1992

Performance analysis of DOA estimation algorithms using physical parameters

Hui Liu
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the [Electrical and Computer Engineering Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Liu, Hui, "Performance analysis of DOA estimation algorithms using physical parameters" (1992).
Dissertations and Theses. Paper 4362.
<https://doi.org/10.15760/etd.6218>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

AN ABSTRACT OF THE THESIS OF Hui Liu for the Master of Science in
Electrical Engineering presented September 30, 1992.

Title: Performance Analysis of DOA Estimation Algorithms Using Physical
Parameters.

APPROVED BY THE MEMBERS OF THE THESIS COMMITTEE:

Fu Li, Chair

George G. Lendaris

Andrew M. Fraser

Analytical performance analysis on Direction-Of-Arrival (DOA) estimation algorithms has attracted much excellent research in recent years, various statistical properties have been revealed. However, in most of these analyses, insights of the performance were masked because of the involvement of singular values and singular vectors which depend on the character of the algorithms and data structures in a complex and nonlinear manner.

In this thesis, we present a performance analysis of several subspace-based DOA estimation algorithms. Compact formulas of Mean-Squared DOA estimation Error (MSE) for different algorithms are first developed. The analysis is then extended to beamspace, forward-backward averaging and spatial smoothing schemes.

All the statistics of these algorithms are expressed in terms of *physical* parameters which reveal the relations between array structures and performance of DOA estimation.

PERFORMANCE ANALYSIS OF DOA ESTIMATION
ALGORITHMS USING PHYSICAL PARAMETERS

by
HUI LIU

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
ELECTRICAL ENGINEERING

Portland State University
1992

TO THE OFFICE OF GRADUATE STUDIES:

The members of the committee approve the thesis of Hui Liu presented
September 30, 1992.

Fu Li, Chair

George G. Lendaris

Andrew M. Fraser

APPROVED:

Rolf Schaumann, Chairman, Department of Electrical Engineering

Roy W. Koch, Vice Provost for Graduate Studies and Research

ACKNOWLEDGEMENT

I wish to express my deepest thanks and gratitude to my advisor, Dr. Fu Li, for his invaluable guidance, encouragement and unrelenting patience during the course of this research.

My sincere appreciation goes to the other members of the committee, Dr. Andrew M. Fraser and Dr. George G. Lendaris for their helpful comments and understanding. I wish to thank Dr. R. J. Vaccaro at University of Rhode Island for his contribution to our work in Chapter III. Thanks are also due to other members of the faculty and staff in Portland State University who have contributed greatly to my graduate education.

Finally, my parents, deserve my special thanks for their continued support, encouragement and sacrifice throughout, and I will forever be indebted to them for all that they have done.

PREFACE

This work started in early 1991 when I first read Dr. Fu Li's dissertation on unified performance analysis of subspace DOA estimation algorithms. My first goal was to analyze the impact of signal correlation on the performance of the DOA estimation. After observing that an important parameter $||\beta_k||^2$ (Chapter 3, Section III.3.4) in our performance formula, was just the diagonal element of signal covariance matrix, we realized that all performance analysis could possibly be expressed in terms of *physical* parameters.

Under the guidance of Dr. Fu Li, we continued our work with this in mind and finished our papers in the summer of 1991 [1] [2]. This research was then extended to beamspace processing [3] [4], forward-backward averaging [5] and spatial smoothing analysis which are covered in this thesis. The continuation of this work on seismic signal processing and detection problem is reported in our latest papers [6].

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENT	iii
PREFACE	iv
LIST OF TABLES	xiii
LIST OF FIGURES	ix
CHAPTER	
I INTRODUCTION	1
I.1 Organization of the Thesis	1
I.2 Nomenclature	2
II DOA ESTIMATION ALGORITHMS	4
II.1 Introduction	4
II.2 DOA Estimation Problem Formulation	4
II.3 Subspace Decomposition	6
II.4 Subspace-based DOA Estimation Methods	8
II.4.1 MUSIC	8
II.4.2 Min-Norm	8
II.4.3 Root-MUSIC Algorithm	9
II.4.4 Subspace Rotation Method	9
III PERFORMANCE ANALYSIS	13
III.1 Introduction	13
III.2 Review of Previous Work	14
III.2.1 Perturbation of Subspaces	14
III.2.2 Performance of Extrema-Searching Algorithms	16

III.2.3	Performance of Polynomial-Rooting Algorithms . . .	17
III.2.4	Performance of Subspace-Rotation Algorithms . . .	19
III.2.5	Statistical Performance of DOA Estimation . . .	20
III.3	Performance Analysis Using Physical Parameters . . .	21
III.3.1	Subspace-Rotation Method . . .	22
III.3.2	Min-Norm . . .	24
III.3.3	MUSIC . . .	25
III.3.4	Simplification of $ \beta_k ^2$. . .	26
III.3.5	Final Expression for MSE . . .	27
III.4	Observations . . .	29
III.4.1	The Effect of Number of Snapshots . . .	29
III.4.2	The Effect of Signal Coherence . . .	29
III.4.3	The Effect of Source Separation . . .	30
III.4.4	The Effect of the Number of Sensors . . .	32
III.5	Numerical Examples . . .	36
III.6	Summary . . .	37
IV	BEAMSPACE ANALYSIS . . .	46
IV.1	Introduction . . .	46
IV.2	Beamspace Processing Review . . .	47
IV.3	Beamspace Performance . . .	48
IV.4	Properties of Beamspace DOA Estimation Algorithms . . .	49
IV.5	Optimal Beamformers . . .	52
IV.6	Numerical Examples . . .	54
IV.7	Summary . . .	55
V	FORWARD-BACKWARD AVERAGING . . .	61
V.1	Introduction . . .	61
V.2	Problem Formulation . . .	62
V.3	Statistical Effect of Forward-Backward Averaging . . .	62
V.3.1	MSE Expression . . .	63
V.3.2	Properties . . .	63

V.3.3	Uncorrelated Signals	65
V.3.4	Correlated Signals	66
V.4	Numerical Examples	68
V.5	Summary	69
VI	SPATIAL SMOOTHING	74
VI.1	Introduction	74
VI.2	Forward Only Spatial Smoothing	75
VI.2.1	Data Formulation	75
VI.2.2	Performance Analysis	76
VI.3	Forward-Backward Spatial Smoothing	77
VI.3.1	Data Formulation	77
VI.3.2	Performance Analysis	78
VI.4	Effect of Spatial Smoothing on Performance	78
VI.4.1	Discussion	79
VI.4.2	Uncorrelated Signals	79
VI.4.3	Correlation Signals	82
VI.5	Numerical Examples	83
VI.6	Summary	84
VII	CONCLUSION	95
	REFERENCES	96
	APPENDICES	
A	PARAMETERS FOR SUBSPACE ROTATION METHOD	102
B	FORWARD BACKWARD AVERAGING	106
C	FORWARD ONLY SPATIAL SMOOTHING	111
D	FORWARD BACKWARD SPATIAL SMOOTHING	114

LIST OF TABLES

TABLE		PAGE
I	Parameters of DOA Estimation Pertubation	38
II	Parameters of DOA Estimation Error	39
III	Parameters of Alternative MSE Expression	39
IV	Parameters of Beamspace MSE	57
V	ϕ_k for Different Algorithms	70

LIST OF FIGURES

FIGURE		PAGE
1	A Typical Array Scene.	11
2	Subspace Decomposition: A Three-sensor Two-source Example. . . .	12
3	Root Mean-Squared Error vs. SNR.	40
4	Root Mean-Squared Error vs. Source Correlation.	41
5	Root Mean-Squared Error vs. Source Separation.	42
6	Root Mean-Squared Error vs. Number of Snapshots.	43
7	Root Mean-Squared Error vs. Number of Sensors for MUSIC.	44
8	Root Mean-Squared Error vs. Number of Sensors for ESPRIT.	45
9	Root MSE vs. SNR.	58
10	Root MSE vs. Dimension of Beamformer 1.	59
11	Root MSE vs. Dimension of Beamformer 2.	60
12	Forward-only and Forward-backward Min-Norm.	71
13	Forward-only and Forward-backward MUSIC.	72
14	Forward-only and Forward-backward Subspace-Rotation.	73
15	The Forward-backward Spatial Smoothing Scheme	85
16	Root Mean-Squared Error vs. Subarray Aperture.	86
17	Root Mean-Squared Error vs. Subarray Aperture.	87
18	Root Mean-Squared Error vs. Subarray Aperture.	88
19	Root Mean-Squared Error vs. Subarray Aperture.	89
20	Root Mean-Squared Error vs. Subarray Aperture.	90

21 Root Mean-Squared Error vs. Subarray Aperture. 91

22 Forward Only vs. Forward-backward Spatial Smoothing. 92

23 Forward Only vs. Forward-backward Spatial Smoothing. 93

24 Forward Only vs. Forward-backward Spatial Smoothing. 94

CHAPTER I

INTRODUCTION

This thesis addresses an aspect of the general problem of arrival time estimation at an array of sensors of a sequence of wavefronts emitted at physically distinct locations. Our interest is in direction-of-arrival (DOA) estimation.

Previous work in performance analysis was based on secondary parameters, parameters which depend on the specifics of the particular data trials, making comparison of distinct algorithms difficult. We here express the DOA statistics in terms of fundamental parameters which are independent of the particular data structures and allow us to observe direct relations between processing performance and signal measurement conditions.

Major results include the following: unified performance analysis of element space DOA estimation algorithms; unified performance analysis of beamspace DOA estimation algorithms; and the effects of forward-backward averaging and spatial smoothing on the performance of DOA estimation.

1.1 ORGANIZATION OF THE THESIS

The thesis is organized as follow:

- Chapter I: General introduction.
- Chapter II: Review of the DOA estimation problem and DOA estimation algorithms.

- Chapter III: Review of subspace perturbation theory and the unified statistical performance analysis. Develop unified formula for the statistics of element-space DOA estimation algorithms using *physical* parameters.
- Chapter IV: Extension of the analysis to the beamspace DOA estimation problem.
- Chapter V: Performance analysis of DOA estimation algorithms using forward-backward averaged data.
- Chapter VI: Performance analysis of DOA estimation algorithms using spatially smoothed data.
- Chapter VII: Conclusion.

I.2 NOMENCLATURE

Matrix Operation:

$(\cdot)^T$ = transpose of matrix $(\cdot)$

$(\cdot)^*$ = complex-conjugate of matrix $(\cdot)$

$(\cdot)''$ = complex-conjugate transpose of matrix $(\cdot)$

$\mathbf{e}_k''$ = $(0 \cdots 1 \cdots 0)$, 1 in the k -th position

$\mathbf{A}^\dagger$ = pseudo-inverse of a full-rank matrix $\mathbf{A}$; $\mathbf{A}^\dagger = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$

$\mathbf{I}$ = identity matrix

$\mathbf{J}$ = permutation matrix, $\mathbf{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$(\cdot)(i, j)$ = the i, j -th element of matrix $(\cdot)$

$\text{Tr}(\cdot)$ = the trace of matrix $(\cdot)$

$\text{Im}(\cdot)$ = imaginary part of matrix $(\cdot)$

$\text{Re}(\cdot)$ = real part of matrix $(\cdot)$

$E(\cdot)$ = the expectation of matrix $(\cdot)$

$\mathbf{A}^\dagger$ = the matrix $\mathbf{A}$ with the first row deleted

$\mathbf{A}^\downarrow$ = the matrix $\mathbf{A}$ with the last row deleted

$\Delta(\cdot)$ = perturbed quantity of $(\cdot)$

$(\tilde{\cdot})$ = noise-corrupted quantity of $(\cdot)$

Principal Symbols:

$\mathbf{S}$ = signal matrix $\mathbf{S} = (\mathbf{s}(1), \dots, \mathbf{s}(M))$

$\mathbf{N}$ = noise matrix $\mathbf{N} = (\mathbf{n}(1), \dots, \mathbf{n}(M))$

$\mathbf{Y}$ = data observed from sensor array $\mathbf{Y} = (\mathbf{y}(1), \dots, \mathbf{y}(M))$

$\mathbf{R}$ = covariance matrix of data

$\mathbf{a}(\theta)$ = steering vector determined by array geometry

$\mathbf{A}$ = array manifold $\mathbf{A} = (\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_P))$

$\Sigma, \mathbf{U}, \mathbf{V}$ = singular values, left- and right- singular vectors

$\Sigma_s, \mathbf{U}_s, \mathbf{V}_s$ = signal singular values, signal left- and right- singular vectors

$\Sigma_o, \mathbf{U}_o, \mathbf{V}_o$ = noise singular values, signal left- and right- singular vectors

Ω_s = signal-subspace $\Omega_s = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$

Ω_o = orthogonal-subspace $\Omega_o = \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$

$\mathbf{R}_s$ = signal covariance

θ = direction-of-arrival plane-wave signal

σ_n^2 = power of observation noise

σ_s^2 = power of source

ρ = correlation coefficient between signals

λ_k = $e^{j \frac{2\pi d}{\lambda_c} \sin \theta_k}$

$\mathbf{D} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_K \end{pmatrix}$

CHAPTER II

DOA ESTIMATION ALGORITHMS

II.1 INTRODUCTION

MUSIC [7] and Min-Norm [8] are subspace-based DOA estimation algorithms which utilize the orthogonality property between the signal and orthogonal subspaces. Both of them locate the DOAs by searching for zeros of the null-spectrum. When a uniform line array is used, Root-MUSIC and Min-Norm can be employed for DOA estimation by searching for the roots of polynomial.

Subspace Rotation [9] (such as State-Space Realization (TAM) [10], ESPRIT [11]) is method for DOA estimation which exploit the rotation invariant property of two subarrays. No searching is involved in these methods, and therefore the computational requirements are less.

In this chapter, we first give the data model for the DOA estimation problem and then introduce the concept of subspace decomposition. Then, the subspace algorithms proposed in the last decade for the problem of estimating directions of arrival (DOA) will be reviewed briefly.

II.2 DOA ESTIMATION PROBLEM FORMULATION

We consider P plane waves simultaneously incident on an array of L sensors (Figure 1). The emitter signals are assumed to be narrowband. The propagation time across the array is small compared to the time variations of the amplitude and

phase modulation of the carrier frequency. Thus the propagation of a wavefront between sensor elements is modeled as a simple phase delay. For the purpose of comparison, we will generally use a uniform line array geometry; other geometries will be specified in different places.

The signal arriving at the i -th sensor at time t is

$$y_i(t) = \sum_{k=1}^P e^{j\frac{2\pi d}{\lambda_c} \sin \theta_k (i-1)} s_k(t)$$

$$= \begin{pmatrix} e^{j\frac{2\pi d}{\lambda_c} (i-1) \sin \theta_1} & \dots & e^{j\frac{2\pi d}{\lambda_c} (i-1) \sin \theta_P} \end{pmatrix} \begin{pmatrix} s_1(t) \\ \vdots \\ s_P(t) \end{pmatrix} \quad (1)$$

where $s_k(t)$ is the k -th narrowband signal (with center wavelength λ_c) arriving at angle θ_k , and d is the spacing between adjacent sensors. The signals arriving at all the sensors at time t are

$$\mathbf{y}(t) = \begin{pmatrix} 1 & \dots & 1 \\ e^{j\frac{2\pi d}{\lambda_c} \sin \theta_1} & \dots & e^{j\frac{2\pi d}{\lambda_c} \sin \theta_P} \\ \vdots & \ddots & \vdots \\ e^{j\frac{2\pi d}{\lambda_c} (L-1) \sin \theta_1} & \dots & e^{j\frac{2\pi d}{\lambda_c} (L-1) \sin \theta_P} \end{pmatrix} \begin{pmatrix} s_1(t) \\ \vdots \\ s_P(t) \end{pmatrix}$$

or in vector form,

$$\mathbf{y}(t) \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{a}(\theta_1) & \dots & \mathbf{a}(\theta_P) \end{pmatrix} \mathbf{s}(t) \stackrel{\text{def}}{=} \mathbf{A}(\theta) \mathbf{s}(t)$$

where $\mathbf{A}(\theta)$ and $\mathbf{a}(\theta)$ (steering vector) represent the array characteristics. Take M snapshots of the signal to form a data matrix as

$$\mathbf{Y} = \begin{pmatrix} y_1(1) & \dots & y_1(M) \\ \vdots & \ddots & \vdots \\ y_L(1) & \dots & y_L(M) \end{pmatrix} = \mathbf{A}(\theta) \begin{pmatrix} \mathbf{s}(1) & \dots & \mathbf{s}(M) \end{pmatrix}. \quad (2)$$

The localization problem is to estimate the DOAs $(\theta_1, \dots, \theta_P)$ from the array output $\mathbf{Y}$. To solve this problem, we make the following assumptions regarding the array, the signals and the noise:

A1: The number of signals is known and is smaller than the number of sensors, namely, $P < L$.

A2: Every set of P steering vectors is linearly independent.

A3: The noise is circular, stationary and ergodic, complex valued Gaussian process with zero mean and covariance matrix $E(\mathbf{N}\mathbf{N}^H) = \sigma_n^2 \mathbf{M}\mathbf{I}$. The circular property imply that $E(\mathbf{N}\mathbf{N}^T) = 0$.

