attached to the aircraft. Also, 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. Section 5 presents simulations that demonstrate the power of the overall frequency and 2-D angle estimation algorithm summarized in the flowchart presented in Figure 5.

2. Spatio-Temporal Sampling and Data Model

The parameters for the prototype sub-Nyquist spatio-temporal sampling system are indicated in Figure 1. We concentrate on signal parameter estimation for a particular 1 GHz baseband 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 is the bandwidth, f_c is the carrier frequency, L is the length of the array, λ is the wavelength, and $\cos\theta$ is the 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 wideband 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 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).

$$\cos(2\pi F_j t + \phi)|_{t=n/F_s} = \cos(2\pi \frac{F_j}{F_s} n + \phi) = \cos[2\pi f_j n + \phi] \quad \text{for } 0 < F_j < \frac{F_s}{2} \quad (1)$$

$$= \cos[2\pi(\frac{F_j}{F_s} - 1)n + \phi] = \cos[2\pi(1 - \frac{F_j}{F_s})n - \phi] = \cos[2\pi f_j n - \phi] \quad \text{for } \frac{F_s}{2} < F_j < F_s$$

$$= \cos[2\pi(1 - \frac{F_j}{F_s})n + \phi] = \cos[2\pi f_j n + \phi] \quad \text{for } F_s < F_j < \frac{3F_s}{2}$$

$$= \cos[2\pi(\frac{F_j}{F_s} - 2)n + \phi] = \cos[2\pi(2 - \frac{F_j}{F_s})n - \phi] = \cos[2\pi f_j n - \phi] \quad \text{for } \frac{3F_s}{2} < F_j < 2F_s$$

⋮

For each range of the analog baseband frequency, the corresponding digital frequency f_j is between 0 and .5, i. e., $0 < f_j < .5$. Continuing this development, we obtain the aliasing function $g(F)$ plotted in Figure 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 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 baseband frequency of each signal despite aliasing, the baseband 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 Figure 1, $W = 1$ GHz and $\tau = .5$ ns $= .5 \times 10^{-9}$ s.

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

$$x_i(n) = \sum_{j=1}^J \left\{ \frac{A_j}{2} e^{j\kappa_j \gamma_{j0}} e^{j\kappa_j \gamma_{ji}} e^{j2\pi f_j n} + \frac{A_j}{2} e^{-j\kappa_j \gamma_{j0}} e^{-j\kappa_j \gamma_{ji}} e^{-j2\pi f_j n} \right\} \quad (2)$$

$$y_i(n) = \sum_{j=1}^J \left\{ \frac{A_j}{2} e^{j\kappa_j \gamma_{j0}} e^{j\kappa_j \gamma_{ji}} e^{-j\kappa_j 2\pi F_j \tau} e^{j2\pi f_j n} + \frac{A_j}{2} e^{-j\kappa_j \gamma_{j0}} e^{-j\kappa_j \gamma_{ji}} e^{j\kappa_j 2\pi F_j \tau} e^{-j2\pi f_j n} \right\}$$

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 baseband bandwidth. A_j is the amplitude of the j -th signal while γ_{j0} is the phase of j -th signal at the origin of the antenna array system. γ_{ji} is the relative phase of the j -th signal arrival at the i -th antenna. If the i -th antenna is at the x - y coordinate pair, (x_i, y_i) , and the j -th source is at an azimuth angle of θ_j and an elevation angle of ϕ_j ,

$$\gamma_j(i) = -\frac{2\pi}{\lambda_j} (x_i \cos\theta_j \sin\phi_j + y_i \sin\theta_j \sin\phi_j) \quad i=1, \dots, M \quad (3)$$

where λ_j is the wavelength of the j -th 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, in the prototype system the observation interval is $.5 \mu s = .5 \times 10^{-6} s$ yielding roughly $N = 128$ samples for each of the M antennas.

As indicated in Figure 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 narrowband passband 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 which is less than the number of antennas. Separate peaks may be processed in parallel. Recall that sources well separated in baseband frequency may be aliased to very nearly the same digital frequency due to under-sampling. Thus, several sources may be contributing to a given spectral peak.

The respective N pt. DFT's of the X and Y data sets for the i -th antenna are denoted $X_i(k)$ and $Y_i(k)$, $i=1, \dots, M$, and may be expressed as

$$X_i(k) = \sum_{j=1}^J \left\{ \frac{A_j}{2} e^{j\kappa_j \gamma_{j0}} e^{j\kappa_j \gamma_j(i)} \text{sinc}_{N_j} \left(f_j - \frac{k}{N} \right) + \frac{A_j}{2} e^{-j\kappa_j \gamma_{j0}} e^{-j\kappa_j \gamma_j(i)} \text{sinc}_{N_j} \left(f_j + \frac{k}{N} \right) \right\} \quad (4)$$

$$Y_i(k) = \sum_{j=1}^J \left\{ \frac{A_j}{2} e^{j\kappa_j \gamma_{j0}} e^{j\kappa_j \gamma_j(i)} e^{-j\kappa_j 2\pi F_j \tau} \text{sinc}_{N_j} \left(f_j - \frac{k}{N} \right) + \frac{A_j}{2} e^{-j\kappa_j \gamma_{j0}} e^{-j\kappa_j \gamma_j(i)} e^{j\kappa_j 2\pi F_j \tau} \text{sinc}_{N_j} \left(f_j + \frac{k}{N} \right) \right\}$$

where N_j is the number of samples for which the j -th 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, 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 $k = k_0$ 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 $2M \times L$ vectors:

$$\mathbf{X}_i(k_o) = [X_i(k_o - L'), \dots, X_i(k_o), \dots, X_i(k_o + L')]^T \quad \mathbf{Y}_i(k_o) = [Y_i(k_o - L'), \dots, Y_i(k_o), \dots, Y_i(k_o + L')]^T \quad (5)$$

To give a perspective on the computational complexity, in the simulations presented in Section 6 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, denoted J' .

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

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

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

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

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

Neglecting the negative frequency contributions, $\mathbf{X}_i(k_o) = \sum_{j=1}^{J'} \frac{A_j}{2} e^{j\kappa_j \gamma_o} e^{j\kappa_j \gamma_j(i)} d(f_j)$ 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 $\mathbf{X} = [\mathbf{X}_1(k_o) : \mathbf{X}_2(k_o) : \dots : \mathbf{X}_M(k_o)]$. $\mathbf{X}$ may be expressed in factored form as

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

where $A_j' = A_j/2$ and $\mathbf{a}(\theta_j, \phi_j, \kappa_j) = [e^{j\kappa_j \gamma_j(1)}, e^{j\kappa_j \gamma_j(2)}, \dots, e^{j\kappa_j \gamma_j(M)}]^T$ 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 , 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 $Y = [Y_1(k_o) : Y_2(k_o) : \dots : Y_M(k_o)]$. Neglecting negative frequency components, Y may be expressed as

$$Y = \sum_{j=1}^{J'} A_j e^{j\kappa_j \gamma_o} e^{-j\kappa_j 2\pi F_j \tau} d(f_j) a^T(\theta_j, \phi_j, \kappa_j) \quad (L \times M) \quad (9)$$

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 Figure 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.

3. ESPRIT Based Frequency Estimation With Temporal Undersampling

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

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

The critical observation for estimating F_j is that when $\mu = e^{-j\kappa_j 2\pi F_j \tau}$, the rank of $Y - \mu X$ drops from J' to $J'-1$ since the l -th term drops out of the sum. Thus, $\mu_j = e^{-j\kappa_j 2\pi F_j \tau}$, $j=1, \dots, J'$, are J' distinct generalized eigenvalues of the $L \times M$ rectangular matrix pencil $\{Y, X\}$.

The argument of the ESPRIT eigenvalue, $\arg\{\mu_j\} = -\kappa_j 2\pi F_j \tau$, is plotted as a function of the baseband frequency F_j for $F_s = 250$ MHz in Figure 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" implementation of ESPRIT for estimating the phase factors, $-\kappa_j 2\pi F_j \tau$, $j=1, \dots, J'$. PRO-ESPRIT operates on the $L \times L$ autocorrelation and cross-correlation matrices $R_{xx} = \frac{1}{M} \sum_{i=1}^M X_i(k_o) X_i^H(k_o) = \frac{1}{M} X X^H$ and $R_{yx} = \frac{1}{M} \sum_{i=1}^M Y_i(k_o) X_i^H(k_o) = \frac{1}{M} Y X^H$. Note 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 R_{xx} : $R_{xx} u_i = \lambda_i u_i$, $i = 1, \dots, L$, 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 U_S as

$$\Sigma_S = \text{diag}\{(\lambda_1 - \lambda_{\min})^{1/2}, (\lambda_2 - \lambda_{\min})^{1/2}, \dots, (\lambda_{J'} - \lambda_{\min})^{1/2}\} \quad (11)$$

$$U_S = [u_1 : u_2 : \dots : u_{J'}] \quad (12)$$

The smallest eigenvalue, $\lambda_{\min}$, is asymptotically equal to the noise power which affects the diagonal elements of the autocorrelation matrix R_{xx} . Note 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

$$\Psi = \Sigma_S^{-1} U_S^H R_{yx} U_S \Sigma_S^{-1} \quad (J' \times J') \quad (13)$$

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

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

$$\begin{aligned} Y - \mu X &= U_S U_S^H Y V_S V_S^H - \mu U_S \Sigma_S V_S^H \\ &= U_S \Sigma_S \{\Sigma_S^{-1} U_S^H Y V_S - \mu I_{J'}\} V_S^H \\ &= U_S \Sigma_S \{\Sigma_S^{-1} U_S^H Y V_S \Sigma_S U_S^H U_S \Sigma_S^{-1} - \mu I_{J'}\} V_S^H \\ &= U_S \Sigma_S \{\Sigma_S^{-1} U_S^H Y X^H U_S \Sigma_S^{-1} - \mu I_{J'}\} V_S^H \end{aligned} \quad (14)$$

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

In the prototype system $\tau = .5 \times 10^{-9}$ s such that F_j may be estimated from the phase of the j -th ESPRIT eigenvalue according to $F_j = |\arg\{\mu_j\}|/2\pi$ (2×10^9) Hz, $j=1, \dots, J'$, 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 a 1 GHz. Simulations presented in Section 6 reveal that the variance of the baseband frequency estimates obtained from ESPRIT in this manner are on the order of 10 MHz while the Cramer Rao Lower Bound (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 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: (i) estimate the digital frequency f_j , (ii) convert f_j to the aliased analog frequency via $F_j^a = f_j F_s$, and (iii) translate F_j^a up to the proper aliasing zone using the phase of the ESPRIT eigenvalue μ_j , in conjunction with Figure 4, where F_j^a is either added to or subtracted from an integer multiple of the sampling rate to estimate the actual baseband 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 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, insurmountable when sources are closely-spaced in frequency after aliasing. Fortunately, 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 j -th $M \times 1$ right generalized eigenvector, r_j , of the $L \times M$ rectangular matrix pencil $\{Y, X\}$ is that vector satisfying $\{Y - \mu_j X\} r_j = 0$. Substituting the noiseless (ideal) forms of the X and Y data matrices:

$$\left\{ \sum_{j=1}^J A_j' e^{j\kappa_j \gamma_0} \{e^{-j\kappa_j 2\pi F_j \tau} - \mu\} d(f_j) a^T(\theta_j, \phi_j, \kappa_j) \right\} r_j = 0 \quad (15)$$

When $\mu = e^{-j\kappa_j 2\pi F_j \tau}$, the l -th term $d(f_l) a^T(\theta_l, \phi_l, \kappa_l)$ drops out of sum such that $a^T(\theta_l, \phi_l, \kappa_l) r_j = 0$ for $l = 1, \dots, J$, $l \neq j$. Hence, r_j can be used to extract $d(f_j)$ to within a scalar multiple:

$$X r_j \propto d(f_j) \quad (Y r_j \propto d(f_j))$$

A key point is that the estimate of $d(f_j)$ 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 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 $\hat{d}(f_j)$ to estimate f_j .

Note, we only desire r_j in order to compute $X r_j$ as our estimate of $d(f_j)$ to within a scalar multiple. We can bypass the computation of r_j and construct $X r_j$ directly from the $J \times 1$ right eigenvectors of Ψ ,

defined by (13), satisfying $\Psi\beta_j = \mu_j\beta_j$, $j=1,\dots,J'$. From (14) and (15), it follows that

$$\hat{d}(f_j) \propto Xr_j = U_S \Sigma_S \beta_j \quad j=1,\dots,J' \quad (16)$$

where Σ_S and U_S are constructed from the J' largest eigenvalues and corresponding eigenvectors of R_{xx} according to (11) and (12), respectively.

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

$$\hat{f}_j = \frac{1}{2\pi} \arg \left\{ \hat{d}^H(f_j) \Delta^* P_{\dagger}^{\dagger} \hat{d}(f_j) \right\} \quad j=1,\dots,J' \quad (17)$$

where $\hat{d}(f_j)$ is computed as in (16) and $P_{\dagger}^{\dagger}$ and Δ are each $L \times L$ matrices defined as

$$P_{\dagger}^{\dagger} = I_L - \frac{1}{L} \mathbf{1} \mathbf{1}^T \quad \Delta = \text{diag} \left\{ e^{-j2\pi \frac{k_0 - L'}{N}}, \dots, e^{-j2\pi \frac{k_0}{N}}, \dots, e^{-j2\pi \frac{k_0 + L'}{N}} \right\} \quad (18)$$

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 $\hat{d}(f_j) = d(f_j)$.

PROOF: First, we need to define some quantities. Let 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, $k_0 - L', \dots, k_0, \dots, k_0 + L'$. Let W_1 and W_2 be composed of the first and last $N-1$ columns of W , respectively. W_1 and W_2 are each $L \times (N-1)$ and related as $W_2 = \Delta W_1$. Finally, let w_N denote the last column of W ; $w_N = \text{diag}(\Delta^{N-1})$, where $\text{diag}(\cdot)$ converts the $L \times L$ diagonal matrix Δ^{N-1} to an $L \times 1$ column vector. Note, the first column of W is $\mathbf{1}$ such that $P_{\dagger}^{\dagger} W = [0_L : P_{\dagger}^{\dagger} W_2] = [0_L : P_{\dagger}^{\dagger} \Delta W_1]$, where 0_L is an $L \times 1$ vector composed of all zeroes.