II.3 SUBSPACE DECOMPOSITION

The subspace decomposition can be performed on the data matrix $\mathbf{Y}$ by a singular value decomposition (SVD), or on the covariance matrix $\mathbf{R} = \mathbf{Y}\mathbf{Y}^H$ by an eigenvalue decomposition. The subspace decomposition using SVD on the direct data matrix $\mathbf{Y}$ is as follows (Figure 2):

$$\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \begin{pmatrix} \mathbf{U}_s & \mathbf{U}_o \end{pmatrix} \begin{pmatrix} \mathbf{\Sigma}_s & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{V}_s^H \\ \mathbf{V}_o^H \end{pmatrix}$$

where $\mathbf{U}_s$ are the singular vectors ($L \times P$ matrix) associated with the P non-zero singular values, while $\mathbf{U}_o$ are the singular vectors ($L \times L - P$ matrix) associated with the zero singular values. The singular vectors are assumed normalized so that $\mathbf{U}^H\mathbf{U} = \mathbf{I}$. Subspace algorithms are based on two important properties of the subspaces:

- The vectors in $\mathbf{U}_s$ span the signal subspace which is defined as the span of the columns of the array manifold matrix $\mathbf{A}(\theta)$. The signal subspace projection matrix can be calculated by

$$\mathbf{\Omega}_s = \mathbf{U}_s\mathbf{U}_s^H = \mathbf{A} \left(\mathbf{A}^H\mathbf{A} \right)^{-1} \mathbf{A}^H$$

- The vectors in $\mathbf{U}_o$ span the orthogonal subspace¹ which is the orthogonal complement of the array manifold. This orthogonality between the orthogonal-subspace and array manifold can be expressed as

$$\mathbf{a}(\theta_k)^H \mathbf{U}_o = 0, \quad k = 1, \dots, P. \quad (3)$$

or

$$\mathbf{a}(\theta_k)^H \boldsymbol{\Omega}_o = 0, \quad k = 1, \dots, P. \quad (4)$$

where

$$\boldsymbol{\Omega}_o = \mathbf{U}_o \mathbf{U}_o^H = \mathbf{I} - \boldsymbol{\Omega}_s$$

is the orthogonal subspace projection matrix.

The above properties hold for arbitrary array geometry where the entry in the i -th row and k -th column (corresponding to sensor located at $[x_i, y_i]$) of $\mathbf{A}(\theta)$ can be written as

$$a_i(\theta_k) = e^{j \frac{2\pi}{\lambda} (x_i \sin \theta_k + y_i \cos \theta_k)}.$$

As shown in [12], the orthogonal-subspace obtained from an eigenvalue decomposition of the covariance matrix and the orthogonal-subspace from a singular value decomposition of the direct data matrix are the same. In this work, all the algorithms used will be based on direct-data for analytical simplicity. The direct-data algorithms will have the same performance as the covariance-based algorithms, with perhaps slightly better numerical properties. One reason that the direct-data algorithm might not be often used in practice is that the dimensions of the matrices grow

¹The term “noise-subspace” has been previously used to describe what we call the orthogonal subspace. Since this subspace is the orthogonal complement of the signal subspace, and since it is well defined whether or not any noise is present in the problem, we believe the term “orthogonal-subspace” is more appropriate.

with the data length, while the covariance matrix has fixed dimensions. However, for the purpose of *analysis*, the direct-data algorithm is much easier to deal with.

II.4 SUBSPACE-BASED DOA ESTIMATION METHODS

II.4.1 MUSIC

MUSIC performs a one-dimensional search for the P zeros over θ of the null-spectrum

$$P_{MU}(\theta) = \mathbf{a}(\theta)^H \mathbf{U}_o \mathbf{U}_o^H \mathbf{a}(\theta) = \mathbf{a}(\theta)^H \boldsymbol{\Omega}_o \mathbf{a}(\theta) \quad (5)$$

In the noise-free case, $P_{MU}(\theta)$ reaches zeros at the true DOAs. When the data is corrupted by noise, the P minimums of the null-spectrum are regarded as the DOA estimated.

In this and future equations, the symbol θ *without* a subscript is a scalar variable which represents a possible direction of arrival, while the subscripted symbol θ_k , $k = 1, \dots, P$ refers to the actual directions of arrival in the noise-free data.

II.4.2 Min-Norm

For uniform line array, Min-Norm searches the roots of polynomial

$$\mathbf{a}(z^{-1})^T \mathbf{d} = \prod_{i=1}^{L-1} (1 - r_i z^{-1}) \quad \text{where} \quad \mathbf{d} = \mathbf{U}_o \frac{\mathbf{c}}{\|\mathbf{c}\|^2} \stackrel{\text{def}}{=} \mathbf{U}_o \bar{\mathbf{c}} = \frac{\boldsymbol{\Omega}_o \mathbf{e}_1}{\|\boldsymbol{\Omega}_o \mathbf{e}_1\|^2} \quad (6)$$

$\mathbf{c}^H = \mathbf{e}_1^H \mathbf{U}_o$ which is the first row of $\mathbf{U}_o$ and $z \stackrel{\text{def}}{=}} e^{j \frac{2\pi d}{\lambda_c} \sin \theta}$. The P roots which have largest amplitudes are chosen as the signal-roots and the rest are regarded as noise roots.

For arbitrary array geometry, the Min-Norm algorithm [13] searches for the P zeros over θ of the null-spectrum

$$P_{MN}(\theta) = \mathbf{a}(\theta)^H \mathbf{d} \mathbf{d}^H \mathbf{a}(\theta). \quad (7)$$

In this case, Min-Norm can be treated as a *Weighted* MUSIC where the null-spectrum can be defined as

$$P_{MN}(\theta) = \mathbf{a}(\theta)^H \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}(\theta) \quad (8)$$

where $\mathbf{W} = \mathbf{I}$ for MUSIC and $\mathbf{W} = \bar{\mathbf{c}}\bar{\mathbf{c}}^H$ for Min-Norm.

II.4.3 Root-MUSIC Algorithm

When a uniform line array is used, Root-MUSIC forms and roots the null-spectrum polynomial

$$P_{RM}(z) = \mathbf{a}(z^{-1})^T \mathbf{U}_o \mathbf{U}_o^H \mathbf{a}(z) = A \prod_{i=1}^{L-1} (1 - r_i z^{-1})(1 - r_i^* z). \quad (9)$$

The polynomial $P_{RM}(z)$ has $2(L-1)$ roots. But unlike Min-Norm, Root-MUSIC always chooses the P roots with largest amplitudes *inside* the unit-circle. This choice results in a bias in the radial direction of the estimated roots since when white observation noise is present the signal-roots will be perturbed inside and outside the unit-circle. However, DOA estimates are only functions of the angles of the roots, not the radii. Thus the radial bias does not affect the DOA estimates obtained by Root-MUSIC.

II.4.4 Subspace Rotation Methods

State-Space Realization [10] and ESPRIT [11] are two typical Subspace Rotation DOA estimation methods among others. The basic assumption here is that the array manifold has the *shift invariant* property:

$$\mathbf{A}^\dagger \mathbf{D} = \mathbf{A}^\dagger \quad (10)$$

where

$$\mathbf{D} = \begin{pmatrix} e^{j\frac{2\pi d}{\lambda_c} \sin \theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{j\frac{2\pi d}{\lambda_c} \sin \theta_P} \end{pmatrix}$$

which contains the DOA information. This property can be found in a uniform linear array or using array composed of matched sensor doublets.

Since $\mathbf{U}_s$ and $\mathbf{A}$ has the same span, if we choose $\Phi = \mathbf{U}_s \mathbf{W}$ ($\mathbf{W}$ is a nonsingular $P \times P$ matrix which could be fixed or data dependent. For example, $\mathbf{W} = \Sigma_s^{\frac{1}{2}}$ for SSR and $\mathbf{I}$ for ESPRIT), then we have the relation

$$\Phi^\dagger \mathbf{F} = \Phi^\dagger \quad (11)$$

where $\mathbf{F}$ is similar to $\mathbf{D}$

$$\mathbf{F} = \mathbf{L} \mathbf{D} \mathbf{L}^{-1}.$$

So by uniquely solving

$$\mathbf{F} = \Phi^{\dagger \dagger} \Phi^\dagger \quad (12)$$

the eigenvalues of $\mathbf{F}$ are the same as that of $\mathbf{D}$ which can be used for the DOA estimation.

Figure 1. A Typical Array Scene.

Figure 2. Subspace Decomposition: A Three-sensor Two-source Example.

CHAPTER III

PERFORMANCE ANALYSIS

III.1 INTRODUCTION

Early study on performance of DOA estimation algorithms were based on simulations, which did not provide the insight of the performance. Pioneered by Kaveh and Barabell [14], analytical performance evaluation has attracted much excellent research, (see [9, 15, 16, 17, 18, 19, 20, 21, 22] among many others), various statistical properties have been revealed. However, the following limitations occur in the performance study: (1) the use of singular values and singular vectors which are nonlinear functions of either the algorithms or data structures. (2) analyzing individual algorithms, making comparison with different methods difficult; (3) assuming unbounded data sets are available (even in analyzing finite data effect induced by observation noise), which may not be a realistic assumption; (4) involving complicated mathematics and statistics, thus resulting in untractable expressions;

By expressing the subspace perturbation in terms of data perturbation using first order linear approximation, Li and Vaccaro[12, 23, 24, 25, 26] developed a performance analysis of the subspace processing algorithms in a unified, and self-contained fashion. The analysis was based on bounded data measurements. The common models they established successfully overcome the second and third limitation listed.

Unfortunately, like most of the previous performance analysis, the formulas

they developed for the statistics of DOA estimation were expressed in terms of unperturbed data parameter such as the singular vectors and singular values. These singular vectors and singular values are derived from the original data via nonlinear mathematical transformation, and are data-trial dependent. The relations between DOAs and source separations and coherence, numbers of sensors and snapshots are not revealed explicitly, rather are contained only in numerical plots. So the ultimate goal is to express the DOA error statistics in terms of fundamental parameters such as array manifold, signal covariance, and number of the snapshots and sensors, rather than data-trial dependent singular values and vectors.

In this chapter, we develop a unified performance analysis.

III.2 REVIEW OF PREVIOUS WORK

In this section, we briefly review the subspace perturbation theory developed Li and Vaccaro [25]. The performance analysis of DOA estimation algorithms using this theory was introduced afterwards.

III.2.1 Perturbation of Subspaces

Various data perturbations are always present in practice which results in the perturbation of estimated subspaces. The perturbed data matrix can be written as

$$\tilde{\mathbf{Y}} = \mathbf{Y} + \Delta\mathbf{Y} = \mathbf{Y} + \mathbf{N} + \Delta\mathbf{A}\mathbf{S}.$$

Where $\mathbf{N}$ is the observation noise and $\Delta\mathbf{A}$ is the sensor error. The subspace decomposition of perturbed data by SVD is

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{U}}\tilde{\Sigma}\tilde{\mathbf{V}}^H = \begin{pmatrix} \tilde{\mathbf{U}}_s & \tilde{\mathbf{U}}_o \end{pmatrix} \begin{pmatrix} \tilde{\Sigma}_s & 0 \\ 0 & \tilde{\Sigma}_o \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{V}}_s^H \\ \tilde{\mathbf{V}}_o^H \end{pmatrix}. \quad (13)$$

We can now write

$$\tilde{\mathbf{U}}_o = \mathbf{U}_o + \Delta\mathbf{U}_o, \text{ and } \tilde{\mathbf{U}}_s = \mathbf{U}_s + \Delta\mathbf{U}_s \quad (14)$$

where $\Delta\mathbf{U}_o$ and $\Delta\mathbf{U}_s$ are the perturbations in the estimated orthogonal- and signal-subspaces. At high SNR (small perturbations), a first-order perturbation expansion can be used to get a linear approximation to the perturbed subspaces.

Lemma III.1 The perturbed orthogonal subspace is spanned by $\mathbf{U}_o + \mathbf{U}_s\mathbf{Q}$ and the perturbed signal-subspace is spanned by $\mathbf{U}_s + \mathbf{U}_o\mathbf{P}$, where $\mathbf{P}$ and $\mathbf{Q}$ are matrices whose norms are of the order of $\|\Delta\mathbf{Y}\|$. The matrix norm can be any submultiplicative norm such as the Euclidean 2-norm or the Frobenius norm.

Proof: See [27, 28].

We now proceed to derive an explicit expression for the matrix $\mathbf{Q}$ which is valid up to first-order. From

$$\tilde{\mathbf{U}}_o^H \tilde{\mathbf{Y}} = \tilde{\Sigma}_o \tilde{\mathbf{V}}_o^H. \quad (15)$$

Using (14) and the fact that $\hat{\Sigma}_o = \Delta\Sigma_o$ (since the noise-free value of Σ_o is $\Sigma_o = 0$) and $\mathbf{V}_o + \mathbf{V}_s\bar{\mathbf{Q}}$ where $\bar{\mathbf{Q}}$ is of the order of $\Delta\mathbf{Y}$, (15) can be written as

$$(\mathbf{U}_o + \mathbf{U}_s\mathbf{Q})^H (\mathbf{Y} + \Delta\mathbf{Y}) = \Delta\Sigma_o (\mathbf{V}_o + \mathbf{V}_s\bar{\mathbf{Q}})^H. \quad (16)$$

By neglecting second order terms and using the fact that $\mathbf{U}_o^H \mathbf{Y} = 0$, we get

$$\mathbf{Q} \doteq -\Sigma_s^{-1} \mathbf{V}_s^H \Delta\mathbf{Y}^H \mathbf{U}_o \quad (17)$$

where we introduce the symbol “ $\doteq$ ” to mean “equal up to first-order terms.” Finally

$$\Delta\mathbf{U}_o \doteq \mathbf{U}_s\mathbf{Q} = -\mathbf{U}_s\Sigma_s^{-1} \mathbf{V}_s^H \Delta\mathbf{Y}^H \mathbf{U}_o. \quad (18)$$

We emphasize the fact that (18) is a general first-order expression for the perturbation of the orthogonal-subspace due to perturbations in the data matrix regardless

of the source of those perturbations. The expression for $\Delta \mathbf{U}_o$ can be used to analyze the performance of *any* algorithm which estimates the orthogonal-subspace from data.

Using the orthogonality between the perturbed orthogonal- and signal- subspaces, we have

$$\tilde{\mathbf{U}}_o^H \tilde{\mathbf{U}}_s = (\mathbf{U}_o + \mathbf{U}_s \mathbf{Q})^H (\mathbf{U}_s + \mathbf{U}_o \mathbf{P}) = 0. \quad (19)$$

Using the orthonormal property of unperturbed subspaces, the above equation reduces to

$$\mathbf{P} = -\mathbf{Q}^H. \quad (20)$$

Thus we have

$$\Delta \mathbf{U}_s \doteq \mathbf{U}_o \mathbf{P} = \mathbf{U}_o \mathbf{U}_o^H \Delta \mathbf{Y} \mathbf{V}_s \Sigma_s^{-1}. \quad (21)$$

III.2.2 Performance of Extrema-Searching Algorithms

A common model for the null spectrum function associated with MUSIC and Min-Norm searching algorithms can be written as

$$P(\theta, \mathbf{U}_o) = \mathbf{a}(\theta)^H \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}(\theta)$$

where the weighting matrix $\mathbf{W}$ equals $\mathbf{I}$ for MUSIC and $\mathbf{W} = \bar{\mathbf{c}}\bar{\mathbf{c}}^H$ for Min-Norm, and θ is a scalar variable which represents a possible direction of arrival. Then by (3)

$$P(\theta_k, \mathbf{U}_o) = 0 \quad \text{for } k = 1, \dots, K. \quad (22)$$

Step 1. Approximate the derivative of the null-spectrum function using the first two terms in its Taylor series expansion about the true angles of arrival at high SNR

$$\frac{\partial P(\hat{\theta}_k, \tilde{\mathbf{U}}_o)}{\partial \theta} = \frac{\partial P(\theta_k, \tilde{\mathbf{U}}_o)}{\partial \theta} + \frac{\partial^2 P(\theta_k, \tilde{\mathbf{U}}_o)}{\partial \theta^2} \Delta \theta_k, \quad k = 1, \dots, K. \quad (23)$$

Since the left-hand side of the above equation is zero, we can solve for $\Delta\theta_k$ from (23) as follows

$$\Delta\theta_k = -\frac{\frac{\partial P(\theta_k, \tilde{\mathbf{U}}_o)}{\partial \theta}}{\frac{\partial^2 P(\theta_k, \tilde{\mathbf{U}}_o)}{\partial \theta^2}}, \quad \text{for } k = 1, \dots, K. \quad (24)$$

Step 2. Substitute $\tilde{\mathbf{U}}_o = \mathbf{U}_o + \Delta\mathbf{U}_o$ and $\tilde{\mathbf{W}} = \mathbf{W} + \Delta\mathbf{W}$ into (24) and approximate (24) by first-order perturbation expressions in $\Delta\mathbf{U}_o$. After some lengthy but straightforward derivation [12], we obtain

$$\Delta\theta_k = \frac{\text{Re}[-\mathbf{a}(\theta_k)^H \Delta\mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k)]}{\mathbf{a}^{(1)}(\theta_k)^H \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k)}. \quad (25)$$

We now substitute (18) into (25) to get

$$\Delta\theta_k = \frac{\text{Im}[j\mathbf{a}(\theta_k)^H \mathbf{U}_s \Sigma_s^{-1} \mathbf{V}_s^H \Delta\mathbf{Y}^H \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k)]}{\mathbf{a}^{(1)}(\theta_k)^H \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k)}.$$

To simplify notation, we define the vectors

$$\alpha_k = j\mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k) \quad \text{and} \quad \beta_k = \mathbf{V}_s \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{a}(\theta_k) \quad (26)$$

and the scalar

$$\gamma_k = \mathbf{a}^{(1)}(\theta_k)^H \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k). \quad (27)$$

Then we have

$$\Delta\theta_k = \frac{\text{Im}[\beta_k^H \Delta\mathbf{Y}^H \alpha_k]}{\gamma_k}.$$

III.2.3 Performance of Polynomial-Rooting Algorithms

A common spectral polynomial for Min-Norm and Root-MUSIC (also Pisarenko's method) can be written as

$$P(z) = \mathbf{a}(z^{-1})^T \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}(z) = A \prod_{i=1}^{L-1} (1 - r_i z^{-1})(1 - r_i^* z) \quad (28)$$

where the weighting matrix $\mathbf{W}$ is the same as in extrema-search algorithm and A is a scaling factor A equals 1 for Min-Norm and $\frac{(L-K)}{\|\mathbf{h}\|^2}$ for MUSIC and $\mathbf{h}$ is the coefficient vector (with first element of unity) for the polynomial $H(z)$. $H(z)$ is the causal spectral factor of $P(z)$, namely, $P(z) = H(z)H(z^{-1})^*$.