Next, define $v(f)$ as the $N \times 1$ Vandermonde vector $v(f) = [1, e^{j2\pi f}, e^{j4\pi f}, \dots, e^{j(N-1)2\pi f}]^T$. Let $v_1(f)$ and $v_2(f)$ be composed of the first and last $N-1$ elements of $v(f)$, respectively. $v_1(f)$ and $v_2(f)$ are each $(N-1) \times 1$ and related as $v_2(f) = e^{j2\pi f} v_1(f)$. With these definitions and relationships, it follows that

$$\begin{aligned} d^H(f_j) \Delta^* P_{\dagger}^{\dagger} d(f_j) &= v^H(f_j) W^H \Delta^* P_{\dagger}^{\dagger} W v(f_j) \\ &= v^H(f_j) W^H \Delta^* P_{\dagger}^{\dagger} \Delta W_1 v_1(f_j) e^{j2\pi f_j} \\ &= \{v_1^H(f_j) W_1^H + e^{-j(N-1)2\pi f_j} w_N^H\} \Delta^* P_{\dagger}^{\dagger} \Delta W_1 v_1(f_j) e^{j2\pi f_j} \end{aligned}$$

$$\begin{aligned}
&= \mathbf{v}_1^H(f_j) \mathbf{W}_1^H \Delta^* \mathbf{P}_1^\dagger \Delta \mathbf{W}_1 \mathbf{v}_1(f_j) e^{j2\pi f_j} + \mathbf{1}^T \mathbf{P}_1^\dagger \Delta \mathbf{W}_1 \mathbf{v}_1(f_j) e^{-j(N-2)2\pi f_j} \\
&= \{\mathbf{v}_1^H(f_j) \mathbf{W}_1^H \Delta^* \mathbf{P}_1^\dagger \Delta \mathbf{W}_1 \mathbf{v}_1(f_j)\} e^{j2\pi f_j}
\end{aligned}$$

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

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

Comparing the Direct and Indirect ESPRIT methods, in the former the phase of the j -th eigenvalue of Ψ is multiplied by $\pi \times 10^9$ while in the latter the phase of $\hat{\mathbf{d}}^H(f_j) \Delta^* \mathbf{P}_1^\dagger \hat{\mathbf{d}}(f_j)$ 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 Cramer Rao Lower Bound (CRB) than the Direct ESPRIT method.

The formula for translating F_j^a up to the proper aliasing zone is dictated by Figure 4 wherein the phase of the ESPRIT eigenvalue $\mu = e^{-jk2\pi F\tau}$, where $\tau = .5 \times 10^{-9}$ 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 Figure 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 Figure 3. The baseband frequency of the j -th source, F_j , is ultimately determined from the aliased frequency estimate, $\hat{F}_j^a$, according to

$$\hat{F}_j = \left| \hat{F}_j^a - 250 \times 10^6 \text{ round} \left\{ \frac{\arg\{\mu_j\} + \pi/16}{\pi/4} \right\} \right| \text{ Hz} \quad \text{for} \quad -15\pi/16 < \arg\{\mu_j\} < \pi \quad (19)$$

$$\hat{F}_j = 1 \times 10^9 - \hat{F}_j^a \text{ Hz} \quad \text{for} \quad -\pi < \arg\{\mu_j\} < -15\pi/16$$

where $\text{round}[x]$ 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 6 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 , while in the other it is subtracted from $(n+1)F_s$.

4. 2-D Angle Estimation With Spatial Undersampling Via PRO-ESPRIT and Integer Search

4.1 Estimation of the Array Manifold for Each Source

In Section 3, we saw that use of the right generalized eigenvectors of the $L \times M$ matrix pencil $\{Y, X\}$ 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 $\{Y, X\}$ 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 j -th left generalized eigenvector of $\{Y, X\}$ is used to extract from the X and Y data an estimate of the array manifold for the j -th source, denoted $a(\theta_j, \phi_j, \kappa_j)$. 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 baseband frequency is located on a downward sloping portion of the aliasing function.

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

$$l_j^H \left\{ \sum_{l=1}^J A_l' e^{j\kappa_l \tau} \{e^{-j\kappa_l 2\pi F_l \tau} - \mu\} d(f_l) a^T(\theta_l, \phi_l, \kappa_l) \right\} = 0 \quad (20)$$

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

$$l_j^H X \propto a^T(\theta_j, \phi_j, \kappa_j) \Rightarrow X^T l_j^* \propto a(\theta_j, \phi_j, \kappa_j) \quad l_j^H Y \propto a^T(\theta_j, \phi_j, \kappa_j) \Rightarrow Y^T l_j^* \propto a(\theta_j, \phi_j, \kappa_j) \quad (21)$$

Thus, applying the j -th left generalized eigenvector allows us to extract an estimate of the array manifold for the j -th source which, in turn, may be operated upon to estimate the azimuth and elevation angles of the j -th 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 l_j is associated with the ESPRIT eigenvalue $\mu_j = e^{-j\kappa_j 2\pi F_j \tau}$, the azimuth and elevation angle estimates obtained by processing the estimate of $a(\theta_j, \phi_j, \kappa_j)$ are automatically paired with the estimate of F_j obtained via the algorithm developed in Section 3. 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 j -th $L \times 1$ left generalized eigenvector, l_j , of the $L \times M$ rectangular matrix pencil $\{Y, X\}$ may be efficiently computed from the $J' \times 1$ j -th left eigenvector, α_j , of Ψ in (13) satisfying $\alpha_j^H \Psi = \mu_j \alpha_j^H$, $j=1, \dots, J'$. From (14) and (20), it follows that

$$l_j = U_S \Sigma_S^{-1} \alpha_j \quad j=1, \dots, J' \quad (22)$$

Recall 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. Fortunately, the isolation of single source components via PRO-ESPRIT 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've 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 1, 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 two-fold: (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. We here 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 6 is depicted in Figure 2. The corresponding array manifold is

$$a(\theta_j, \phi_j, \kappa_j) = \quad (23)$$

$$\left[e^{j \kappa_j 2\pi \frac{2d_2}{\lambda_j} v_j}, e^{j \kappa_j 2\pi \frac{2d_2-d_1}{\lambda_j} v_j}, e^{j \kappa_j 2\pi \frac{d_2}{\lambda_j} v_j}, e^{j \kappa_j 2\pi \frac{d_1}{\lambda_j} v_j}, 1, e^{j \kappa_j 2\pi \frac{d_1}{\lambda_j} u_j}, e^{j \kappa_j 2\pi \frac{d_2}{\lambda_j} u_j}, e^{j \kappa_j 2\pi \frac{2d_2-d_1}{\lambda_j} u_j}, e^{j \kappa_j 2\pi \frac{2d_2}{\lambda_j} u_j} \right]^T$$

where u_j and v_j are the direction cosines of the j -th source relative to x -axis and y -axis, respectively, and λ_j

$$v_{\tilde{l}} = v_{j_0} + \tilde{l} \frac{\lambda_j}{d_2} \quad \tilde{l} \in \left\{ \text{ceiling} \left[-\frac{d_2}{\lambda_j} (1+v_{j_0}) \right], \text{floor} \left[\frac{d_2}{\lambda_j} (1-v_{j_0}) \right] \right\} \quad (28)$$

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 $v_{\tilde{l}}$

$$v_{j_0} + \tilde{k} \frac{\lambda_j}{d_1} = v_{j_0} + \tilde{l} \frac{\lambda_j}{d_2} \quad \rightarrow \quad \frac{d_2}{d_1} = \frac{\tilde{l}}{\tilde{k}} \quad (29)$$

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 prime, i. e., have no common factors other than unity. The set of ambiguous angles is then $v_n = v_{j_0} + n k' \frac{\lambda_j}{d_1} = v_{j_0} + n l' \frac{\lambda_j}{d_2}$ for any n for which $-1 < v_n < 1$. Consider the case of $n=1$. If we make sure that either $v_{j_0} + k' \frac{\lambda_j}{d_1}$ or $v_{j_0} + l' \frac{\lambda_j}{d_2}$ 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*

$$k' \notin \left\{ \text{ceiling} \left[-\frac{d_1}{\lambda_j} (1+v_{j_0}) \right], \text{floor} \left[\frac{d_1}{\lambda_j} (1-v_{j_0}) \right] \right\} \quad \text{or} \quad l' \notin \left\{ \text{ceiling} \left[-\frac{d_2}{\lambda_j} (1+v_{j_0}) \right], \text{floor} \left[\frac{d_2}{\lambda_j} (1-v_{j_0}) \right] \right\}$$

These conditions depend on the direction cosine of the source. To remove the data dependence, we over-specify and let $v_{j_0} = 1$ for the lower bound limit and $v_{j_0} = -1$ 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.