Step 1. Take the derivative of $P(z, \tilde{r})$ over z , substitute $\tilde{r} = r + \Delta r$ into $\frac{\partial P(z, \tilde{r})}{\partial z}$ and evaluate it at $z = r_k$. The first-order terms yields

$$\frac{\partial P(z, \tilde{r})}{\partial z} \Big|_{z=r_k} = 2j A r_k^* \text{Im}(\Delta r_k r_k^*) G(r_k) \text{ where } G(r_k) = \prod_{\substack{j=1 \\ j \neq k}}^{L-1} (1 - r_j r_k^{-1})(1 - r_j^* r_k). \quad (29)$$

Step 2. Take the derivative of $P(z, \tilde{\mathbf{U}}_o)$ with respect to z , substitute $\tilde{\mathbf{U}}_o = \mathbf{U}_o + \Delta \mathbf{U}_o$ and $\tilde{\mathbf{W}} = \mathbf{W} + \Delta \mathbf{W}$ into $\frac{\partial P(z, \tilde{\mathbf{U}}_o)}{\partial z}$ and evaluate it at $z = r_k$. The first-order terms yields

$$\frac{\partial P(z, \tilde{\mathbf{U}}_o)}{\partial z} \Big|_{z=r_k} = -2j r_k^* \text{Im}[r_k^* \mathbf{a}(r_k)^H \Delta \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(r_k)]. \quad (30)$$

Step 3. Equate (29) and (30), we obtain

$$2j A r_k^* \text{Im}\left(\frac{\Delta r_k}{r_k}\right) G(r_k) = -2j r_k^* \text{Im}[r_k^* \mathbf{a}(r_k)^H \Delta \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(r_k)]. \quad (31)$$

Using $r_k^* = r_k^{-1}$ and angle-root relation given in [20]

$$\Delta \theta_k = C_k \text{Im}\left[\frac{\Delta r_k}{r_k}\right] = C_k \frac{-\text{Im}[r_k^{-1} \mathbf{a}(r_k)^H \Delta \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(r_k)]}{AG(r_k)} \quad (32)$$

where $C_k = \frac{\lambda_c}{2\pi d \cos \theta_k}$. Substitute (18) into (32) to obtain

$$\Delta \theta_k = \frac{\text{Im}[C_k r_k^{-1} \mathbf{a}(r_k)^H \mathbf{U}_s \Sigma_s^{-1} \mathbf{V}_s^H \Delta \mathbf{Y}^H \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(r_k)]}{AG(r_k)}. \quad (33)$$

If we define

$$\alpha_k = \mathbf{U}_o \mathbf{W} \mathbf{U}_o^H \mathbf{a}^{(1)}(r_k) C_k r_k^{-1} \text{ and } \beta_k = \mathbf{V}_s \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{a}(r_k)$$

then (33) can be simplified as

$$\Delta\theta_k = \frac{\text{Im}[\beta_k^H \Delta \mathbf{Y}^H \alpha_k]}{AG(r_k)}. \quad (34)$$

III.2.4 Performance of Subspace-Rotation Algorithms

Common model for Subspace-Rotation can be defined as

$$\Phi^\dagger \mathbf{F} = \Phi^\dagger \quad (35)$$

where $\Phi^\dagger$ and $\Phi^\dagger$ for uniform line array are the first $L - 1$ rows and last $L - 1$ rows, respectively, of matrix Φ

$$\Phi = \mathbf{U}_s \mathbf{W}. \quad (36)$$

The weighting matrix $\mathbf{W}$ can be specialized for different Matrix-Shifting algorithms: $\mathbf{W} = \Sigma_s^{\frac{1}{2}}$ for SSR, $\mathbf{W} = \mathbf{I}$ for ESPRIT. This type of weighting matrices for SSR and ESPRIT was first used in [9]. Since the true and estimated signal roots (eigenvalues of $\mathbf{F}$ or elements in diagonal matrix $\mathbf{D}$) are identical for SSR and ESPRIT through a similarity transform, so their statistical performances are identical too.

Step 1. Calculate the perturbation of the transition matrix $\Delta \mathbf{F}$ due to data perturbations. We have

$$(\Phi^\dagger + \Delta \Phi^\dagger)(\mathbf{F} + \Delta \mathbf{F}) = (\Phi^\dagger + \Delta \Phi^\dagger). \quad (37)$$

Cancel $\Phi^\dagger \mathbf{F}$ and $\Phi^\dagger$ as (35) and neglect the second-order term $\Phi^\dagger \Delta \mathbf{F}$ to obtain

$$\Delta \mathbf{F} = \Phi^{\dagger \dagger} (\Phi^\dagger - \Phi^\dagger \mathbf{F}).$$

Step 2. Calculate the first-order perturbation of the eigenvalues of $\tilde{\mathbf{F}}$ due to $\Delta \mathbf{F}$ is [29]

$$\Delta \lambda_k = \mathbf{v}_k \Delta \mathbf{F} \mathbf{u}_k = \mathbf{v}_k \Phi^{\dagger \dagger} (\Phi^\dagger - \Phi^\dagger \mathbf{F}) \mathbf{u}_k = \mathbf{v}_k \Phi^{\dagger \dagger} (\Phi^\dagger - \lambda_k \Phi^\dagger) \mathbf{u}_k \quad (38)$$

where $\mathbf{v}_k$ and $\mathbf{u}_k$ are left- and right- orthonormal eigenvectors of $\mathbf{F}$ associated with λ_k . Further

$$\begin{aligned}\Delta\theta_k &= C_k \text{Im}\left(\frac{\Delta\lambda_k}{\lambda_k}\right) \\ &= C_k \text{Im}[\mathbf{v}_k \mathbf{W}^{-1} \mathbf{U}_s^{\dagger\top} (\Delta \mathbf{U}_s^{\dagger} \lambda_k^{-1} - \Delta \mathbf{U}_s^{\dagger}) \mathbf{W} \mathbf{u}_k].\end{aligned}\quad (39)$$

Substitute (21) into (39), we get

$$\Delta\theta_k = \text{Im}[C_k \mathbf{v}_k \mathbf{W}^{-1} \mathbf{U}_s^{\dagger\top} (\mathbf{U}_o^{\dagger} \lambda_k^{-1} - \mathbf{U}_o^{\dagger}) \mathbf{U}_o^H \Delta \mathbf{Y} \mathbf{V}_s \Sigma_s^{-1} \mathbf{W} \mathbf{u}_k].\quad (40)$$

Define $\alpha_k^H = C_k \mathbf{v}_k \mathbf{W}^{-1} \mathbf{U}_s^{\dagger\top} (\mathbf{U}_o^{\dagger} \lambda_k^{-1} - \mathbf{U}_o^{\dagger}) \mathbf{U}_o^H$ and $\beta_k = \mathbf{V}_s \Sigma_s^{-1} \mathbf{W} \mathbf{u}_k$, we obtain

$$\Delta\theta_k = \text{Im}[\beta_k^H \Delta \mathbf{Y}^H \alpha_k].\quad (41)$$

III.2.5 Statistical Performance of DOA Estimation

The general expression for perturbation of DOA estimate is

$$\Delta\theta_k = \frac{\text{Im}[\alpha_k^H \Delta \mathbf{Y} \beta_k]}{\gamma_k}\quad (42)$$

where the parameters α_k , β_k , and γ_k are given in Table I.

Since we have used a first-order perturbation analysis in which the element of data perturbation $\Delta \mathbf{Y}$ is assumed to be *iid* and of zero mean, the predicted bias of an estimated direction of arrival is zero.

The mean-squared error of an estimated direction of arrival is given by

$$E_{\Delta \mathbf{Y}}(\Delta\theta_k)^2 = \frac{E_{\Delta \mathbf{Y}}[\text{Im}(\beta_k^H \Delta \mathbf{Y}^H \alpha_k)]^2}{\gamma_k^2}.\quad (43)$$

The right-hand side of the above equation is a function of the perturbation matrix $\Delta \mathbf{Y}$. Very often this perturbation is due to the noise matrix $\mathbf{N}$ which is an uncorrelated circular random variables with equal variances $\sigma_n^2/2$ (for real and imaginary

part respectively), then (see [12])

$$\begin{aligned}
 E_N(\Delta\theta_k)^2 &= \frac{E[|\beta_k^H \mathbf{N}^H \alpha_k|^2]}{2\gamma_k^2} \\
 &= \frac{||\alpha_k||^2 ||\beta_k||^2 \sigma_n^2}{2\gamma_k^2}.
 \end{aligned} \tag{44}$$

Several analytical conclusion can be observed:

- Using the Schwarz inequality, we can see the mean-squared error of the Min-Norm Algorithm is bounded below by that of MUSIC.
- Using L'Hospital's rule, we can see that the mean-squared error of the extrema-searching approach is identical to that of the polynomial-rooting approach if the sources are distinct in the spectrum.
- Using a similarity transformation, we can see that State-Space Realization and ESPRIT have identical performance when a uniform line array is used.

III.3 PERFORMANCE ANALYSIS USING PHYSICAL PARAMETERS

As discussed earlier on, these parameters in Table I are in forms of unperturbed singular values, and left- and right- singular vectors, and these singular values and vectors obtained from highly nonlinear transformation, namely, SVD prevent us from seeing the contributions from the real factors, such as source separations and signal coherence, numbers of sensors and snapshots, thus offers little insight for design. In this chapter we will significantly improve the performance analysis by making unification of those parameters and further simplification of MSE expression to remove the data-dependent singular values and vectors. In what follows, we will analyze first Subspace-Rotation (State-Space Realization and ESPRIT), and then

Min-Norm and MUSIC. For MUSIC and Min-Norm, we have shown that extrema-searching approach and polynomial approach have the identical performance, therefore we will only show the derivation of extrema-searching approach.

Before proceed further, it is very useful to distinguish the *physical parameters* which are independent of data realization (trials)

1. Number of sensors L (embedded in $\mathbf{A}$)
2. Number of snapshots M
3. Source separation $\theta_k - \theta_i$ (embedded in $\mathbf{A}$)
4. Source covariance $\mathbf{R}_s$
5. Array Manifold $\mathbf{A}$
6. Signal-Subspace $\mathbf{\Omega}_s = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$
7. Orthogonal-Subspace $\mathbf{\Omega}_o = \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$
8. Noise power σ_n^2

and the *indirect parameters* which are dependent to data realization and decomposition

1. Singular values $\mathbf{\Sigma}_s$
2. Left Singular vectors $\mathbf{U}_s$
3. Right Singular vectors $\mathbf{V}_s$.

The physical parameters are desired for the final expression. Indirect parameters are not needed for the final expression.

III.3.1 Subspace-Rotation Method

The subspace shift-invariance relation can be expressed as

$$\mathbf{A}^\dagger \mathbf{D} = \mathbf{A}^\dagger \quad (45)$$

while in general $\Phi = \mathbf{U}_s \mathbf{W}$, and the eigenvalues of transition matrix $\mathbf{F}$ give the information of DOA, thus $\mathbf{F}$ is similar to $\mathbf{D}$

$$\mathbf{F} = \mathbf{L} \mathbf{D} \mathbf{L}^{-1} \quad (46)$$

where $\mathbf{L}$ is eigenvector matrix of $\mathbf{F}$. Substitute (36) and (46) into (35), we get

$$(\mathbf{U}_s \mathbf{W})^\dagger \mathbf{L} \mathbf{D} \mathbf{L}^{-1} = (\mathbf{U}_s \mathbf{W})^\dagger$$

or

$$(\mathbf{U}_s \mathbf{W} \mathbf{L})^\dagger \mathbf{D} = (\mathbf{U}_s \mathbf{W} \mathbf{L})^\dagger. \quad (47)$$

Compare (47) with (45), we have

$$\mathbf{A} = \mathbf{U}_s \mathbf{W} \mathbf{L} \quad (48)$$

so that

$$\mathbf{L} = \mathbf{W}^{-1} \mathbf{U}_s^H \mathbf{A}. \quad (49)$$

Also it is easy to verify that the inverse of $\mathbf{L}$ can take the form of

$$\mathbf{L}^{-1} = \mathbf{R}_s \mathbf{A}^H \mathbf{U}_s \Sigma_s^{-1} \mathbf{W}. \quad (50)$$

From the derivation in Appendix A, we can get the simplified MSE expression for Subspace-Rotation algorithms

$$E_{\Delta \mathbf{Y}}(\Delta \theta_k)^2 = \frac{||\beta_k||^2 \sigma_n^2}{2||\bar{\alpha}_k \Omega_{\phi} \mathbf{a}^{(1)}(\theta_k)||^2} \quad (51)$$

where $\bar{\alpha}_k = \frac{\alpha_k}{\|\alpha_k\|}$ and

$$\alpha_k^H = C_k \mathbf{e}_k^H (\mathbf{A}^{\uparrow \dagger} \mathbf{I}^\uparrow - \mathbf{A}^{\downarrow \dagger} \mathbf{I}^\downarrow) \quad (52)$$

as defined in Appendix A

We now have the following observations:

1. α_k and β_k are independent of $\mathbf{W}$ and thus are irrelevant to the algorithm (either SSR or ESPRIT) used. In other words, SSR and ESPRIT both have the same performance in mean-squared error sense, this agrees with [23, 30].
2. β_k for SSR and ESPRIT is the same as β_k of MUSIC and Min-Norm. This is the further unification of the mean-squared error expression.

III.3.2 Min-Norm

In Min-Norm algorithm

$$\bar{\mathbf{c}} = \frac{\mathbf{c}}{\|\mathbf{c}\|^2} = \frac{\mathbf{U}_o \mathbf{e}}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^2} \quad (53)$$

where $\mathbf{e} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^H$, and $\mathbf{c}^H$ is the first row of $\mathbf{U}_o$. We can now express α_k in Table I as

$$\alpha_k = j \mathbf{U}_o \bar{\mathbf{c}} \bar{\mathbf{c}}^H \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k) = \frac{j \boldsymbol{\Omega}_o \mathbf{e} \mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^4}. \quad (54)$$

Take norm-square of α_k

$$\begin{aligned} \|\alpha_k\|^2 &= \frac{\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{e} \mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{e} \mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^8} \\ &= \frac{(\mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{e}) [\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{e} \mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)]}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^8} \end{aligned} \quad (55)$$

since $\mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{e}$ is a scalar, we can have

$$\|\alpha_k\|^2 = \frac{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^2 \|\mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^8} = \frac{\|\mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^6} \quad (56)$$

and γ_k in Table I is

$$\gamma_k = \mathbf{a}^{(1)}(\theta_k)^H \mathbf{U}_o \bar{\mathbf{c}} \bar{\mathbf{c}}^H \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k) = \frac{\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{e} \mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^4} \quad (57)$$

so that

$$\gamma_k = \frac{\|\mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^4} = \|\alpha_k\|^2 \|\mathbf{e}^H \boldsymbol{\Omega}_o\|^2. \quad (58)$$

The mean-squared error is

$$E_{\Delta Y}(\Delta \theta_k)^2 = \frac{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|^2 \|\beta_k\|^2 \sigma_n^2}{2 \|\mathbf{e}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2}. \quad (59)$$

Now introducing a new vector $\bar{\mathbf{e}}$

$$\bar{\mathbf{e}} = \frac{\boldsymbol{\Omega}_o \mathbf{e}}{\|\mathbf{e}^H \boldsymbol{\Omega}_o\|} \quad (60)$$

(59) can rewritten as

$$E_{\Delta Y}(\Delta \theta_k)^2 = \frac{\|\beta_k\|^2 \sigma_n^2}{2 \|\bar{\mathbf{e}}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2}. \quad (61)$$

III.3.3 MUSIC

For MUSIC, it is rather straightforward to get

$$\alpha_k = j \mathbf{U}_o \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k) = j \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) \quad (62)$$

and

$$\gamma_k = \mathbf{a}^{(1)}(\theta_k)^H \mathbf{U}_o \mathbf{U}_o^H \mathbf{a}^{(1)}(\theta_k) = \mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) \quad (63)$$

thus we have

$$\gamma_k = \|\boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2 = \|\alpha_k\|^2. \quad (64)$$

The expression of mean-squared error is

$$E_{\Delta Y}(\Delta \theta_k)^2 = \frac{\|\beta_k\|^2 \sigma_n^2}{2 \|\boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2}. \quad (65)$$

III.3.4 Simplification of $||\beta_k||^2$

For the unperturbed data $\mathbf{Y}$

$$\mathbf{Y} = \mathbf{A}\mathbf{S} = \mathbf{U}_s \Sigma_s \mathbf{V}_s \quad (66)$$

we define the estimate data covariance matrix $\hat{\mathbf{R}}$ as

$$\begin{aligned} \hat{\mathbf{R}} &= \mathbf{Y}\mathbf{Y}^H = \mathbf{A}\mathbf{S}\mathbf{S}^H\mathbf{A}^H \\ &= \mathbf{A}\hat{\mathbf{R}}_s\mathbf{A}^H = \mathbf{U}_s \Sigma_s^2 \mathbf{U}_s^H \end{aligned} \quad (67)$$

where $\hat{\mathbf{R}}_s \stackrel{\text{def}}{=} \mathbf{S}\mathbf{S}^H$ is the estimate signal covariance matrix. The generalized inverse of $\hat{\mathbf{R}}$ is

$$\hat{\mathbf{R}}^\dagger = \mathbf{A}^H \hat{\mathbf{R}}_s^{-1} \mathbf{A}^\dagger = \mathbf{U}_s \Sigma_s^{-2} \mathbf{U}_s^H. \quad (68)$$

Now we can get

$$\begin{aligned} \mathbf{V}_s^H \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{A} &= \mathbf{V}_s^H \Sigma_s \mathbf{U}_s^H \mathbf{U}_s \Sigma_s^{-2} \mathbf{U}_s^H \mathbf{A} = \mathbf{S}^H \mathbf{A}^H \hat{\mathbf{R}}_s^{-1} \mathbf{A} \\ &= \mathbf{S}^H \mathbf{A}^H \mathbf{A}^H \hat{\mathbf{R}}_s^{-1} \mathbf{A}^\dagger \mathbf{A} = \mathbf{S}^H \hat{\mathbf{R}}_s^{-1} \end{aligned} \quad (69)$$

and

$$(\mathbf{V}_s^H \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{A})^H (\mathbf{V}_s^H \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{A}) = \hat{\mathbf{R}}_s^{-1} \mathbf{S}\mathbf{S}^H \hat{\mathbf{R}}_s^{-1} = \hat{\mathbf{R}}_s^{-1} \mathbf{S}\mathbf{S}^H \hat{\mathbf{R}}_s^{-1} = \hat{\mathbf{R}}_s^{-1}. \quad (70)$$

Since β_k is the k -th column of $\mathbf{V}_s^H \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{A}$, $||\beta_k||^2 = \beta_k^H \beta_k$ is the k -th diagonal element of $\hat{\mathbf{R}}_s^{-1}$.

The relationship between the true and estimated covariance matrices of the source signals is

$$\mathbf{R}_s = E_s(\mathbf{S}\mathbf{S}^H) = \frac{1}{M} \sum_{m=1}^M E_s(\mathbf{s}_m \mathbf{s}_m^H) = \frac{1}{M} E_s(\hat{\mathbf{R}}_s) \quad (71)$$

here $E_s(\cdot)$ is the expectation with respect to the random variable (signal) s . If we assume the source signals are ergodic such that

$$\mathbf{R}_s = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M \mathbf{s}_m \mathbf{s}_m^H = \lim_{M \rightarrow \infty} \frac{1}{M} \hat{\mathbf{R}}_s.$$

We now define

$$\Delta \mathbf{R}_s = \frac{1}{M} \hat{\mathbf{R}}_s - \mathbf{R}_s \quad (72)$$

and

$$E_s(\Delta \mathbf{R}_s) = \frac{1}{M} E_s(\hat{\mathbf{R}}_s) - \mathbf{R}_s = \frac{1}{M} M \mathbf{R}_s - \mathbf{R}_s = 0.$$

Using (72), we can express the $\hat{\mathbf{R}}_s$ as

$$\hat{\mathbf{R}}_s = M \mathbf{R}_s (\mathbf{I} - \mathbf{R}_s^{-1} \Delta \mathbf{R}_s). \quad (73)$$

Assuming

$$\|\mathbf{R}_s^{-1} \Delta \mathbf{R}_s\| < 1 \quad (74)$$

then the $\hat{\mathbf{R}}_s^{-1}$ can be written as

$$\hat{\mathbf{R}}_s^{-1} = \frac{1}{M} \mathbf{R}_s^{-1} [\mathbf{I} + \mathbf{R}_s^{-1} \Delta \mathbf{R}_s - (\mathbf{R}_s^{-1} \Delta \mathbf{R}_s)^2 + \dots]. \quad (75)$$

If we take the expectation of (75)

$$\begin{aligned} E_s(\hat{\mathbf{R}}_s^{-1}) &= \frac{1}{M} \mathbf{R}_s^{-1} [\mathbf{I} + \mathbf{R}_s^{-1} E_s(\Delta \mathbf{R}_s) - E_s(\mathbf{R}_s^{-1} \Delta \mathbf{R}_s)^2 + \dots] \\ &= \frac{1}{M} \mathbf{R}_s^{-1} [\mathbf{I} - E_s(\mathbf{R}_s^{-1} \Delta \mathbf{R}_s)^2 + \dots]. \end{aligned} \quad (76)$$

Because of (74), we can drop the second- and higher- order terms of (76). So to the first-order

$$E_s(\hat{\mathbf{R}}_s^{-1}) \doteq \frac{1}{M} \mathbf{R}_s^{-1}. \quad (77)$$

III.3.5 Final Expression for MSE

Denoting $\mathbf{R}_s(k, k) = \mathbf{e}_k^H \mathbf{R}_s^{-1} \mathbf{e}_k$, we get our final expression

$$E(\Delta \theta_k)^2 = E_s\{E_{\Delta \mathbf{Y}}[(\Delta \theta_k)^2 | \mathbf{s}]\} = \frac{\mathbf{R}_s^{-1}(k, k) \sigma_n^2}{2M\gamma_k} \quad (78)$$

where γ_k is specified in Table II, and $E[\cdot]$ is conditional expectation.

Notice that since $\mathbf{\Omega}_o = \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ so that $\mathbf{\Omega}_o$ is an fundamental parameters, unlike $\mathbf{U}_o$ which is indirect parameter from SVD.

An alternative expression which has some advantages in comparison is

$$E(\Delta\theta_k)^2 = \frac{\mathbf{R}_s^{-1}(k, k) \sigma_n^2}{2M[\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o \mathbf{W} \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)]} \quad (79)$$

where weighting matrix $\mathbf{W}$ can be specified for different algorithms as in Table III.

Lemma III.2 Given $\|\mathbf{W}\| = 1$, $|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o \mathbf{W} \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)|$ reaches maximum when $\mathbf{W} = \mathbf{I}$. This means MUSIC has the least MSE compared with Min-Norm and Subspace-Rotation.

Proof: Using Cauchy-Schwarz inequality, we have

$$|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o \mathbf{W} \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)| \leq \|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o\| \cdot \|\mathbf{W} \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|. \quad (80)$$

Using submultiplicative property of matrix norm

$$\begin{aligned} & \|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o\| \cdot \|\mathbf{W} \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\| \\ & \leq \|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o\| \cdot \|\mathbf{W}\| \cdot \|\mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\| \\ & = |\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)| \cdot \|\mathbf{W}\| \end{aligned} \quad (81)$$

since $\|\mathbf{W}\| = 1$

$$|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)| \geq |\mathbf{a}^{(1)}(\theta_k)^H \mathbf{\Omega}_o \mathbf{W} \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)|. \quad (82)$$

□

Table III shows that $\|\mathbf{W}\| = 1$ for all the algorithms, this concludes that MUSIC has smaller mean-squared error than all other algorithms mentioned. This

result is a general combination of the previous results on comparison of MUSIC and Min-Norm [31, 12] and on comparison of MUSIC and ESPRIT [9].

III.4 OBSERVATIONS

It is obvious that (78) and Table II (or (79) and Table III) have significant advantages over (44) and Table I. For instance, the factors which contribute to mean-squared error are explicitly decoupled. The contribution of the source signals is in $\mathbf{R}_s^{-1}(k, k)$, the contribution of the number of snapshots is M , the contributions of the number of sensors and source separations are in γ_k . This advantage will enable us to study the relationship between the mean-squared error performance and various contributing factors as shown in what follows.

III.4.1 The Effect of Number of Snapshots

From (78) we can see that $\mathbf{R}_s^{-1}$ and the norm-square term in denominator are not the functions of M , the mean-square error of DOA estimation is inversely proportional to the number of snapshots. (78) also tells us that M contributes all algorithms equally.

III.4.2 The Effect of Signal Coherence

Source coherence always exist which degrades the performance of subspace based DOA estimation. The analysis proposed in this paper allow us to investigate the the effect of source coherence to mean-squared estimation error. Now consider the case of existing two coherent sources with the power of σ_{s1}^2 and σ_{s2}^2 , respectively, if we define the source correlation coefficient $\rho = \frac{E_s(s_1 s_2^*)}{\sigma_{s1} \sigma_{s2}}$, we have

$$\mathbf{R}_s = \begin{pmatrix} \sigma_{s1}^2 & \rho \sigma_{s1} \sigma_{s2} \\ \rho^* \sigma_{s1} \sigma_{s2} & \sigma_{s2}^2 \end{pmatrix}.$$

The inverse of $\mathbf{R}_s$ is

$$\mathbf{R}_s^{-1} = \frac{1}{\sigma_{s1}^2 \sigma_{s2}^2 (1 - |\rho|^2)} \begin{pmatrix} \sigma_{s2}^2 & -\rho \sigma_{s1} \sigma_{s2} \\ -\rho^* \sigma_{s1} \sigma_{s2} & \sigma_{s1}^2 \end{pmatrix}. \quad (83)$$

Substitute (83) into (78), we can get

$$\begin{aligned} E(\Delta\theta_k)^2 &= \frac{\sigma_n^2}{2M\sigma_{sk}^2(1 - |\rho|^2)\gamma_k} \\ &= \frac{1}{2M(\frac{\sigma_{sk}}{\sigma_n})^2(1 - |\rho|^2)\gamma_k} \quad \text{for } k = 1, 2 \end{aligned} \quad (84)$$

where $\frac{\sigma_{sk}}{\sigma_n}$ is the signal-to-noise ratio. We can see that the mean-squared error increases as the source correlation (coefficient ρ) increases. When ρ approaches to 1 (highly correlated), the mean-squared error increases very rapidly. When the signal is fully correlated *i.e.*, coherent ($\rho = 1$), the mean-squared error becomes infinity, in other words, the algorithm fails. Similar derivations can provide us with the mean-square error versus source correlation relationship for the cases existing more than two coherent sources. Other conclusion includes that SNR and signal coherence contribute to all algorithms equally.

III.4.3 The Effect of Source Separation

Using L'Hospital's rule, we can show that the polynomial of r_k

$$\gamma_k = \mathbf{a}^{(1)}(r_k)^H \mathbf{\Omega}_o \mathbf{W} \mathbf{\Omega}_o \mathbf{a}^{(1)}(r_k) \quad (85)$$

can be factored into

$$A \prod_{\substack{j \neq k \\ j=1}}^{L-1} (1 - r_k r_j^{-1})(1 - r_k^* r_j) \stackrel{\text{def}}{=} AG(r_k) \quad (86)$$

where $r_k = e^{j\frac{2\pi d}{\lambda_c} \sin \theta_k}$. Notice that $G(r_k)$ is a function of the separations between the k -th signal root r_k and all other roots r_j . When another signal root, say r_i is

very close to r_k , The mean-squared error will be dominantly affected by the angular separation of these two closely spaced roots. Let us look at the factor involved r_k and r_i of $G(r_k)$:

$$\begin{aligned} G(r_k) &= 2 - r_k r_i^{-1} - r_k^* r_i \\ &= 2 - 2 \cos\left[\frac{2\pi d}{\lambda_c}(\sin \theta_k - \sin \theta_i)\right] \end{aligned} \quad (87)$$

when the angular separation between the two roots is small, we get

$$\sin \theta_k - \sin \theta_i \doteq \cos \theta_k \Delta \theta_k$$

where

$$\Delta \theta_k = \theta_k - \theta_i.$$

Then this factor can be written as

$$2\left[1 - \cos\left(\frac{2\pi d}{\lambda_c} \cos \theta_k \Delta \theta_k\right)\right] \doteq \left(\frac{2\pi d}{\lambda_c} \cos \theta_k \Delta \theta_k\right)^2 \quad (88)$$

here we used the Taylor expansion of the cosine function and only keep up to the second-order terms. We can now show that

$$E(\theta_k)^2 = \bar{A} \frac{\mathbf{R}_s^{-1}(k, k) \sigma_n^2}{2M(1 - \rho^2) \bar{G}(r_k) (\cos \theta_k \Delta \theta_k)^2} \quad (89)$$

where

$$\bar{A} = A \left(\frac{2\pi d}{\lambda_c}\right)^2$$

and

$$\bar{G}(r_k) = \prod_{\substack{j \neq k, i \\ j=1}}^{L-1} (1 - r_k r_j^{-1})(1 - r_k^* r_j).$$

(89) reveals an important fact that mean-squared error of DOA estimation is very sensitive to the small angular separation of the closely spaced sources as it is reversely proportional to the square of the small angular separation.

III.4.4 The Effect of the Number of Sensors

We can see that in (78) only γ_k the function of the number of sensors L , so we can single γ_k out for the analysis of the effect of the number of sensors. We also notice γ_k is not an explicit function of L . L only contributes to the performance implicitly, in term of the number of the factors involved in (86). This fact makes the an analytical expression of sensor number effect rather difficult, if not impossible at all, to reach. In what follows, we will only able to show the monotonical property of mean-squared error versus number of sensor for MUSIC and non-overlapped ESPRIT. Other situations will be discussed numerically.

Lemma III.3 MUSIC *The mean-squared error of MUSIC DOA estimation monotonically decreases as the number of sensors increases.*

Proof: The mean-squared error is inversely proportional to γ_k

$$E(\theta_k)_L^2 = \frac{\mathbf{R}_s^{-1}(k, k) \sigma_n^2}{2M \mathbf{a}^{(1)}(\theta_k)_L^H \boldsymbol{\Omega}_{oL} \mathbf{a}^{(1)}(\theta_k)_L} \quad (90)$$

where subscript L stands for the case where an array of L sensors is used. Similarly we have

$$E(\theta_k)_{L-1}^2 = \frac{\mathbf{R}_s^{-1}(k, k) \sigma_n^2}{2M \mathbf{a}^{(1)}(\theta_k)_{L-1}^H \boldsymbol{\Omega}_{oL-1} \mathbf{a}^{(1)}(\theta_k)_{L-1}} \quad (91)$$

for the case an array of $L - 1$ sensors are used. Because of the orthogonality in $L - 1$ dimensions

$$\boldsymbol{\Omega}_{oL-1} \mathbf{A}_{L-1} = 0 \quad (92)$$

or equivalently

$$\mathbf{I}^1{}^H \boldsymbol{\Omega}_{oL-1} \mathbf{I}^1 \mathbf{A}_L = 0. \quad (93)$$

We can say $\mathbf{I}^H \boldsymbol{\Omega}_{oL-1} \mathbf{I}^\dagger$ is in the subspace of $\boldsymbol{\Omega}_{oL}$. From

$$\begin{aligned}
 & \mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_{oL} \mathbf{a}^{(1)}(\theta_k) - \mathbf{a}^{(1)}(\theta_k)^H \mathbf{I}^H \boldsymbol{\Omega}_{oL-1} \mathbf{I}^\dagger \mathbf{a}^{(1)}(\theta_k) \\
 &= \mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_{oL} (\mathbf{I} - \mathbf{I}^H \boldsymbol{\Omega}_{oL-1} \mathbf{I}^\dagger) \boldsymbol{\Omega}_{oL} \mathbf{a}^{(1)}(\theta_k) \\
 &\geq 0
 \end{aligned} \tag{94}$$

where the inequality is because

$$\mathbf{I} - \mathbf{I}^H \boldsymbol{\Omega}_{oL-1} \mathbf{I}^\dagger$$

is nonnegative-definite, thus we can show

$$\begin{aligned}
 \frac{E(\theta_k)_{L-1}^2}{E(\theta_k)_L^2} &= \frac{\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_{oL} \mathbf{a}^{(1)}(\theta_k)}{\mathbf{a}^{(1)}(\theta_k)^H \mathbf{I}^H \boldsymbol{\Omega}_{oL-1} \mathbf{I}^\dagger \mathbf{a}^{(1)}(\theta_k)} \\
 &\geq 1.
 \end{aligned} \tag{95}$$

So for MUSIC the mean-squared error monotonically decreases as the number of sensors L increases.

□

Lemma III.4 ESPRIT (Non-overlapping Subarray Case) *The mean-squared error of ESPRIT DOA estimation monotonically decreases as the number of sensors increases.*

Proof: For uniform line array of L sensors, we used the first and last $L - 1$ sensors as the two subarrays. These two subarray are maximally overlapped.