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

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 true source direction cosine, v_{j_0} .

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 $l' \notin \{-14, 14\}$. 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.

4.3 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 source direction cosine. We now develop an algorithm to do such. With respect to Figure 2, element pairs 1-2 and 4-5 provide two measurements of $\psi_1 = \arg\{e^{j 2\pi \frac{d_1}{\lambda_j} v_{jo}}\}$. The candidate estimates of v_{jo} in the "visible" region $-1 \leq v \leq 1$ are

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

Let k^* be that for which $v_{jo} = \frac{\lambda_j}{2\pi d_1} \psi_1 + k^* \frac{\lambda_j}{d_1}$. 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, 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 overriding goal of keeping the computational load as small as possible.

Element pairs 1-3 and 3-5 provide two measurements of $\psi_2 = \arg\{e^{j 2\pi \frac{d_2}{\lambda_j} v_{jo}}\}$. The candidate estimates of v_{jo} in the "visible" region are

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

Let l^* denote that value of l such that $v_l^{(2)} = v_{jo}$. 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^* then is as follows. For each integer k in $\{\text{ceiling} \left[\frac{d_1}{\lambda_j} \frac{\psi_1}{2\pi} \right], \text{floor} \left[\frac{d_1}{\lambda_j} \frac{\psi_1}{2\pi} \right]\}$, compute the corresponding l according to (33). Select k^* as that for which $|l - \text{round}[l]|$ is minimum, where $\text{round}[l]$ 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.

So far we've 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 $\psi_3 = \arg\{e^{j 2\pi \frac{(d_2-d_1)}{\lambda_j} v_{j0}}\}$. 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 Figure 2, $\frac{d_2-d_1}{d_1} = \frac{3.0}{2.3} = \frac{30}{23} = \frac{m'}{k'}$. (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 $\psi_4 = \arg\{e^{j 2\pi \frac{(2d_2-d_1)}{\lambda_j} v_{j0}}\}$. 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 solving for n yields

$$n = k \frac{2d_2-d_1}{d_1} + \frac{2d_2-d_1}{d_1} \frac{\psi_1}{2\pi} - \frac{\psi_4}{2\pi} \quad (35)$$

Relative to the prototype array in Figure 2, $\frac{2d_2-d_1}{d_1} = \frac{8.3}{2.3} = \frac{83}{23} = \frac{n'}{k'}$. (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 $\{\text{ceiling} \left[\frac{d_1}{\lambda_j} \frac{\psi_1}{2\pi} \right], \text{floor} \left[\frac{d_1}{\lambda_j} \frac{\psi_1}{2\pi} \right]\}$, compute the corresponding values of l , m , and n , according to (33),

(34), and (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), (34), and (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 Figure 5. The computational simplicity is evident. Note, 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.

5. Simulation Examples

The performance of the frequency and 2-D angle estimation algorithm summarized in the flowchart in Figure 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 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 Figure 2. Note, at 18 GHz the wavelength is roughly 2/3 in. such that the smallest interelement spacing in the L-array, $d_1 = 2.3$ in, is roughly 7 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 Figures 1 and 2 were common to all simulation runs. In all cases, the signal scenario was composed of equi-powered RF pulsed signals (monochromatic planewaves) that were "turned on" during the entire .5 μ s 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 Equations (36) and (37) of Appendix A, prior to computing the 128 pt. DFT. Finally, the Cramer Rao Lower Bound 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, while 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 Figure 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 4x4 EVD followed by a 2x2 EVD. Sample statistics computed from 250 independent runs for each of a number of different SNR's are plotted in Figures 6, 7, 9, and 10.

Figures 6 and 9 reveal the high variance of the Direct ESPRIT frequency estimates, three orders of magnitude greater than the CRB, in accordance with the discussion in Section 3. 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 which are in the band 2-18 GHz.

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 Figures 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 a thousandth of a degree. It may be very difficult to achieve this kind of accuracy in practice even if it is achieved in simulation. Note, 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 0.1° in all cases, even at 0 dB SNR, while the sample biases obtained in the 17-18 GHz range were less 0.01° in all cases, even at 0 dB SNR.

Again, for signals in the 17-18 GHz band, the smallest interelement spacing in the L-array employed is roughly 7 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, 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.

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 Figure 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 5x5 EVD followed by a 4x4 EVD. Sample statistics computed from 250 independent runs for each of a number of different SNR's are plotted in Figures 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 Figure 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 remarkable performance.

6. Final Comments

The frequency and 2-D angle estimation algorithm developed within and summarized in Figure 5 is not able to handle sources that are aliased to exactly the same frequency. Examining Figure 3, this will occur if (i) two sources are separated in frequency by nF_s , or (ii) one source is at $nF_s - \Delta F$ while another source 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.

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. Stavrinos, "Sensor Array Signal Processing via a Procrustes 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.

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 antennas. By raw data, we mean that prior to any processing including the FFT (or DFT). Let $x(n)$ denote the $M \times 1$ vector the i -th component of which is the raw data output from the reference channel of the i -th antenna, $i=1, \dots, M$, at the n -th sampling instant, $n=0, 1, \dots, N-1$. Let $y(n)$ be

defined similarly relative to the time-delayed channel at each antenna. From the initial development in Section 2, it follows that $\mathbf{x}(n)$ and $\mathbf{y}(n)$ may be expressed as

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

$$\mathbf{y}(n) = \text{Re}\{\mathbf{A}\mathbf{\Omega}^n \mathbf{\Phi} \mathbf{c}\} + \mathbf{n}_y(n) \quad n = 0, 1, \dots, N-1 \quad (37)$$

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

$$\mathbf{A} = [\mathbf{a}(\theta_1, \phi_1) : \mathbf{a}(\theta_2, \phi_2) : \dots : \mathbf{a}(\theta_J, \phi_J)] \quad (38)$$

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

$$\mathbf{c} = [c_1, c_2, \dots, c_J]^T = \bar{\mathbf{c}} + j\tilde{\mathbf{c}} \quad (39)$$

where $c_j = A_j e^{j\gamma_{j0}}$ is the complex amplitude of the j -th source at time $n = 0$ at the reference element. $\mathbf{\Omega}$ is the $J \times J$ diagonal matrix

$$\mathbf{\Omega} = \text{diag}\{e^{j\omega_1 T}, e^{j\omega_2 T}, \dots, e^{j\omega_J T}\} \quad (40)$$