We defined $\mathbf{I}^\dagger$ and $\mathbf{I}^H$ as

$$\mathbf{I}^\dagger = \begin{pmatrix} \mathbf{I}_{L-1, L-1} & \mathbf{0}_{L-1, 1} \end{pmatrix}$$

and

$$\mathbf{I}^\dagger = \begin{pmatrix} \mathbf{0}_{L-1,1} & \mathbf{I}_{L-1,L-1} \end{pmatrix}$$

thus the α_k was expressed as

$$\alpha_k^H = \mathbf{e}_k^H (\mathbf{D}^{-1} \mathbf{A}^{\downarrow \dagger} \mathbf{I}^\dagger - \mathbf{A}^{\downarrow \dagger} \mathbf{I}^\dagger).$$

When the non-overlapped array is used, especially in the case of non-uniform array with sensor-doublets, we can easily modify the above result by defining $\mathbf{A}_F$ and $\mathbf{A}_S$ as the array manifold for the first- and second- subarray, and $\mathbf{I}_F$ and $\mathbf{I}_S$ as

$$\mathbf{I}_F = \begin{pmatrix} \mathbf{I}_{\frac{L}{2}, \frac{L}{2}} & \mathbf{0}_{\frac{L}{2}, \frac{L}{2}} \end{pmatrix}$$

and

$$\mathbf{I}_S = \begin{pmatrix} \mathbf{0}_{\frac{L}{2}, \frac{L}{2}} & \mathbf{I}_{\frac{L}{2}, \frac{L}{2}} \end{pmatrix}$$

then the α_k can be expressed as

$$\alpha_k^H = \mathbf{e}_k^H (\mathbf{D}^{-1} \mathbf{A}_F^\dagger \mathbf{I}_S - \mathbf{A}_F^\dagger \mathbf{I}_F).$$

Using the fact that $\mathbf{I}_F \mathbf{I}_S^H = \mathbf{0}$, we have

$$\begin{aligned} \alpha_k^H \alpha_k &= \mathbf{e}_k^H (\mathbf{D}^{-1} \mathbf{A}_F^\dagger \mathbf{A}_F^{\dagger H} \mathbf{D} + \mathbf{A}_F^\dagger \mathbf{A}_F^{\dagger H}) \mathbf{e}_k \\ &= 2\mathbf{R}_A^{-1}(k, k) \end{aligned} \tag{96}$$

where $\mathbf{R}_A = \mathbf{A}_F^H \mathbf{A}_F$ and we used the fact that

$$\mathbf{e}_k^H \mathbf{D}^{-1} \mathbf{A}_F^\dagger \mathbf{A}_F^{\dagger H} \mathbf{D} \mathbf{e}_k = \mathbf{e}_k^H \mathbf{A}_F^\dagger \mathbf{A}_F^{\dagger H} \mathbf{e}_k.$$

Now we use the subscript $\frac{L}{2}$ for dimension of the matrix, and partition

$$\mathbf{A}_{F \frac{L}{2}} = \begin{pmatrix} \mathbf{A}_{F \frac{L}{2}-1} \\ \mathbf{x} \end{pmatrix}$$

then using matrix inversion lemma

$$\mathbf{R}_{A\frac{L}{2}}^{-1} = (\mathbf{A}_F^H \mathbf{A}_{F\frac{L}{2}-1} + \mathbf{x}^H \mathbf{x})^{-1} = \mathbf{R}_{A\frac{L}{2}-1}^{-1} - \frac{\mathbf{R}_{A\frac{L}{2}-1}^{-1} \mathbf{x}^H \mathbf{x} \mathbf{R}_{A\frac{L}{2}-1}^{-1}}{1 + \mathbf{x} \mathbf{R}_{A\frac{L}{2}-1}^{-1} \mathbf{x}^H}.$$

Now we can get

$$\mathbf{R}_{A\frac{L}{2}}^{-1}(k, k) - \mathbf{R}_{A\frac{L}{2}-1}^{-1}(k, k) = -\mathbf{e}_k^H \frac{\mathbf{R}_{A\frac{L}{2}-1}^{-1} \mathbf{x}^H \mathbf{x} \mathbf{R}_{A\frac{L}{2}-1}^{-1}}{1 + \mathbf{x} \mathbf{R}_{A\frac{L}{2}-1}^{-1} \mathbf{x}^H} \mathbf{e}_k \leq 0$$

because $\mathbf{R}_{A\frac{L}{2}-1}^{-1} \mathbf{x}^H \mathbf{x} \mathbf{R}_{A\frac{L}{2}-1}^{-1}$ is non-negative definite.

$$\begin{aligned} \frac{E(\theta_k)_{L-1}^2}{E(\theta_k)_L^2} &= \frac{\mathbf{R}_{A\frac{L}{2}-1}^{-1}(k, k)}{\mathbf{R}_{A\frac{L}{2}}^{-1}(k, k)} \\ &\geq 1. \end{aligned} \tag{97}$$

So for Non-ESPRIT the mean-squared error monotonically decreases as the number of sensors L increases.

□

We have shown (and we are only able to show) the monotonical property of mean-squared error versus number of sensor for MUSIC and non-overlapped ESPRIT. The reason we failed to show the same property for Min-Norm and ESPRIT is not because the analyses of them are difficult (they are!), but because such property does not exist in Min-Norm and overlapped ESPRIT and SSR. This can be observed from our theoretical prediction using (78), the numerical result is in Figure 5. This is exactly the strength of performance analysis: one does not have to run hundreds or thousands trials of simulation to discovered a phenomenon.

III.5 NUMERICAL EXAMPLES

In this section, we will present numerically the statistical performance of mean-squared error of DOA estimation predicted by our analysis, and compare them with the extensive computer simulation.

The general configuration of the experiments is: an uniform line array of eight sensors (with $d = \lambda_c/2$) with two sources at 0.2 and 0.35 radians (angles are measured with respect to the normal of the array). The signals are $s_k(n) = e^{j(\frac{2\pi}{\lambda_c}n + \phi_{kn})}$, where ϕ_{kn} are independent random phase angles uniformly distributed in the interval $(-\pi, \pi)$. Twenty snapshots of array data were taken for 500 hundred trials.

In Figure 3, we show the root mean-squared error of DOA estimation versus signal-to-noise-ratio (SNR). SNR is defined as

$$SNR = 10 \log \frac{\sigma_s^2}{\sigma_n^2}.$$

In Figure 4, we show the root mean-squared error of DOA estimation versus source correlation. The correlation coefficient is define as

$$\rho = \frac{E_s(s_1 s_2^*)}{\sigma_{s1} \sigma_{s2}}.$$

In Figure 5, we show the root mean-squared error of DOA estimation versus source separation. As the minimal separation in Figure 2 is as small as 0.01 radians (about half degree), we use 40dB SNR. The first source is place at $\theta_1 = 0.2$ (radians) with respect to normal, and second source is place at

$$\theta_2 = \theta_1 + \Delta\theta.$$

In Figure 6, we show the root mean-squared error of DOA estimation versus number of snapshots.

In Figure 7, we show the root mean-squared error of DOA estimation versus number of sensors for MUSIC, Min-Norm, and overlapped ESPRIT. In Figure 8, we show the root mean-squared error of DOA estimation versus number of sensors for non-overlapping ESPRIT. We used a third source at $\theta_3 = .5$ (radians) and 30 snapshots. Figures 7 and 8 show that the mean-squared error monotonically decreases as the number of sensors increases. Figure 7 also shows the non-monotonical behavior of root mean-squared error of DOA estimation versus number of sensors for Min-Norm and overlapped ESPRIT.

III.6 SUMMARY

In this chapter, we provide an unified analysis for subspace-based DOA estimation algorithms by expressing the DOA perturbation in terms of physical data parameters such as numbers of sensors and snapshots, and source separations and coherence. Many important results have been drawn from our compact formula of the expectation of mean-squared DOA estimation error for different algorithms. The techniques developed here are quite general and should turn out to be useful to analyze performance characteristics of other subspace-based techniques.

TABLE I
PARAMETERS OF DOA ESTIMATION PERTURBATION

	α_k	β_k	γ_k
MUSIC	$j\mathbf{U}_o\mathbf{U}_o^H\mathbf{a}^{(1)}(\theta_k)$	$\mathbf{V}_s\Sigma_s^{-1}\mathbf{U}_s^H\mathbf{a}(\theta_k)$	$\mathbf{a}^{(1)}(\theta_k)^H\mathbf{U}_o\mathbf{U}_o^H\mathbf{a}^{(1)}(\theta_k)$
Min-Norm Searching	$j\mathbf{U}_o\bar{\mathbf{c}}\bar{\mathbf{c}}^H\mathbf{U}_o^H\mathbf{a}^{(1)}(\theta_k)$	$\mathbf{V}_s\Sigma_s^{-1}\mathbf{U}_s^H\mathbf{a}(\theta_k)$	$\mathbf{a}^{(1)}(\theta_k)^H\mathbf{U}_o\bar{\mathbf{c}}\bar{\mathbf{c}}^H\mathbf{U}_o^H\mathbf{a}^{(1)}(\theta_k)$
Min-Norm	$C_k\mathbf{U}_o\bar{\mathbf{c}}\bar{\mathbf{c}}^H\mathbf{U}_o^H\mathbf{a}^{(1)}(r_k)$	$\mathbf{V}_s\Sigma_s^{-1}\mathbf{U}_s^H\mathbf{a}(r_k)$	$Ar_k\prod_{j=1}^{L-1} (1-r_jr_k^{-1}) ^2$
Root-MUSIC	$C_k\mathbf{U}_o\mathbf{W}\mathbf{U}_o^H\mathbf{a}^{(1)}(r_k)$	$\mathbf{V}_s\Sigma_s^{-1}\mathbf{U}_s^H\mathbf{a}(r_k)$	$Ar_k\prod_{j=1}^{L-1} (1-r_jr_k^{-1}) ^2$
SSR	$C_k\mathbf{v}_k\Sigma_s^{-\frac{1}{2}}\mathbf{U}_s^{\dagger}(\mathbf{U}_o^{\dagger}\lambda_k^{-1}-\mathbf{U}_o^{\dagger})\mathbf{U}_o^H$	$\mathbf{V}_s\Sigma_s^{-\frac{1}{2}}\mathbf{u}_k$	1
ESPRIT	$C_k\mathbf{v}_k\mathbf{U}_{sz}^{\dagger}(\mathbf{U}_{ox}\lambda_k^{-1}-\mathbf{U}_{oz})\mathbf{U}_o^H$	$\mathbf{V}_s\Sigma_s^{-1}\mathbf{u}_k$	1

TABLE II
PARAMETERS OF DOA ESTIMATION ERROR

	α_k^H	β_k^H	γ_k
MUSIC	$-j\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o$	$\mathbf{e}_k^H \mathbf{R}^{-1} \mathbf{S}$	$ \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) ^2$
Min-Norm	$-j\mathbf{a}^{(1)}(\theta_k)^H \frac{\bar{\mathbf{e}}_1 \bar{\mathbf{e}}_1^H}{ \mathbf{e} \boldsymbol{\Omega}_o ^2}$	$\mathbf{e}_k^H \mathbf{R}^{-1} \mathbf{S}$	$ \alpha_k ^2 \mathbf{e} \boldsymbol{\Omega}_o ^2$
ESPRIT	$C_k \mathbf{e}_k^H \mathbf{A}^{\downarrow \dagger} (\mathbf{I}^{\downarrow} \lambda_k^{-1} - \mathbf{I}^{\downarrow})$	$\mathbf{e}_k^H \mathbf{R}^{-1} \mathbf{S}$	1

TABLE III
PARAMETERS OF ALTERNATIVE MSE EXPRESSION

	MUSIC	Min-Norm	Subspace-Rotation
$\mathbf{W}$	$\mathbf{I}$	$\bar{\mathbf{e}} \bar{\mathbf{e}}^H$	$\bar{\alpha}_k \bar{\alpha}_k^H$

Figure 3. Root Mean-Squared Error vs. SNR.

Figure 4. Root Mean-Squared Error vs. Source Correlation.

Figure 5. Root Mean-Squared Error vs. Source Separation.

Figure 6. Root Mean-Squared Error vs. Number of Snapshots.

Figure 7. Root Mean-Squared Error vs. Number of Sensors for MUSIC.

Figure 8. Root Mean-Squared Error vs. Number of Sensors for ESPRIT.

CHAPTER IV

BEAMSPACE ANALYSIS

IV.1 INTRODUCTION

The subspace-based DOA algorithms all involve eigenvalue decomposition or SVD. The computational load for the data from arrays containing large numbers of sensors is high, thus the subspace algorithms may be prohibitive. A common approach to this problem is to map the array data from full dimension (element-space) into lower dimension (beam-space) through *beamforming*. This reduces the computational complexity [32]. Additionally, beam-space processing reduces the sensitivity to wavefront distortion [33] and noise structure [34], improves robustness [35] and adaptive response time. Beam-space processing also increases the SNR resolution threshold for MUSIC [36, 37]. In situations where sensor measurements are not accessible (e.g. phase array radar) or where real-time processing is required, beam-space approximation is adequate for DOA estimation [38] [39]. These advantages inspired researchers to apply subspace algorithms in beam-space. However, the performance of beam-space processing remains largely unexplored except for some pioneering analyses in [36] for spectrum estimation error and in [37] for resolution threshold, both of which are limited to MUSIC.

In this chapter, we analyze beam-space performance of DOA estimation for several subspace processing algorithms in a unified fashion based on a finite set of measurements. The mean-squared errors (MSE) of DOA estimation for these algo-

rithms are expressed in a single formula using *physical parameters*. The important relationship between MSE and properties of beamformer is revealed, and optimality of beamforming is discussed.

IV.2 BEAMSPACE PROCESSING REVIEW

The beamspace data are beamformed element-space data

$$\mathbf{Y}_B = \mathbf{B}^H \mathbf{Y} \quad (98)$$

where the beamformer $\mathbf{B}$ is of dimension $L \times B$ and $B \leq L$. Without loss the generality, we assume that the column vectors of $\mathbf{B}$ are orthonormal

$$\mathbf{B}^H \mathbf{B} = \mathbf{I}. \quad (99)$$

After beamforming, the noisy data become

$$\begin{aligned} \tilde{\mathbf{Y}}_B &= \mathbf{B}^H \mathbf{A} \mathbf{S} + \mathbf{B}^H \mathbf{N} \\ &= \begin{pmatrix} \tilde{\mathbf{U}}_{sB} & \tilde{\mathbf{U}}_{oB} \end{pmatrix} \begin{pmatrix} \tilde{\Sigma}_{sB} & 0 \\ 0 & \tilde{\Sigma}_{oB} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{V}}_{sB}^H \\ \tilde{\mathbf{V}}_{oB}^H \end{pmatrix}. \end{aligned}$$

Generally, B is much less than L so that the array output is compressed from an $L \times M$ matrix to a $B \times M$. Thus the computational load involved in SVD or EVD is dramatically reduced.

For simplicity, we define the beamspace steering vector, array manifold, projection matrix and beamspace noise as

$$\begin{aligned} \mathbf{a}_B(\theta) &= \mathbf{B}^H \mathbf{a}(\theta) \\ \mathbf{A}_B &= \mathbf{B}^H \mathbf{A} \\ \Omega_{oB} &= \mathbf{A}_B \left(\mathbf{A}_B^H \mathbf{A}_B \right)^{-1} \mathbf{A}_B^H \\ \mathbf{N}_B &= \mathbf{B}^H \mathbf{N}. \end{aligned} \quad (100)$$

Similar to element-space, the orthogonality between $\mathbf{A}_B$ and $\mathbf{U}_{oB}$ (or $\mathbf{\Omega}_{oB}$) exists in beamspace, so that MUSIC and Min-Norm algorithms can be applied to either search the null of $\|\mathbf{a}_B(\theta)\mathbf{\Omega}_{oB}\|^2$ or root polynomial $\|\mathbf{a}(z_k)^H \mathbf{B} \mathbf{\Omega}_{oB}\|^2 = 0$; *unlike* for element-space, the Vandermonde structure does not exist for beamspace array manifold $\mathbf{A}_B$ (this implies that shift-invariance does not exist among the rows of $\mathbf{U}_{sB}$) so that TAM cannot be applied. However, ESPRIT is valid if we choose

$$\mathbf{B}_{B \times L} = \begin{pmatrix} \bar{\mathbf{B}} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{B}} \end{pmatrix} \text{ so } \mathbf{A}_B = \begin{pmatrix} \mathbf{A}_B \\ \mathbf{A}_B \mathbf{D} \end{pmatrix} \quad (101)$$

where the number of columns for $\bar{\mathbf{B}}$ is $\frac{L}{2}$ for non-overlapped subarrays, and is from $\frac{L}{2}$ to $L - 1$ for uniform line array, we consider only non-overlapped subarrays in this paper.

IV.3 BEAMSPACE PERFORMANCE

For MUSIC, Min-Norm, and ESPRIT, the derivation for relationship of DOA estimation error and observation noise are identical for both element-space and beamspace processing. Similar to (42) in element-space, in beamspace we have

$$\Delta\theta_{kB} = \frac{\text{Im}[\alpha_{kB}^H \mathbf{N}_B \beta_{kB}]}{\gamma_{kB}} = \frac{\text{Im}[\alpha_{kB}^H \mathbf{B}^H \mathbf{N} \beta_{kB}]}{\gamma_{kB}}. \quad (102)$$

Since $\mathbf{N}$ is white as we previously assumed, we have

$$E(\Delta\theta_{kB}^2) = \frac{\|\alpha_{kB}^H \mathbf{B}^H\|^2 \|\beta_{kB}\|^2 \sigma_n^2}{2\gamma_{kB}^2} = \frac{\|\alpha_{kB}\|^2 \|\beta_{kB}\|^2 \sigma_n^2}{2\gamma_{kB}^2} \quad (103)$$

where $\alpha_{kB}^H \mathbf{B}^H \mathbf{B} \alpha_{kB} = \alpha_{kB}^H \alpha_{kB}$ is because (99). We can see that (103) is identical to (44), so that the derivation in Chapter 3 applies to beamspace performance analysis which leads to the following MSE expression

$$E(\Delta\theta_{kB})^2 = \frac{\mathbf{R}_s^{-1}(k, k) \sigma_n^2}{2M[\mathbf{a}^{(1)}(\theta_k)_B^H \mathbf{\Omega}_{oB} \mathbf{W} \mathbf{\Omega}_{oB} \mathbf{a}^{(1)}(\theta_k)_B]} \quad (104)$$

where weighting matrix $\mathbf{W}$ can be specified for different algorithms as in Table IV and

$$\mathbf{e}_B = \frac{\Omega_{oB}\mathbf{e}_1}{\|\Omega_{oB}\mathbf{e}_1\|} \quad \text{and} \quad \bar{\alpha}_{Bk} = \frac{\alpha_{Bk}}{\|\alpha_{Bk}\|}.$$

For ESPRIT, we have

$$\alpha_k^H = C_k \mathbf{e}_k^H (\mathbf{D}^{-1} \mathbf{A}_B^\dagger [0_{B \times B} \quad \mathbf{I}_{B \times B}] - \mathbf{A}_B^\dagger [\mathbf{I}_{B \times B} \quad 0_{B \times B}])$$

so that

$$\alpha_k^H \alpha_k = C_k^2 \mathbf{e}_k^H (\mathbf{D}^{-1} \mathbf{A}_B^\dagger \mathbf{A}_B^\dagger \mathbf{D} + \mathbf{A}_B^\dagger \mathbf{A}_B^\dagger) \mathbf{e}_k \stackrel{\text{def}}{=} 2C_k^2 \mathbf{R}_{\mathbf{A}_B}^{-1}(k, k) \quad (105)$$

where

$$\mathbf{R}_{\mathbf{A}_B} \stackrel{\text{def}}{=} \mathbf{A}_B^H \mathbf{A}_B. \quad (106)$$

So for ESPRIT, the MSE is

$$E(\Delta\theta_{kB})^2 = \sigma_n^2 C_k^2 \mathbf{R}_{\mathbf{A}_B}^{-1}(k, k) \mathbf{R}_s^{-1}(k, k) \quad (107)$$

IV.4 PROPERTIES OF BEAMSPACE DOA ESTIMATION ALGORITHMS

To observe some important properties of beamspace processing, we assume to have two beamformers $\mathbf{B}_1$ (of dimension $L \times B_1$) and $\mathbf{B}_2$ (of dimension $L \times B_2$), and $B_1 \geq B_2$, then compare the performance of these two beamformers when the column space spanned by $\mathbf{B}_2$ is a subspace of the column space spanned by $\mathbf{B}_1$ (denoted as $\mathbf{B}_2 \subset \mathbf{B}_1$).

Lemma IV.1 (MUSIC): *If $\mathbf{B}_2 \subset \mathbf{B}_1$, then the MSE of DOA estimation using beamformer $\mathbf{B}_2$ will always be larger than or equal to that of using $\mathbf{B}_1$.*

Proof: Since $\mathbf{B}_2 \subset \mathbf{B}_1$, we can express

$$\mathbf{B}_2 = \mathbf{B}_1 \mathbf{T}$$

where the $B_1 \times B_2$ matrix $\mathbf{T}$ is assumed to be $\mathbf{T}^H \mathbf{T} = \mathbf{I}$. Then from (104) for MUSIC, we can get

$$\begin{aligned} \frac{E(\Delta\theta_k)_{\mathbf{B}_1}^2}{E(\Delta\theta_k)_{\mathbf{B}_2}^2} &= \frac{\|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_2 \boldsymbol{\Omega}_{o\mathbf{B}_1}\|^2}{\|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1}\|^2} \\ &= \frac{\|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \mathbf{T} \boldsymbol{\Omega}_{o\mathbf{B}_2}\|^2}{\|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1}\|^2}. \end{aligned} \quad (108)$$

Since

$$\mathbf{A}_{B_2}^H \boldsymbol{\Omega}_{oB_2} = \mathbf{A}^H \mathbf{B}_1 \mathbf{T} \boldsymbol{\Omega}_{oB_2} = 0 \quad (109)$$

$\mathbf{T} \boldsymbol{\Omega}_{oB_2}$ is in the subspace of $\boldsymbol{\Omega}_{o\mathbf{B}_1}$. So that

$$\mathbf{T} \boldsymbol{\Omega}_{oB_2} = \boldsymbol{\Omega}_{o\mathbf{B}_1}^H \mathbf{T} \boldsymbol{\Omega}_{oB_2} \quad (110)$$

then (108) becomes

$$\frac{E(\Delta\theta_{k\mathbf{B}_1})^2}{E(\Delta\theta_{k\mathbf{B}_2})^2} = \frac{\|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1} \mathbf{T} \boldsymbol{\Omega}_{o\mathbf{B}_2}\|^2}{\|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1}\|^2}. \quad (111)$$

We can show that

$$\begin{aligned} &\|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1}\|^2 - \|\mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1} \mathbf{T} \boldsymbol{\Omega}_{o\mathbf{B}_2}\|^2 \\ &= \mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1} \boldsymbol{\Omega}_{o\mathbf{B}_1}^H \mathbf{B}_1^H \mathbf{a}^{(1)}(\theta_k) \\ &\quad - \mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1} \mathbf{T} \boldsymbol{\Omega}_{o\mathbf{B}_2} \mathbf{T}^H \boldsymbol{\Omega}_{o\mathbf{B}_1}^H \mathbf{B}_1^H \mathbf{a}^{(1)}(\theta_k) \\ &= \mathbf{a}^{(1)}(\theta_k)^H \mathbf{B}_1 \boldsymbol{\Omega}_{o\mathbf{B}_1} (\mathbf{I} - \mathbf{T} \boldsymbol{\Omega}_{o\mathbf{B}_2} \mathbf{T}^H) \boldsymbol{\Omega}_{o\mathbf{B}_1}^H \mathbf{B}_1^H \mathbf{a}^{(1)}(\theta_k) \\ &\geq 0 \end{aligned} \quad (112)$$

where $\mathbf{T} \boldsymbol{\Omega}_{o\mathbf{B}_2} \mathbf{T}^H$ is a projection matrix and that $\mathbf{I} - \mathbf{T} \boldsymbol{\Omega}_{o\mathbf{B}_2} \mathbf{T}^H$ is nonnegative-definite matrix. We now have $E(\Delta\theta_{k\mathbf{B}_1})^2 \leq E(\Delta\theta_{k\mathbf{B}_2})^2$

Corollary IV.1 (MUSIC): *The MSE of DOA estimation in any beamspace will always be larger than or equal to that of element-space.*

Proof: Let $\mathbf{B}_1 = \mathbf{I}$ and $\mathbf{B}_2 = \mathbf{B}$, apply Lemma IV.1.

□

Lemma IV.2 (ESPRIT): *If $\bar{\mathbf{B}}_2 \subset \bar{\mathbf{B}}_1$, then the MSE of DOA estimation using beamspace $\bar{\mathbf{B}}_2$ will always be larger than or equal to that of using $\bar{\mathbf{B}}_1$.*

Proof: Without loss the generality, we assume two subarray beamformers $\bar{\mathbf{B}}_1$ and $\bar{\mathbf{B}}_2$ with relationship of their dimensions $B_2 = B_1 - 1$. If $\bar{\mathbf{B}}_2 \subset \bar{\mathbf{B}}_1$, then $\bar{\mathbf{B}}_2$ and $\bar{\mathbf{B}}_1$ can be related as

$$\bar{\mathbf{B}}_2 = \bar{\mathbf{B}}_1 \mathbf{T}_{(B_1 \times B_2)} \quad \text{and} \quad \mathbf{A}_{\bar{\mathbf{B}}_2} = \mathbf{T}_{(B_2 \times B_1)}^H \mathbf{A}_{\bar{\mathbf{B}}_1}$$

where $\mathbf{T}^H \mathbf{T} = \mathbf{I}$. By properly selecting a vector $\mathbf{b}$ of dimension $B_1 \times 1$, we can construct a $B_1 \times B_1$ full rank unitary matrix $(\mathbf{T} \ \mathbf{b})$. From (107) we can see for different beamformers, the only factor contributes to MSE is $\mathbf{R}_{\mathbf{A}_{\bar{\mathbf{B}}}}^{-1}(k, k)$.

$$\begin{aligned} \mathbf{R}_{\mathbf{A}_{\bar{\mathbf{B}}_1}}^{-1} &= (\mathbf{A}_{\bar{\mathbf{B}}_1}^H \mathbf{A}_{\bar{\mathbf{B}}_1})^{-1} = \left(\mathbf{A}_{\bar{\mathbf{B}}_1}^H (\mathbf{T} \ \mathbf{b}) \begin{pmatrix} \mathbf{T}^H \\ \mathbf{b}^H \end{pmatrix} \mathbf{A}_{\bar{\mathbf{B}}_1} \right)^{-1} \\ &= (\mathbf{A}_{\bar{\mathbf{B}}_1}^H \mathbf{T} \mathbf{T}^H \mathbf{A}_{\bar{\mathbf{B}}_1} + \mathbf{A}_{\bar{\mathbf{B}}_1}^H \mathbf{b} \mathbf{b}^H \mathbf{A}_{\bar{\mathbf{B}}_1})^{-1} \end{aligned} \quad (113)$$

notice that

$$\begin{aligned} \mathbf{A}_{\bar{\mathbf{B}}_1}^H \mathbf{T} \mathbf{T}^H \mathbf{A}_{\bar{\mathbf{B}}_1} &= \mathbf{A}^H \bar{\mathbf{B}}_1 \mathbf{T} \mathbf{T}^H \bar{\mathbf{B}}_1^H \mathbf{A} \\ &= \mathbf{A}^H \bar{\mathbf{B}}_2 \bar{\mathbf{B}}_2^H \mathbf{A} = \mathbf{A}_{\bar{\mathbf{B}}_2}^H \mathbf{A}_{\bar{\mathbf{B}}_2} = \mathbf{R}_{\mathbf{A}_{\bar{\mathbf{B}}_2}} \end{aligned} \quad (114)$$

and define $\mathbf{v} = \mathbf{b}^H \mathbf{A}_{\bar{\mathbf{B}}_1}$, then (113) becomes

$$\mathbf{R}_{\mathbf{A}_{\bar{\mathbf{B}}_1}}^{-1} = (\mathbf{R}_{\mathbf{A}_{\bar{\mathbf{B}}_2}} + \mathbf{v} \mathbf{v}^H)^{-1}. \quad (115)$$

Using the matrix lemma, we have

$$\mathbf{R}_{\mathbf{A}_{\mathbf{B}_1}}^{-1} = \mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1} - \frac{\mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1} \mathbf{v} \mathbf{v}^H \mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1}}{1 + \mathbf{v}^H \mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1} \mathbf{v}} \quad (116)$$

so that we can obtain

$$\mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1}(k, k) - \mathbf{R}_{\mathbf{A}_{\mathbf{B}_1}}^{-1}(k, k) = \left[\frac{\mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1} \mathbf{v} \mathbf{v}^H \mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1}}{1 + \mathbf{v}^H \mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1} \mathbf{v}} \right] (k, k) \geq 0. \quad (117)$$

The last step is because $\mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1} \mathbf{v} \mathbf{v}^H \mathbf{R}_{\mathbf{A}_{\mathbf{B}_2}}^{-1}$ is nonnegative-definite. This inequality gives us $E(\Delta\theta_{k\bar{\mathbf{B}}_1})^2 \leq E(\Delta\theta_{k\bar{\mathbf{B}}_2})^2$.

□

Corollary IV.2 (ESPRIT): *The MSE of DOA estimation in any beamspace will always be larger than or equal to that of element-space.*

Proof: Let $\bar{\mathbf{B}}_1 = \mathbf{I}$ and $\bar{\mathbf{B}}_2 = \bar{\mathbf{B}}$, apply Lemma IV.2.

□

Remarks (Min-Norm): We have shown the monotonic property of MSE versus dimension of beamspace for MUSIC and ESPRIT when the beamspace associated with $\mathbf{B}_2$ is in the subspace of the beamspace associated with $\mathbf{B}_1$. This monotonic property does not exist for Min-Norm. This can be predicted from our theoretical expression using (104).

IV.5 OPTIMAL BEAMFORMERS

The Optimal Beamformers discussed here are in the sense of least MSE.

Lemma IV.3 (ESPRIT): *In beamspace, ESPRIT attains its minimum MSE of DOA estimation when $\bar{\mathbf{B}} \subset \mathbf{A}_s$.*

Proof: Let $\bar{\mathbf{B}}_1 = \mathbf{I}$ and $\bar{\mathbf{B}}_2 = \bar{\mathbf{B}}$, the equality of (117) holds when when $\mathbf{v} = \mathbf{b}^H \mathbf{A}_s = 0$ since $\mathbf{b}\mathbf{b}^H = \mathbf{I} - \bar{\mathbf{B}}^H \bar{\mathbf{B}}$, so equivalently we have

$$\mathbf{A}_s^H \bar{\mathbf{B}} \bar{\mathbf{B}}^H = \mathbf{A}_s^H.$$

□

The minimum MSE beamformer for ESPRIT can expressed as

$$\bar{\mathbf{B}} = \mathbf{Q}(\mathbf{Q}^H \mathbf{Q})^{-\frac{1}{2}} \quad \text{with} \quad \mathbf{Q} = [\mathbf{A}_s \ \cdots] \quad (118)$$

Lemma (MUSIC): *In beamspace, MUSIC attains its minimum MSE of DOA estimation when $\mathbf{B}$ span $[\mathbf{A}, \mathbf{A}^{(1)}]$.*

Proof: Let $\mathbf{B}_1 = \mathbf{I}$ and $\mathbf{B}_2 = \mathbf{B}$, then the numerator of (111) becomes

$$\begin{aligned} & \|\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{B} \boldsymbol{\Omega}_o \mathbf{B}\|^2 \\ &= \mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{B} \boldsymbol{\Omega}_o \mathbf{B}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) \\ &= \mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{B} \mathbf{B}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) \\ &- \mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{B} \boldsymbol{\Omega}_s \mathbf{B}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) \\ &\leq \|\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o\|^2 - \|\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{B} \boldsymbol{\Omega}_s \mathbf{B}\|^2 \\ &\leq \|\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o\|^2. \end{aligned} \quad (119)$$

The equality $\|\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{B} \boldsymbol{\Omega}_o \mathbf{B}\|^2 = \|\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o\|^2$ hold if and only if

$$(i) \quad \mathbf{B} \mathbf{B}^H \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)$$

$$(ii) \quad \mathbf{\Omega}_s \mathbf{B} \mathbf{B}^H \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = 0$$

Here we show the condition (ii) is automatically met given the condition (i):

Assuming $\mathbf{B} \mathbf{B}^H \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)$, we will have

$$\mathbf{A}^H \mathbf{B} \mathbf{B}^H \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = \mathbf{A}^H \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = 0$$

since $\mathbf{A}^H \mathbf{B}$ and $\mathbf{\Omega}_s \mathbf{B}$ share the same row space, so that

$$\mathbf{\Omega}_s \mathbf{B} \mathbf{B}^H \mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = 0$$

Now we can expressed

$$\mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = \mathbf{a}^{(1)}(\theta_k) - \sum_{i=1}^P \mathbf{u}_i \mathbf{u}_i^H \mathbf{a}^{(1)}(\theta_k) \quad (120)$$

where $\mathbf{u}_i$'s are the signal singular vectors. So $\mathbf{\Omega}_o \mathbf{a}^{(1)}(\theta_k)$ is the linear combination of $[\mathbf{a}^{(1)}(\theta_k), \mathbf{u}_1, \dots, \mathbf{u}_P]$ or, $[\mathbf{a}^{(1)}(\theta_k), \mathbf{A}]$. To meet the condition (i), The span of $\mathbf{B} \mathbf{B}^H$ has to include span of $[\mathbf{A} \mathbf{A}^{(1)}]$.

□

The minimum MSE beamformer for MUSIC can expressed as

$$\mathbf{B} = \mathbf{Q}(\mathbf{Q}^H \mathbf{Q})^{-\frac{1}{2}} \quad \text{with} \quad \mathbf{Q} = [\mathbf{A} \mathbf{A}^{(1)} \dots] \quad (121)$$

It should be noticed that this selection of MUSIC beamformer for minimum MSE agrees with the selection for best resolution [37].

IV.6 NUMERICAL EXAMPLES

The configuration of the experiments is: a uniform line array of ten sensors (with $d = \lambda_c/2$) with three sources at 0.2, 0.35, and 0.5 radians (angles are measured

with respect to the normal of the array). Twenty snapshots of array data were taken for 500 hundred trials. When ESPRIT is used, another array of ten sensor is added in as the second subarray. We use the prolate spheroidal sequences beamformer as defined in [35]. The direction sector is selected at $[0, 0.7]$ radians.

Figure 9 shows the root MSE of DOA estimation versus SNR when the dimension of the beamformer is 6.

Figure 10 shows the root MSE of DOA estimation versus the dimensions of beamformer. When we use the prolate spheroidal sequences beamformer, $\mathbf{B}_2 \subset \mathbf{B}_1$ when $B_2 < B_1$. We can see the monotonic property for MUSIC and ESPRIT.

In Figure 11, we use a two sources configuration with DOAs at 0.2 and 0.35 radians to illustrate the non-monotonic property using certain beamformer. The beamformer we use here is the same as used by [40]

$$\mathbf{B} = (\mathbf{a}(\theta_{ini}), \mathbf{a}(\theta_{ini} + \Delta\theta), \dots, \mathbf{a}(\theta_{end}))$$

where $[\theta_{ini}, \theta_{end}]$ is the sector of interest and

$$\Delta\theta = \frac{\mathbf{a}(\theta_{en}) - \mathbf{a}(\theta_{in})}{Q - 1}$$

Using this beamformer, the span of the beamformer doesn't increase with the increase of dimension of the beamformer. So we can observe the monotonic property in the root MSE of DOA estimation versus beamformer dimension figure.

IV.7 SUMMARY

In this chapter, the performance of beamspace DOA estimation algorithms are examined in a unified fashion. Compact formula for the expectation of mean-squared DOA estimation error has been derived and expressed using *physical* pa-

rameters. Important properties of beamformer and insights on beamformer design have been offered.

TABLE IV
PARAMETERS OF BEAMSPACE MSE

	MUSIC	Min-Norm	ESPRIT
W	I	$\bar{\mathbf{e}}_B \bar{\mathbf{e}}_B^H$	$\bar{\alpha}_{Bk} \bar{\alpha}_{Bk}^H$

Figure 9. Root MSE vs. SNR.

Figure 10. Root MSE vs. Dimension of Beamformer 1.

Figure 11. Root MSE vs. Dimension of Beamformer 2.

CHAPTER V

FORWARD-BACKWARD AVERAGING

V.1 Introduction

Most subspace-based DOA estimation algorithms use a standard sensors versus snapshots (or forward-only) data formulation. The Min-Norm algorithm uses a forward-backward averaged data. This forward-backward averaging was observed experimentally in the past not only reducing the mean-squared DOA estimation error, but also reducing the correlation between two interfered signals. One question to be answered is why the other methods do not also use forward-backward averaging?

These questions have attracted significant effort in performance analyses (see [24] for references). Among others [17, 41, 42, 43] addressed the performance due to the effect of forward-backward averaging.

In this chapter, we apply our unified analysis to the effects of forward-backward averaging to the statistical performance of DOA estimation algorithms. Results on the manner in which forward-backward averaging decorrelate sources correlation and their effects on different algorithms are provided.