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 . $\mathbf{\Phi}$ is the $J \times J$ diagonal matrix

$$\mathbf{\Phi} = \text{diag}\{e^{j\omega_1 \tau}, e^{j\omega_2 \tau}, \dots, e^{j\omega_J \tau}\} \quad (41)$$

where τ is the time delay equal to $.5 \text{ ns} = .5 \times 10^{-9} \text{ s}$ in our prototype system. $\mathbf{n}_x(n)$ and $\mathbf{n}_y(n)$, $n=0, 1, \dots, N-1$, 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 $\mathbf{x}(n) \sim \mathcal{N}(\text{Re}\{\mathbf{A}\mathbf{\Omega}^n \mathbf{c}\}, \sigma_n^2 \mathbf{I}_M)$ and $\mathbf{y}(n) \sim \mathcal{N}(\text{Re}\{\mathbf{A}\mathbf{\Omega}^n \mathbf{\Phi} \mathbf{c}\}, \sigma_n^2 \mathbf{I}_M)$. The log-likelihood function is

$$\ln L(\omega, \theta, \phi, \bar{\mathbf{c}}, \tilde{\mathbf{c}}, \sigma_n^2) = \text{constant} - NM \ln \sigma_n^2 \quad (42)$$

$$-\frac{1}{2\sigma_n^2} \sum_{n=0}^{N-1} \|\mathbf{x}(n) - \text{Re}\{\mathbf{A}\mathbf{\Omega}^n \mathbf{c}\}\|^2 - \frac{1}{2\sigma_n^2} \sum_{n=0}^{N-1} \|\mathbf{y}(n) - \text{Re}\{\mathbf{A}\mathbf{\Omega}^n \mathbf{\Phi} \mathbf{c}\}\|^2$$

Let α denote the set of parameters that the log-likelihood function depends on. α contains $5J+1$ parameters which we group as follows: $\omega = [\omega_1, \omega_2, \dots, \omega_J]^T$, $\theta = [\theta_1, \theta_2, \dots, \theta_J]^T$, $\phi = [\phi_1, \phi_2, \dots, \phi_J]^T$, $\bar{\mathbf{c}} = [A_1 \cos \gamma_{10}, \dots, A_J \cos \gamma_{J0}]$, $\tilde{\mathbf{c}} = [A_1 \sin \gamma_{10}, \dots, A_J \sin \gamma_{J0}]$, 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 i -th parameter α_i is $[\mathbf{J}^{-1}]_{ii}$, i. e., the i -th diagonal element of the inverse of $\mathbf{J}$. Taking into account symmetry, the Fisher Information Matrix may be built up from the the

(1,1) element $E \left[\left[\frac{\partial}{\partial \sigma_n^2} (\ln L) \right]^2 \right] = \frac{MN}{\sigma_n^4}$, the five $1 \times J$ blocks, $E \left[\frac{\partial}{\partial \sigma_n^2} (\ln L) \nabla_{\epsilon}^T (\ln L) \right]$, $E \left[\frac{\partial}{\partial \sigma_n^2} (\ln L) \nabla_{\theta}^T (\ln L) \right]$, $E \left[\frac{\partial}{\partial \sigma_n^2} (\ln L) \nabla_{\phi}^T (\ln L) \right]$, and $E \left[\frac{\partial}{\partial \sigma_n^2} (\ln L) \nabla_{\omega}^T (\ln L) \right]$, all of which are equal to 0, and the $J \times J$ blocks $E[\nabla_{\epsilon} \nabla_{\epsilon}^T]$, $E[\nabla_{\epsilon} \nabla_{\theta}^T]$, $E[\nabla_{\epsilon} \nabla_{\phi}^T]$, $E[\nabla_{\epsilon} \nabla_{\omega}^T]$, $E[\nabla_{\theta} \nabla_{\epsilon}^T]$, $E[\nabla_{\theta} \nabla_{\theta}^T]$, $E[\nabla_{\theta} \nabla_{\phi}^T]$, $E[\nabla_{\theta} \nabla_{\omega}^T]$, $E[\nabla_{\phi} \nabla_{\epsilon}^T]$, $E[\nabla_{\phi} \nabla_{\theta}^T]$, $E[\nabla_{\phi} \nabla_{\phi}^T]$, and $E[\nabla_{\phi} \nabla_{\omega}^T]$, where it is understood that the function that the gradient is operating upon 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 fifteen $J \times J$ blocks. As an example, though,

$$E[\nabla_{\theta} \nabla_{\phi}^T] = \frac{1}{\sigma_n^2} \sum_{n=0}^{N-1} \text{Re}\{\Omega^{*n} C^* A_{\theta}^H\} \text{Re}\{A_{\phi} \Omega^n C\} + \frac{1}{\sigma_n^2} \sum_{n=0}^{N-1} \text{Re}\{\Omega^{*n} C^* \Phi^* A_{\theta}^H\} \text{Re}\{A_{\phi} \Phi \Omega^n C\} \quad (43)$$

where C , A_{θ} , and A_{ϕ} are defined below.

$$C = \text{diag}\{c_1, c_2, \dots, c_J\} \quad (44)$$

$$A_{\theta} = \left[\frac{\partial}{\partial \theta} a(\theta, \phi) \Big|_{(\theta, \phi) = (\theta_j, \phi_j)}, j=1, \dots, J \right] \quad A_{\phi} = \left[\frac{\partial}{\partial \phi} a(\theta, \phi) \Big|_{(\theta, \phi) = (\theta_j, \phi_j)}, j=1, \dots, J \right] \quad (45)$$

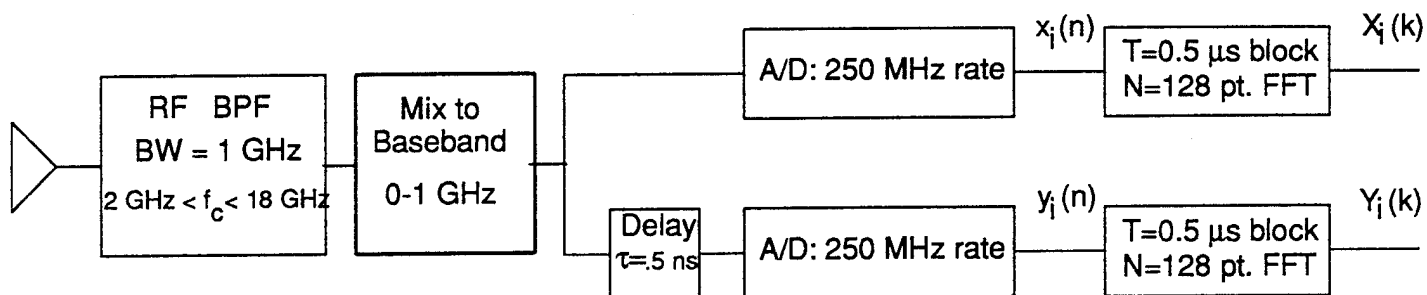


Figure 1: Receiver module and front end signal processing for i -th antenna in prototype system.

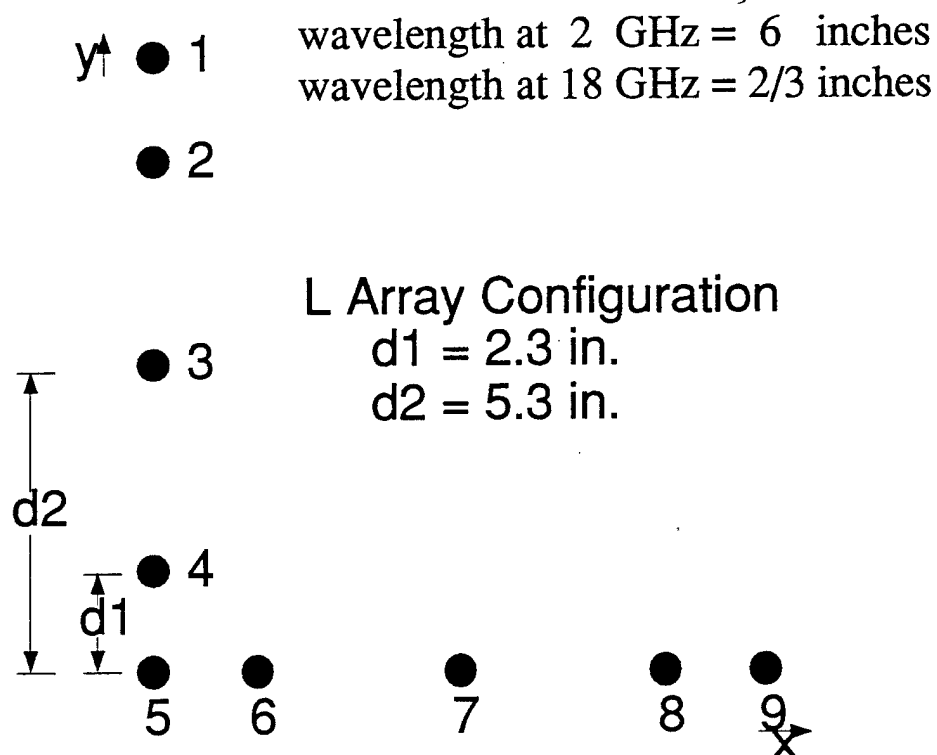


Figure 2: L-shaped antenna array employed in simulations for azimuth/elevation angle estimation over 2-18 GHz spectrum. Each leg is symmetric about its center.

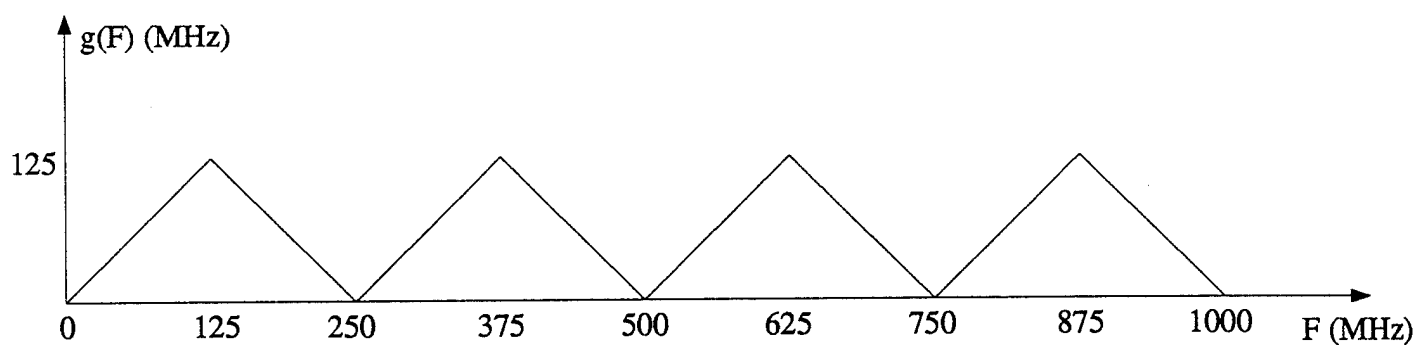


Figure 3: Aliasing function: aliased frequency as a function of baseband frequency after sampling at 250 MHz with real processing (no I and Q).

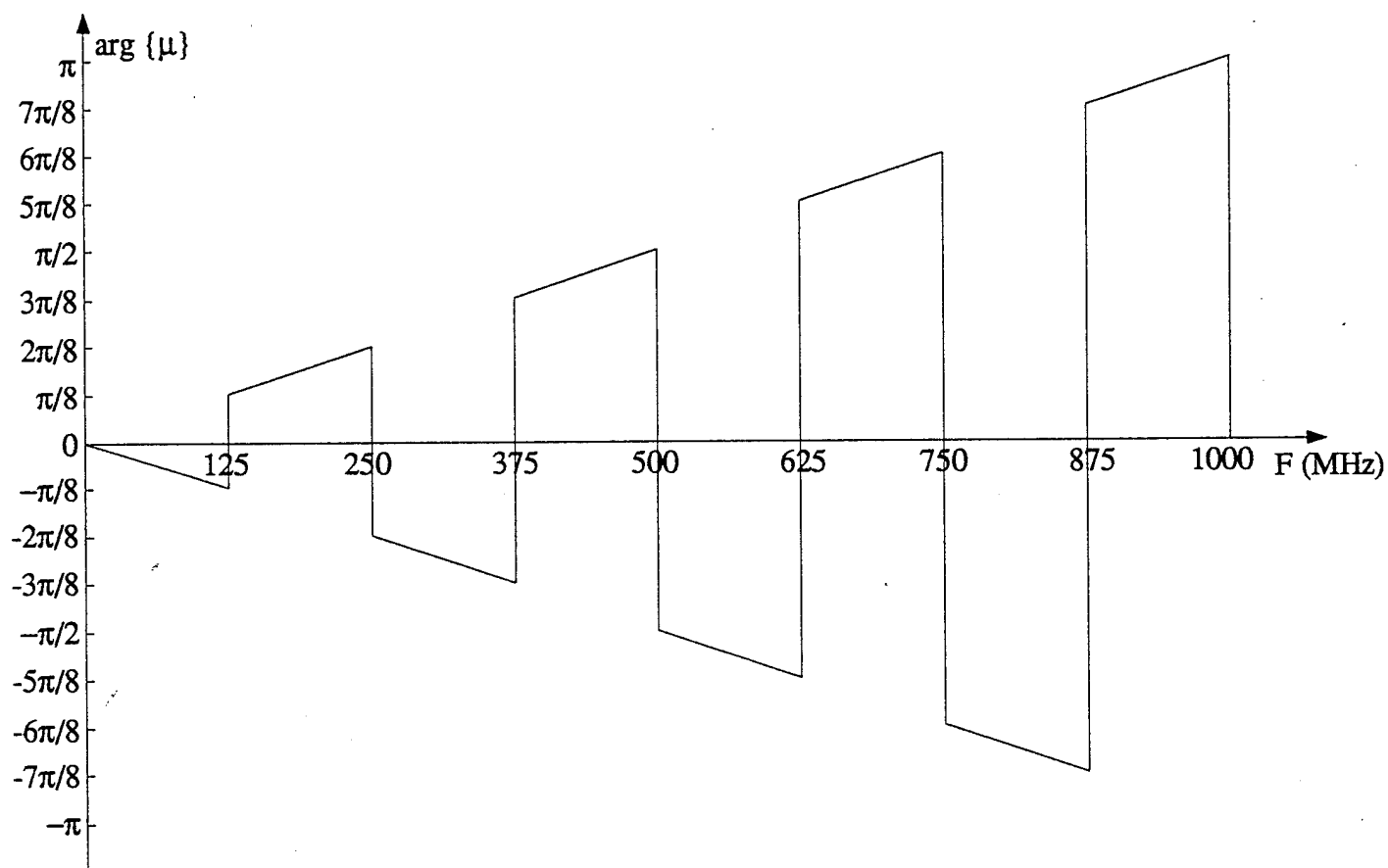


Figure 4: Phase of ESPRIT eigenvalue as a function of analog baseband frequency with 250 MHz sampling rate.