V.2 PROBLEM FORMULATION

The forward-backward averaging utilizes the data as the follows

$$\mathbf{Y}_{FB} = \left(\begin{array}{ccc|ccc} y_1(1) & \cdots & y_1(M) & | & y_L(1)^* & \cdots & y_L(M)^* \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ y_L(1) & \cdots & y_L(M) & | & y_1(1)^* & \cdots & y_1(M)^* \end{array} \right) \quad (122)$$

the subscript FB stands for Forward-Backward, or, in matrix form

$$\mathbf{Y}_{FB} = [\mathbf{Y} \ \mathbf{J}\mathbf{Y}^*] = \mathbf{A}[\mathbf{S} \ \mathbf{D}^{-L+1}\mathbf{S}^*] \stackrel{\text{def}}{=} \mathbf{A}\mathbf{S}_{FB} \quad (123)$$

where $\mathbf{J}$ is the permutation matrix.

The estimate covariance data matrix is

$$\mathbf{R}_{FB} = \mathbf{Y}_{FB}\mathbf{Y}_{FB}^H = \mathbf{C} + \mathbf{J}\mathbf{C}^*\mathbf{J} = \mathbf{A}\mathbf{R}_{sFB}\mathbf{A}^H \quad (124)$$

where estimated signal covariance matrix is

$$\mathbf{R}_{sFB} = (\mathbf{S}\mathbf{S}^H + \mathbf{D}^{-L+1}\mathbf{S}^*\mathbf{S}^H\mathbf{D}^{L-1}) = \mathbf{R} + \mathbf{D}^{-L+1}\mathbf{R}^*\mathbf{D}^{L-1}. \quad (125)$$

V.3 STATISTICAL EFFECT OF FORWARD-BACKWARD AVERAGING

In the noisy case, the forward-backward averaged data can be written as

$$\hat{\mathbf{Y}}_{FB} = \mathbf{Y}_{FB} + \mathbf{N}_{FB} = \mathbf{A}\mathbf{S}_{FB} + \mathbf{N}_{FB}$$

We can simply replace the forward-only parameters in (42) by forward-backward averaged parameters.

$$\Delta\theta_k = \frac{\text{Im}[\beta_{FBk}^H \mathbf{N}_{FB}^H \alpha_k]}{\gamma_k} \quad (126)$$

where

$$\beta_{FBk}^H = \mathbf{e}_k \mathbf{R}_{FB}^{-1} \mathbf{S}_{FB}$$

here $||\beta_{FBk}||^2 \stackrel{\text{def}}{=} \mathbf{R}_{FB}^{-1}(k, k)$ and $\mathbf{N}_{FB} = (\mathbf{N} \quad \mathbf{J}\mathbf{N}^*)$. But α_k is the same as in forward-only data formulation because it's only a function of array manifold $\mathbf{A}$.

V.3.1 MSE Expression

By long but straightforward derivation in Appendix B, we get the expression of MSE for forward-backward averaging as

$$E(\Delta\theta_k)^2 = \frac{\sigma_n^2[\mathbf{R}_{FB}^{-1} + 2\text{Re}(\phi_k \bar{\mathbf{R}}_{FB}^{-1})](k, k)}{2[\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{W} \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)]} \quad (127)$$

where we define

$$\phi_k = -z_k^{L-1} \frac{(\alpha_k^H \mathbf{J} \alpha_k^*)}{||\alpha_k||^2}. \quad (128)$$

and

$$\bar{\mathbf{R}}_{FB}^{-1} = \mathbf{R}_{FB}^{-1*} \mathbf{R}^* \mathbf{R}_{FB}^{-1}. \quad (129)$$

Using Schwartz Inequality

$$||-z_k^{L-1}(\alpha_k^H \mathbf{J} \alpha_k^*)|| \leq ||-z_k^{L-1}|| \cdot ||\alpha_k|| \cdot ||\mathbf{J}|| \cdot ||\alpha_k|| = ||\alpha_k||^2$$

it is easy to see $||\phi_k|| \leq 1$.

V.3.2 Properties

Property V.1 *For MUSIC*, $||\phi_k|| = 1$.

Proof: Using the properties

$$\mathbf{A} = \mathbf{J} \mathbf{A}^* \mathbf{D}^{L-1} \quad \text{and} \quad \mathbf{A}^\dagger = \mathbf{D}^{1-L} \mathbf{A}^{*\dagger} \mathbf{J} \quad (130)$$

we can get

$$\begin{aligned} \boldsymbol{\Omega}_o &= \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \\ &= \mathbf{I} - \mathbf{J} \mathbf{A}^* \mathbf{D}^{L-1} (\mathbf{D}^{1-L} \mathbf{A}^T \mathbf{J} \mathbf{J} \mathbf{A}^* \mathbf{D}^{L-1})^{-1} \mathbf{D}^{1-L} \mathbf{A}^T \mathbf{J} \\ &= \mathbf{I} - \mathbf{J} \mathbf{A}^* (\mathbf{A}^T \mathbf{A}^*)^{-1} \mathbf{A}^T \mathbf{J} = \mathbf{J} \boldsymbol{\Omega}_o^* \mathbf{J} \end{aligned} \quad (131)$$

or $\mathbf{J}\boldsymbol{\Omega}_o^* = \boldsymbol{\Omega}_o\mathbf{J}$ so that we can rewrite

$$\begin{aligned} -\mathbf{J}\alpha_k^* z_k^{L-1} &= j\mathbf{J}\boldsymbol{\Omega}_o^* \mathbf{a}^{(1)}(\theta_k)^* z_k^{L-1} \\ &= j\boldsymbol{\Omega}_o\mathbf{J}\mathbf{a}^{(1)}(\theta_k)^* z_k^{L-1}. \end{aligned} \quad (132)$$

Since

$$\begin{aligned} &\mathbf{a}^{(1)}(\theta_k) - \mathbf{J}\mathbf{a}^{(1)}(\theta_k)^* z_k^{L-1} \\ &= \begin{pmatrix} 0 \\ C_k z_k \\ \vdots \\ (L-1)C_k z_k \end{pmatrix} - \begin{pmatrix} -(L-1)C_k z_k \\ \vdots \\ -C_k z_k \\ 0 \end{pmatrix} z_k \\ &= (L-1)C_k \mathbf{a}(\theta_k) \end{aligned} \quad (133)$$

pre-multiply (133) by $\boldsymbol{\Omega}_o$, and notice that $\boldsymbol{\Omega}_o \mathbf{a}(\theta_k) = 0$, we have

$$j\boldsymbol{\Omega}_o\mathbf{J}\mathbf{a}^{(1)}(\theta_k)^* z_k^{L-1} = j\boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k) = \alpha_k \quad (134)$$

so that

$$-\alpha_k^H \mathbf{J}\alpha_k^* z_k^{L-1} = ||\alpha_k||^2 \quad (135)$$

$\phi_k = 1$ for MUSIC is proved.

□

Property V.2 *For Subspace-Rotation algorithms, $||\phi_k|| = 1$.*

Proof:

$$\alpha_k^H = \mathbf{e}_k \mathbf{A}^{\downarrow \dagger} (\mathbf{I}^{\uparrow} z_k^{-1} - \mathbf{I}^{\downarrow})$$

so that

$$\begin{aligned}
 z_k^{L-1} \alpha_k^H \mathbf{J} \alpha_k^* &= z_k^{L-1} \mathbf{e}_k \mathbf{A}^{\downarrow \dagger} (\mathbf{I}^\dagger z_k^{-1} - \mathbf{I}^\downarrow) \mathbf{J} \alpha_k^* \\
 &= z_k^{L-1} \mathbf{e}_k \mathbf{A}^{\downarrow \dagger} \mathbf{J} (\mathbf{I}^\downarrow z_k^{-1} - \mathbf{I}^\dagger) \alpha_k^*
 \end{aligned} \tag{136}$$

where we used $\mathbf{I}^\downarrow \mathbf{J} = \mathbf{J} \mathbf{I}^\dagger$ and $\mathbf{I}^\dagger \mathbf{J} = \mathbf{J} \mathbf{I}^\downarrow$, further

$$\begin{aligned}
 -z_k^{L-1} \alpha_k^H \mathbf{J} \alpha_k^* &= \mathbf{e}_k (z_k^{L-2} \mathbf{A}^{\downarrow \dagger} \mathbf{J}) (\mathbf{I}^\dagger z_k - \mathbf{I}^\downarrow) \alpha_k^* \\
 &= \mathbf{e}_k (\mathbf{A}^{\downarrow \dagger})^* (\mathbf{I}^\dagger z_k - \mathbf{I}^\downarrow) \alpha_k^* \\
 &= \alpha_k^T \alpha_k^* = \|\alpha_k\|^2
 \end{aligned} \tag{137}$$

where $z_k^{L-2} \mathbf{A}^{\downarrow \dagger} \mathbf{J} = (\mathbf{A}^{\downarrow \dagger})^*$ is because of (130). $\phi_k = 1$ for SSR/ESPRIT is proved.

□

ϕ_k for different algorithms can be found in Table V.

V.3.3 Uncorrelated Signals

When the signals are uncorrelated, $\mathbf{R}$ is diagonal such that (125) becomes

$$\mathbf{R}_{FB} = \mathbf{R} + \mathbf{D}^{-L+1} \mathbf{R}^* \mathbf{D}^{L-1} = 2\mathbf{R} \tag{138}$$

so that we have

$$\mathbf{R}_{FB}^{-1} = \frac{1}{2} \mathbf{R}^{-1} \tag{139}$$

and

$$\bar{\mathbf{R}}_{FB}^{-1} = \mathbf{R}_{FB}^{-1*} \mathbf{R}^* \mathbf{R}_{FB}^{-1} = \frac{1}{4} \mathbf{R}^{-1} \tag{140}$$

where $\mathbf{R}^* = \mathbf{R}$ for uncorrelated signals, and $\bar{\mathbf{R}}_{FB}^{-1}$ is real and diagonal, so that (127) becomes

$$E(\Delta\theta_k)^2 = \frac{\sigma_n^2[1 + \text{Re}(\phi_k)][\mathbf{R}^{-1}](k, k)}{4[\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{W} \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)]} \quad (141)$$

oindent Lemma V.1 *For uncorrelated signals, MUSIC and Subspace-Rotation have the same MSE using forward-only and forward-backward averaged data.*

Proof: For MUSIC and Subspace Rotation algorithms, $\phi_k = 1$ such that

$$E(\Delta\theta_k)^2 = \frac{\sigma_n^2 ||\alpha_k||^2}{2\gamma_k^2} \mathbf{R}^{-1}(k, k). \quad (142)$$

□

Equation (142) is the same as forward-only data formulations. Forward-backward averaging brings no improvement to uncorrelated signals when use MUSIC and SSR/ESPRIT algorithms.

Lemma V.2 *For uncorrelated signals, Min-Norm has smaller MSE using forward-backward averaged data than using forward-only data.*

Proof: For Min-Norm algorithm $\text{Re}(\phi_k) \leq ||\phi_k|| \leq 1$

$$\begin{aligned} E(\Delta\theta_k)^2 &= \frac{\sigma_n^2[1 + \text{Re}(\phi_k)][\mathbf{R}^{-1}](k, k)}{4[\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{W} \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)]} \\ &\leq \frac{\sigma_n^2[\mathbf{R}^{-1}](k, k)}{2[\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{W} \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)]} \end{aligned} \quad (143)$$

□

This means forward-backward averaging does improve performance for uncorrelated signals over forward-only formulation when use Min-Norm algorithm.

V.3.4 Correlated Signals

For correlated signals, as state in [41, 44]. The forward-backward averaging has the effect to decorrelate the correlation between signal when the number of correlated sources is 2. Similar to [45, 44], we have, in the case of two correlated sources with correlation coefficient ρ . Notice here, for finite data analysis, the ρ is the equivalent correlation coefficient of sample correlation coefficient. With the increase of the snapshot, it will goes to its expected value.

$$\begin{aligned}
 \mathbf{R}_{FB} &= \mathbf{R}_s + \mathbf{D}^{-L-1} \mathbf{R}_s^* \mathbf{D}^{L-1} \\
 &\approx M \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho \\ \sigma_1 \sigma_2 \rho^* & \sigma_2^2 \end{pmatrix} \\
 &+ M \mathbf{D}^{-L+1} \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho^* \\ \sigma_1 \sigma_2 \rho & \sigma_2^2 \end{pmatrix} \mathbf{D}^{L-1} \\
 &= 2M \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \frac{\rho + \rho^* z_1^{1-L} z_2^{L-1}}{2} \\ \sigma_1 \sigma_2 \frac{\rho^* + \rho z_1^{L-1} z_2^{1-L}}{2} & \sigma_2^2 \end{pmatrix} \\
 &\stackrel{\text{def}}{=} 2M \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \bar{\rho} \\ \sigma_1 \sigma_2 \bar{\rho}^* & \sigma_2^2 \end{pmatrix}. \tag{144}
 \end{aligned}$$

We can see, after forward-backward averaging, the equivalent correlation coefficient goes to $\bar{\rho}$ and

$$|\bar{\rho}|^2 = \frac{|\rho|^2 + \text{Re}(\rho^2 z_1^{1-L} z_2^{L-1})}{2} \leq |\rho|^2 \tag{145}$$

so that the sources are decorrelated and the performance will be improved. It is interesting to notice that $\bar{\rho}$ depends the angular separation of the signals. When ρ is real, the maximum decorrelation occurs when two sources are closely spaced.

It has been experimentally observed that $\mathbf{R}_{FB}^{-1}(k, k) \leq \frac{1}{2} \mathbf{R}^{-1}(k, k)$, so that it can be shown

$$E(\Delta \theta_k)_{FB}^2 \leq \frac{\sigma_n^2 \|\alpha_k\|^2}{\gamma_k^2} \mathbf{R}_{FB}^{-1}(k, k)$$

$$\leq \frac{\sigma_n^2 \|\alpha_k\|^2}{2\gamma_k^2} \mathbf{R}^{-1}(k, k) = E(\Delta\theta_k)_F^2 \quad (146)$$

where we have used

$$\begin{aligned} E[Im(\beta^H \mathbf{N}_{FB}^H \alpha_k)]^2 &\leq E \|\beta^H \mathbf{N}_{FB}^H \alpha_k\|^2 \\ &= \sigma_n^2 \|\alpha_k\|^2 \mathbf{R}_{FB}^{-1}(k, k) \end{aligned} \quad (147)$$

(146) tells us that the mean-squared error of using forward-backward averaged data will not be larger than that of using forward-only data for all subspace-based DOA estimation algorithms regardless of the coherence of the signals.

V.4 NUMERICAL EXAMPLES

We use an uniform line array of eight sensors (with $d = \lambda_c/2$) with two sources at 0.2 and 0.35 radians (angles are measured with respect to the normal of the array). The signals are $s_k(n) = e^{j(\frac{2\pi}{\lambda_c}n + \phi_{kn})}$, where ϕ_{kn} are independent random phase angles uniformly distributed in the interval $(-\pi, \pi)$. Twenty snapshots of array data were taken for 500 hundred trials.

Figure 12, 13 and 14 show the root mean-squared error of DOA estimation versus source correlation using forward-only and forward-backward data formulation. The correlation coefficient is define as

$$\rho = \frac{E_s(s_1 s_2^*)}{\sigma_{s1} \sigma_{s2}}.$$

We can see that the forward-backward averaging improves the performance of Min-Norm even for uncorrelated sources, however, it only improves the performance of MUSIC and Subspace-Rotation for correlated signals.

V.5 SUMMARY

In this chapter, we examine the effect of forward-backward averaging on the performance of DOA estimation algorithms. Unified formula for mean-squared DOA estimation error has been derived. The reason why for uncorrelated signals, forward-backward averaging improves the performance of Min-Norm but not for MUSIC and Subspace-Rotation has been revealed. We further conclude that, in general, forward-backward averaging is preferable for all subspace-based DOA estimation algorithms.

TABLE V
 ϕ_k FOR DIFFERENT ALGORITHMS

MUSIC	Min-Norm	Subspace-Rotation
$\phi_k = 1$	$ \phi_k \leq 1$	$\phi_k = 1$

Figure 12. Forward-only and Forward-backward Min-Norm.

Figure 13. Forward-only and Forward-backward MUSIC.

Figure 14. Forward-only and Forward-backward Subspace-Rotation.

CHAPTER VI

SPATIAL SMOOTHING

VI.1 INTRODUCTION

From the previous discussion, we can see that, when multiple signals are linear dependent, the subspace algorithms have difficulty in determining the *DOA* because of the collapse of the dimensionality of the so-called signal subspace. Evans *et al* [46] and later Shan *et al* [47] proposed a spatial smoothing method for a uniform linear array to restore the dimensionality of the signal subspace. A modified scheme, uses both forward and backward subarrays for averaging which required less sensors was proposed by Evans *et al* [46] and Pillai [41]. Since then, a number of papers appeared dealing with performance analysis of the spatial smoothing algorithm [44, 48, 49, 50, 43, 51, 52]. Some of them analyzed only one algorithm [44, 50, 43, 51] while the other studied only bias and resolution [44]. In this chapter, we examine the effect of using data both in forward and backward manner, along with spatial smoothing, on the performance on DOA estimation algorithms. A analysis of the forward only scheme is carried out first, then, the formula for forward-backward scheme is derived using some of the results in Chapter IV on the forward-backward averaging. Several observations on the performance improvement effect of spatial smoothing are provided.