With $1 < L < 6$ DFT values centered at a DFT spectral peak in both X and Y data at i -th antenna, $i=1, \dots, M$, construct $L \times 1$ DFT vectors ($L' = \text{floor}[(L-1)/2]$):

$$\mathbf{X}_{i_0}(k_0) = [X_i(k_0 - L'), \dots, X_i(k_0), \dots, X_i(k_0 + L')]^T \quad \mathbf{Y}_{i_0}(k_0) = [Y_i(k_0 - L'), \dots, Y_i(k_0), \dots, Y_i(k_0 + L')]^T$$

Form $L \times L$ covariance and cross-covariance matrices:

$$\mathbf{R}_{xx} = \sum_{i=1}^M \mathbf{X}_{i_0}(k_0) \mathbf{X}_{i_0}^H(k_0) / M \quad \mathbf{R}_{yx} = \sum_{i=1}^M \mathbf{Y}_{i_0}(k_0) \mathbf{X}_{i_0}^H(k_0) / M$$

Compute EVD of $\mathbf{R}_{xx}$. Determine no. of sources, J , ($1 < J < L$) contributing to spectral peak at k_0 by applying statistical test (e.g., AIC) to eigenvalues.

$$\Sigma_S = \text{diag} \{(\lambda_1 - \lambda_{\min})^{1/2}, \dots, (\lambda_J - \lambda_{\min})^{1/2}\} (J \times J); \quad \mathbf{U}_S = [\mathbf{u}_1, \dots, \mathbf{u}_J] (L \times J)$$

eigenvalues, μ_j
left eigenvectors, α_j

$$\text{Compute EVD of } \Psi = \Sigma_S^{-1} \mathbf{U}_S^H \mathbf{R}_{yx} \mathbf{U}_S \Sigma_S^{-1} (J \times J)$$

eigenvalues, μ_j
right eigenvectors, β_j

for each source, $j=1, \dots, J$, estimate analog baseband frequency, $\hat{F}_j$, and direction cosine, $\hat{v}_j$:

for each interelement spacing, d_i , $i=1, \dots, I$, represented in leg, estimate corr. phase differential:
 $\varphi_i = \arg \{ \alpha_j^H \Sigma_S^{-1} \mathbf{U}_S^H \mathbf{X}_m(k_0) \mathbf{X}_n^H(k_0) \mathbf{U}_S \Sigma_S^{-1} \alpha_j \} / 2\pi$
 m -th and n -th antennas are separated by d_i
 d_1 is smallest interelement spacing in leg

determine n_1^* as that integer in range
 $n_1 \in \{ \text{ceiling}[-d_1/\hat{\lambda}_j - \varphi_1], \text{floor}[d_1/\hat{\lambda}_j - \varphi_1] \}$
for which $\sum_{i=2}^I |n_i - \text{round}[n_i]|$ is minimum
where: $n_i = (d_i/d_1)(n_1 + \varphi_1) - \varphi_i$ $i=2, \dots, I$

$$\kappa_j = -\text{sign} \{ \arg \{ \mu_j \} \}$$

$$\hat{v}_j = \kappa_j \frac{1}{I} \sum_{i=1}^I \hat{v}_j^{(i)}; \quad \hat{v}_j^{(i)} = \frac{\hat{\lambda}_j}{d_i} \varphi_i + n_i^* \frac{\hat{\lambda}_j}{d_i}$$

$\hat{F}_j$ and $\hat{v}_j$ automatically paired

aliased freq. estimate ($0 < F_j^a < 125$ MHz):

$$\hat{F}_j^a = \frac{1}{2\pi} \arg \{ \beta_j^H \Sigma_S \mathbf{U}_S^H \Delta^* \mathbf{P}_1^{\perp} \mathbf{U}_S \Sigma_S \beta_j \} \cdot F_S$$

baseband freq. estimate ($0 < F_j < 1$ GHz):

$$\hat{F}_j = \left| \hat{F}_j^a - F_S \cdot \text{round} \left[\frac{\arg \{ \mu_j \} + \pi/16}{\pi/4} \right] \right|$$

estimate of wavelength:

$$\hat{\lambda}_j = c / (F_{\text{mix}}^{\text{RF}} + \hat{F}_j) \quad c: \text{speed of light}$$

mixer frequency: $2 < F_{\text{mix}}^{\text{RF}} < 17$ GHz

defined quantities (computed a-priori):

$$F_S = 250 \text{ MHz} \quad w_N = \exp[-j2\pi/N]$$

$$\Delta = \text{diag} \left\{ w_N^{k_0 - L'}, \dots, w_N^{k_0}, \dots, w_N^{k_0 + L'} \right\}$$

$$\mathbf{P}_1^{\perp} = \mathbf{I} - \frac{1}{L} \mathbf{1} \mathbf{1}^T \quad (\mathbf{1}: L \times 1 \text{ composed of ones})$$

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

Table I. Signal Parameters for Simulation Example 1.

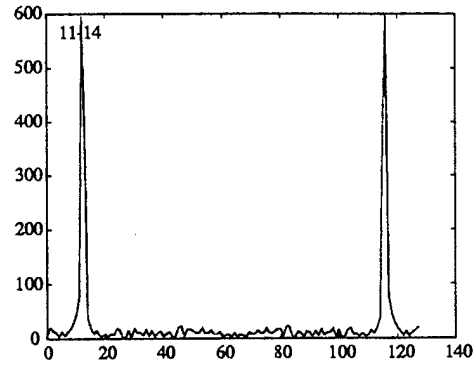


Figure 8: Sample DFT spectrum of X data for Ex. 1.

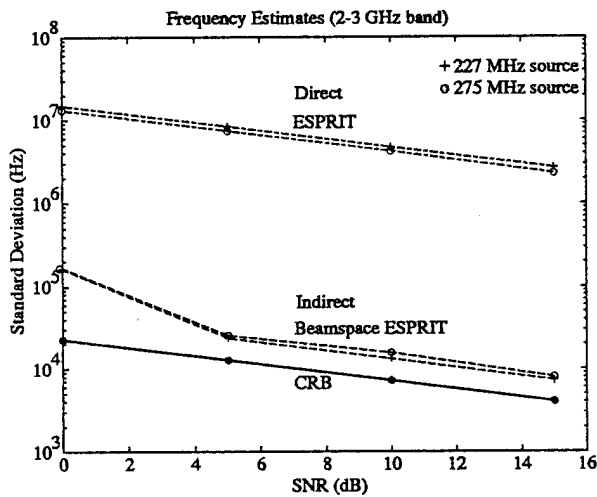


Figure 6: Frequency estimation performance for Example 1 with signals in 2-3 GHz band.

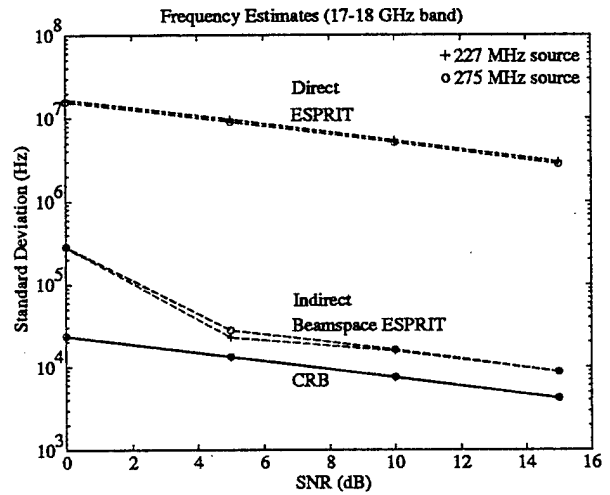


Figure 9: Frequency estimation performance for Example 1 with signals in 17-18 GHz band.

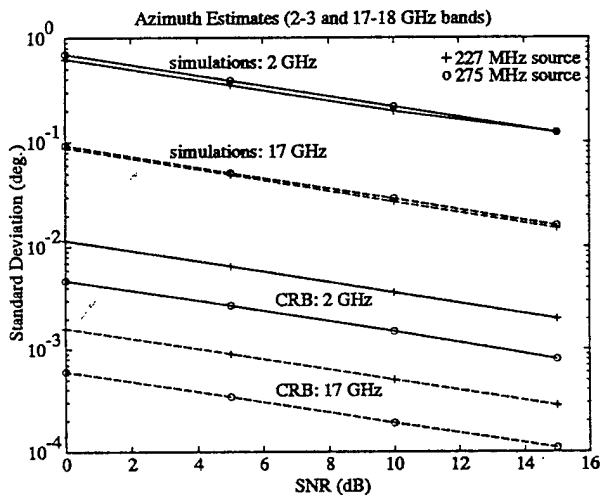


Figure 7: Azimuth estimation performance for Example 1.

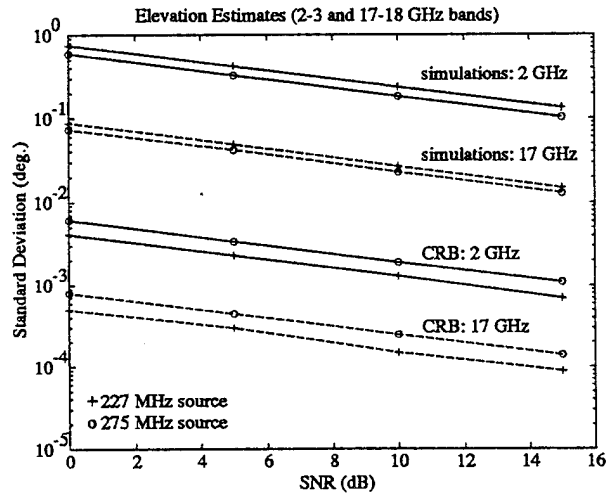


Figure 10: Elevation estimation performance for Ex. 1.

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).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

Table II. Signal Parameters for Simulation Example 2.

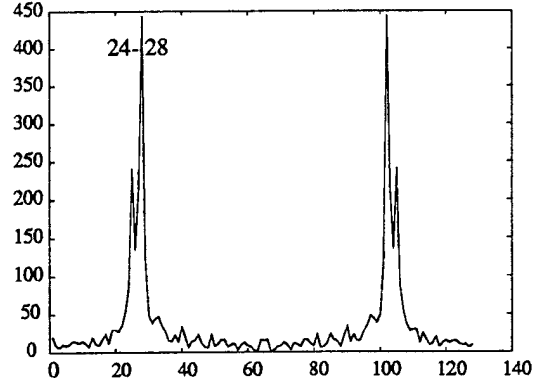


Figure 13: Sample DFT spectrum of X data for Ex. 2.

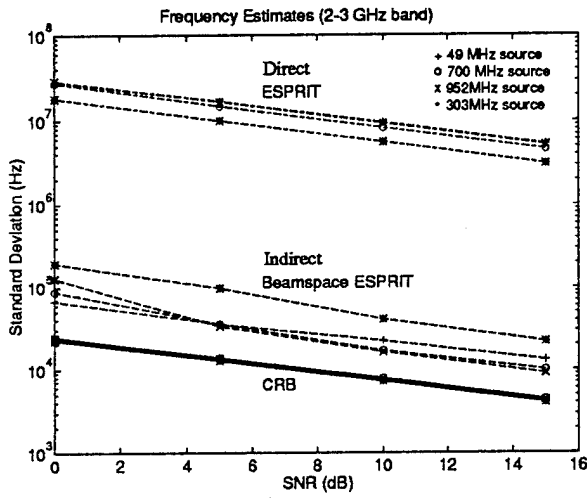


Figure 11: Frequency estimation performance for Example 2 with signals in 2-3 GHz band.

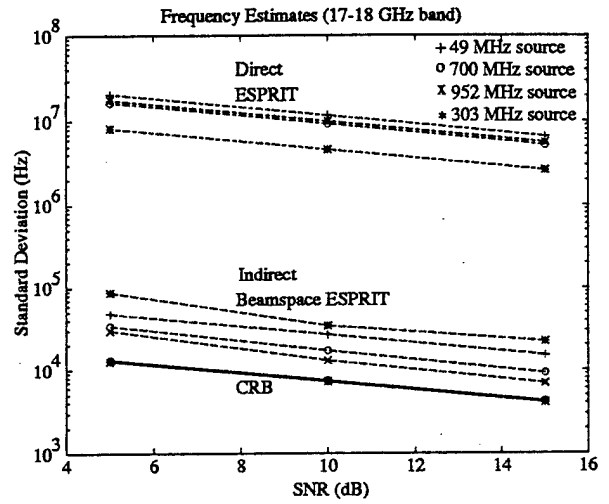


Figure 14: Frequency estimation performance for Example 2 with signals in 17-18 GHz band.

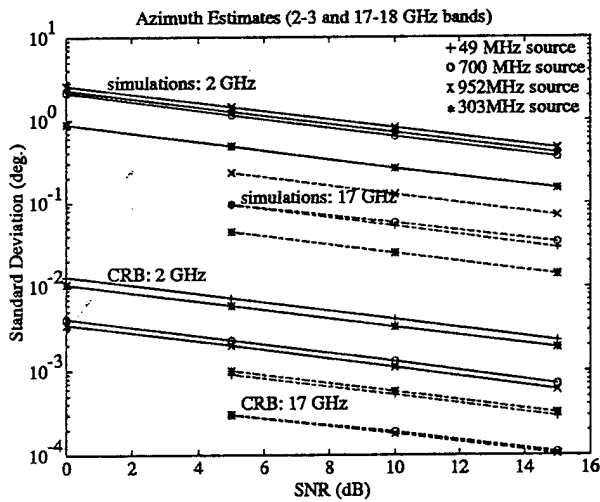


Figure 12: Azimuth estimation performance for Example 2.

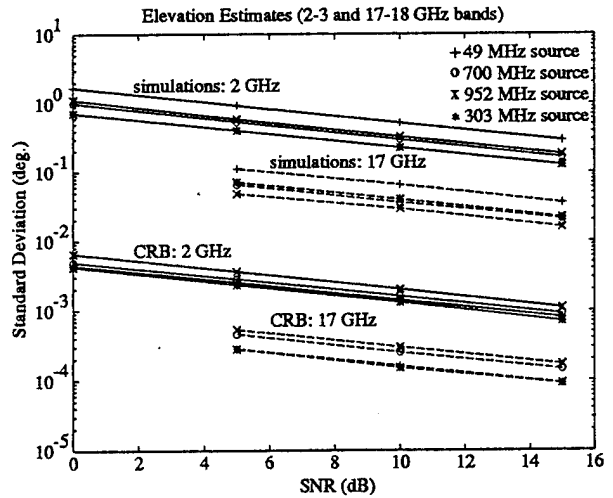


Figure 15: Elevation estimation performance for Ex. 2.