VI.2 FORWARD ONLY SPATIAL SMOOTHING

VI.2.1 Data Formulation

Assume that a uniform linear array of L sensors is divided into overlapping subarrays of size Q (Figure 15). The total number of subarrays is $L - Q + 1$. Each subarray gives a covariance matrix and the forward spatially smoothed covariance matrix is obtained by averaging of these sub-covariance matrix:

$$\mathbf{R}_F = \sum_{i=1}^{L-Q+1} \mathbf{Y}_i \mathbf{Y}_i^H = \sum_{i=1}^{L-Q+1} \mathbf{A} \mathbf{D}^{i-1} \mathbf{S} \mathbf{S}^H \mathbf{D}^{i-1H} \mathbf{A}^H \quad (148)$$

where $\mathbf{A}$ is the subarray manifold of size $Q \times P$ and $\mathbf{Y}_i$ is the data output of the subarray:

$$\mathbf{Y}_i \stackrel{\text{def}}{=} \begin{pmatrix} y_{i,1} & y_{i,2} & \cdots & y_{i,M} \\ \vdots & \vdots & \ddots & \vdots \\ y_{i+Q-1,1} & y_{i+Q-1,2} & \cdots & y_{i+Q-1,M} \end{pmatrix} = \mathbf{A} \mathbf{D}^{i-1} \mathbf{S}. \quad (149)$$

Accordingly, we define the forward spatially smoothed sources covariance matrix as

$$\mathbf{R}_{sF} \stackrel{\text{def}}{=} \sum_{i=1}^{L-Q+1} \mathbf{D}^{i-1} \mathbf{S} \mathbf{S}^H \mathbf{D}^{i-1H}. \quad (150)$$

For the convenience of analysis, a equivalent direct data expression which gives the same covariance matrix can be defined as

$$\begin{aligned} \mathbf{Y}_F &= [\mathbf{Y}_1 \ \mathbf{Y}_2 \ \dots \ \mathbf{Y}_{L-Q+1}] \\ &\stackrel{\text{def}}{=} \mathbf{A} [\mathbf{S} \ \mathbf{D} \mathbf{S} \ \dots \ \mathbf{D}^{L-Q} \mathbf{S}] \stackrel{\text{def}}{=} \mathbf{A} \mathbf{S}_F \end{aligned} \quad (151)$$

on which we can apply the singular value decomposition and use subspace algorithms for DOA estimation. A equivalent noise matrix can be formed in the same fashion

$$\mathbf{N}_F = [\mathbf{N}_1 \ \mathbf{N}_2 \ \dots \ \mathbf{N}_{L-Q+1}]$$

VI.2.2 Performance Analysis

To start our performance analysis, we rewrite the noise corrupted data in the form of

$$\tilde{\mathbf{Y}}_F = \mathbf{Y}_F + \mathbf{N}_F.$$

Base on the perturbation theory derived in Chapter III, we simply substitute the parameters in equation (42) with their spatially smoothed counterparts and get the bias of DOA estimation

$$\Delta\theta_k = \frac{\Im[\beta_k^H \mathbf{N}_F^H \alpha_k]}{\gamma_k} \quad (152)$$

where α_k , γ_k are defined in Table II. They do not change with the spatial smoothing since they are only functions of subarray manifold $\mathbf{A}$. However β_k is now

$$\beta_k^H = \mathbf{e}_k^H \mathbf{R}_{sF}^{-1} \mathbf{S}_F \quad (153)$$

owing to the spatial smoothing.

The mean-squared error of the angular estimation is given by

$$E(\Delta\theta_k)^2 = \frac{E[\Im(\beta_k^H \mathbf{N}_F^H \alpha_k)]^2}{\gamma_k^2}. \quad (154)$$

Expanding the numerator of above equation, we have

$$\begin{aligned} & E[\Im(\beta_k^H \mathbf{N}_F^H \alpha_k)]^2 \\ &= \frac{1}{2} E(\beta_k^H \mathbf{N}_F \alpha_k)^2 - \frac{1}{4} E(\beta_k^H \mathbf{N}_F^H \alpha_k \beta_k^H \mathbf{N}_F^H \alpha_k) - \frac{1}{4} E(\alpha_k^H \mathbf{N}_F \beta_k \alpha_k^H \mathbf{N}_F \beta_k). \end{aligned} \quad (155)$$

Take the expectation with respect to noise, the second and third term go to zero because of the circular property of the noise. But

$$E[\mathbf{N}_F^H \alpha_k \alpha_k^H \mathbf{N}_F] \neq \sigma_n^2 \mathbf{I} ||\alpha_k||^2$$

because of the spatial smoothing. By noticing the above property, we derive the MSE expression for forward only spatial smoothing in Appendix C Here is the final formula for that:

$$E(\Delta\theta_k)^2 = \frac{\sigma_n^2}{2} \frac{\mathbf{R}_{sF}^{-1} \mathbf{R}_W \mathbf{R}_{sF}^{-1}(k, k)}{\gamma_k^2} \quad (156)$$

$\mathbf{R}_W$ is the *Weighted* signal correlation matrix defined in Appendix C. Again, the express only involves *physical* parameters. The effect of spatial smoothing on performance can be seen directly from this formula. Without spatial smoothing, the MSE is proportional to $\mathbf{R}_s^{-1}(k, k)$ which becomes infinity when the signal are coherent. All subspace-based algorithms will fail to resolve the DOA due to the dimensionality collapse of the signal subspace. After applying the spatial smoothing to the data matrix, $\mathbf{R}_s^{-1}(k, k)$ is substituted by *smoothed* signal covariance matrix $\mathbf{R}_{sF}^{-1}(k, k)$ which is no longer rank deficient as shown in [47]. This enable those subspace-based DOA estimation algorithm to estimate the DOAs with *finite* MSE.

VI.3 FORWARD-BACKWARD SPATIAL SMOOTHING

VI.3.1 Data Formulation

Similar to forward-backward averaging, in forward-backward spatial smoothing, the covariance matrix is obtained by averaging the forward and backward subarray covariance matrix:

$$\mathbf{R}_{FB} = \mathbf{R}_F + \mathbf{J} \mathbf{R}_F^* \mathbf{J}. \quad (157)$$

Accordingly, We define the forward-backward spatially smoothed noise and signal data matrix as

$$\begin{aligned} \mathbf{N}_{FB} &= [\mathbf{N}_F \mathbf{J} \mathbf{N}_F^*] \\ \mathbf{S}_{FB} &= [\mathbf{S}_F \mathbf{D}^{1-L} \mathbf{S}_F^*]. \end{aligned}$$

The equivalent direct data expression for perturbed data is given by:

$$\tilde{\mathbf{Y}}_{FB} = \mathbf{A}\mathbf{S}_{FB} + \mathbf{N}_{FB}. \quad (158)$$

It is shown in [51] that, in general, for all subspace-based methods, forward-backward is preferable to forward only spatial smoothing. We will develop the formula for the performance of forward-backward spatial smoothing and provide reason to support the above argument.

VI.3.2 Performance Analysis

The formulas for the estimation of bias and mean-squared error in (42) and (43) hold except that β_k and $\mathbf{N}$ should be substituted by their forward-backward spatially smoothed counterparts.

The derivation of the explicit expression of MSE can be found in Appendix D. We here only give the final formula for the MSE of DOA estimation using forward-backward spatial smoothing.

$$E(\Delta\theta_k)^2 = \frac{\sigma_n^2}{2\gamma_k^2} (\mathbf{R}_{sFB}^{-1} \mathbf{R}_{WFB} \mathbf{R}_{sFB}^{-1} - \Re(\mathbf{R}_{sFB}^{-1} {}^*\overline{\mathbf{R}}_{WFB} \mathbf{R}_{sFB}^{-1}))(k, k) \quad (159)$$

where $\mathbf{R}_{WFB}$ is forward-backward averaged *Weighted* signal covariance matrix

$$\mathbf{R}_{WFB} \stackrel{\text{def}}{=} \mathbf{R}_W + \mathbf{D}^{1-L} \mathbf{R}_W^* \mathbf{D}^{L-1}$$

and $\overline{\mathbf{R}}_{WFB}$ is defined in Appendix D, which is also a *Weighted* signal covariance matrix due to the backward spatial smoothing.

VI.4 EFFECT OF SPATIAL SMOOTHING ON PERFORMANCE

In this section, we analyze the effect of spatial smoothing on the performance of subspace algorithms. We first have an overall discussion on spatial smoothing and

then compare the forward only and forward-backward schemes for uncorrelated and highly correlated sources, respectively.

VI.4.1 Discussion

Spatial smoothing decorrelates the sources at the cost of reducing of effective array aperture. It has been proven in our previous work that the MSE is inversely proportional to the array aperture in element space (without spatial smoothing). Spatial smoothing actually has a negative effect on the performance of subspace algorithms for lowly correlated sources (Figure 16). On the other hand, spatial smoothing is a must for all the subspace-based algorithms when the signals are coherent.

At this point in our discussion, it is natural to ask that how can we reach the best performance by using spatial smoothing for different source correlation?

Due to the involvement of many other measurement conditions, it is hard to isolate the tradeoff effect between array aperture and source decorrelation. An analytic answer to the above question is still expected. However, by analyzing our theoretical plot outs for different source correlation, we can roughly conclude that, for MUSIC and ESPRIT, minimal spatial smoothing is desirable because of the dominant effect of the size of aperture on the performance of the algorithm. On the other hand, for Min-Norm, the optimal is reached when the spatial smoothing make the aperture $\frac{3}{4}$ of the maximum aperture [17]. At this point, the difference of the minimum mean-squared error between MUSIC and Min-Norm reaches its minima.

VI.4.2 Uncorrelated Signals

For forward only spatial smoothing, we can simplify the expression of the

mean-squared error. Since

$$\mathbf{R}_W = \sum_{i=1}^M \bar{\mathbf{S}}_i \mathbf{B} \mathbf{B}^H \bar{\mathbf{S}}_i^H.$$

When the number of snapshot is large, it can be approximated by its expectation value

$$\mathbf{R}_W \doteq M \bar{\mathbf{S}} \mathbf{B} \mathbf{B}^H \bar{\mathbf{S}}^H. \quad (160)$$

Because of independence between source signals,

$$\bar{\mathbf{S}} \mathbf{B} \mathbf{B}^H \bar{\mathbf{S}}^H \doteq \begin{pmatrix} \mathbf{a}_{L-Q}(\theta_1)^H \mathbf{B} \mathbf{B}^H \mathbf{a}_{L-Q}(\theta_1) \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \mathbf{a}_{L-Q}(\theta_P)^H \mathbf{B} \mathbf{B}^H \mathbf{a}_{L-Q}(\theta_P) \sigma_P^2 \end{pmatrix} \quad (161)$$

where $\sigma_1^2, \dots, \sigma_P^2$ are signal powers and $\mathbf{a}_{L-Q}(\theta_i)$ is the steering vector with aperture $L - Q$ (not Q , which is the aperture of subarray).

$$\mathbf{a}_{L-Q}(\theta_i) = \begin{pmatrix} 1 & \lambda_i & \dots & \lambda_i^{L-Q} \end{pmatrix}^T.$$

The $\mathbf{R}_s$ in equation (156) can also be approximated by its expectation value.

$$\mathbf{R}_s^{-1}(i, i) \doteq \frac{1}{M(L - Q + 1)\sigma_i^2} \quad (162)$$

so the simplified expression of (156) is given by

$$E(\Delta\theta_k)^2 = \frac{\|\mathbf{a}_{L-Q}(\theta_i)^H \mathbf{B}\|^2 \sigma_n^2}{2M(L - Q + 1)^2 \gamma_k^2 \sigma_k^2}. \quad (163)$$

For forward-backward spatial smoothing, we will prove the following lemma similar to that in forward-backward averaging discussed in Chapter V.

Lemma VI.1 For uncorrelated signals, MUSIC and Subspace Rotation algorithms have the same MSE using forward-backward spatial smoothed and forward only spatial smoothed data.

Proof: Notice that the sources are uncorrelated, from equation (161) (162), $\mathbf{R}_W$ and $\mathbf{R}_s$ are diagonal. So the first term in Expression (159) is approximately

$$\mathbf{R}_{sFB}^{-1} \mathbf{R}_{WFB} \mathbf{R}_{sFB}^{-1} \doteq \frac{1}{2} \mathbf{R}_s^{-1} \mathbf{R}_W \mathbf{R}_s^{-1}. \quad (164)$$

For the second term in (159), take the approximation of $\mathbf{R}_{sFB}^{-1}$ yields.

$$\begin{aligned} & \mathbf{R}_{sFB}^{-1} {}^* \overline{\mathbf{R}}_{WFB} \mathbf{R}_{sFB}^{-1} {}^H (k, k) \\ & \doteq \frac{1}{4M^2(L-Q+1)^2} \overline{\mathbf{R}}_{WFB}(k, k). \end{aligned} \quad (165)$$

Now from the definition of $\overline{\mathbf{R}}_{WFB}$ and using Property V.1, V.2, we get

$$\begin{aligned} & \overline{\mathbf{R}}_{WFB}(k, k) \\ & \doteq \sum_{i=1}^M \overline{\mathbf{S}}_i z_k^{L-1} \overline{\mathbf{B}}^* \mathbf{B}^H \overline{\mathbf{S}}_i^H + \overline{\mathbf{S}}_i^* \mathbf{B}^* \overline{\mathbf{B}}^H z_k^{L-1} \overline{\mathbf{S}}_i^T(k, k) \\ & = \sum_{i=1}^M \overline{\mathbf{S}}_i \mathbf{B} \mathbf{B}^H \overline{\mathbf{S}}_i^H + \overline{\mathbf{S}}_i^* \mathbf{B} \mathbf{B}^H \overline{\mathbf{S}}_i^T(k, k) \\ & \doteq 2\mathbf{R}_W. \end{aligned} \quad (166)$$

Combine (164) (165) and (166), the expression for MSE using MUSIC and Subspace Rotation for uncorrelated sources becomes

$$E(\Delta\theta_k)^2 = \frac{||\mathbf{a}_{L-Q}(\theta_i)^H \mathbf{B}||^2 \sigma_n^2}{2M(L-Q+1)^2 \gamma_k^2 \sigma_k^2} \quad (167)$$

which is the same as using forward only spatial smoothing.

□

VI.4.3 Correlation Signals

For correlated signals, both $\mathbf{R}_{sF}$, $\mathbf{R}_{sFB}$ and $\mathbf{R}_W$ are functions of signal correlation coefficients. We will discuss a simple two sources case to illustrate the decorrelation effect of spatial smoothing.

or forward only spatial smoothing, let σ_1^2 , σ_2^2 be the sources powers and ρ the correlation coefficient between two sources, the expectation of the forward only spatially smoothed sources covariance matrix is

$$\begin{aligned}\mathbf{R}_{sF} &= M \sum_{i=0}^{L-Q} \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho (\lambda_1 \lambda_2^{-1})^i \\ \sigma_1 \sigma_2 \rho^* (\lambda_2 \lambda_1^{-1})^i & \sigma_2^2 \end{pmatrix} \\ &= M(L-Q+1) \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \bar{\rho} \\ \sigma_1 \sigma_2 \bar{\rho}^* & \sigma_2^2 \end{pmatrix}\end{aligned}\quad (168)$$

where

$$\bar{\rho} = \frac{1 - (\lambda_1 \lambda_2^{-1})^{L-Q+1}}{(1 - \lambda_1 \lambda_2^{-1})(L-Q+1)} \rho$$

is the equivalent sources correlation coefficient after smoothing. It is obvious that

$$\frac{|\bar{\rho}|^2}{|\rho|^2} \leq 1$$

which means spatial smoothing can decorrelate sources correlation, and recover the rank of sources covariance matrix for the case of coherent signals.

For forward-backward spatial smoothing,

$$\begin{aligned}\mathbf{R}_{sFB} &= \mathbf{R}_{Fs} + \mathbf{D}^{1-L} \mathbf{R}_{Fs}^* \mathbf{D}^{L-1} \\ &= 2M(L-Q+1) \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \hat{\rho} \\ \sigma_1 \sigma_2 \hat{\rho}^* & \sigma_2^2 \end{pmatrix}\end{aligned}\quad (169)$$

where

$$\hat{\rho} = \frac{\bar{\rho} + \bar{\rho}^* (\lambda_1 \lambda_2)^{L-1}}{2}$$

is the equivalent sources correlation coefficient after forward-backward spatial smoothing.

We can relate the magnitude-square of equivalent correlation coefficient for forward only and forward-backward spatial smoothing by

$$|\hat{\rho}|^2 = \frac{|\bar{\rho}|^2 + \Re(\bar{\rho}\lambda_1^{1-L}\lambda_2^{L-1})}{2} \leq |\bar{\rho}|^2.$$

So forward-backward spatial smoothing has better effect of decorrelate the sources correlation, than forward only spatial smoothing.

VI.5 NUMERICAL EXAMPLES

We present some numerical examples to illustrate effect of spatial smoothing on the performance and further support our theoretical results. A twenty-element uniform linear array is used. The sources are equal power with covariance matrix given by

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad (170)$$

and DOA = 0.2, 0.35 radians, SNR = 20 dB, Number of snapshots = 20. All the simulations were conducted using 500 trials. The other configuration of data structure are the same as in Chapter III. In all the figures, only the root MSE for DOA at 0.35 radian is plotted out.

For forward only spatial smoothing, Figure 16, 17 and 18 show the Root MSE of DOA estimation vs. the aperture of subarray with source correlation $\rho = 0, 0.75$ and 1, respectively. We can easily see that, when the source correlation approaches 1, spatial smoothing becomes a must for subspace-based algorithms.

Figure 19, 20 and 21 show the same results for the forward-backward spatial smoothing scheme. Compare these figures to the figures for forward only spatial

smoothing, we can see that generally forward-backward spatial smoothing outperforms forward only spatial smoothing.

For independent sources, Figure 22, 23 and 24 compare the performance of forward only and forward-backward smoothing for MUSIC, Min-Norm and ESPRIT, respectively. They verify our conclusion in Lemma VI.1.

VI.6 SUMMARY

A unified analysis of different subspace-based DOA estimation algorithms implemented on spatially smoothed data was presented in this chapter. The forward only spatial smoothing scheme is studied first. The analysis was then extended to forward-backward spatial smoothing scheme. Important insights into the effect of spatial smoothing on the performance of DOA estimation algorithms was provided.

Figure 15. The Forward-backward Spatial Smoothing Scheme

Figure 16. Root Mean-Squared Error vs. Subarray Aperture.

Figure 17. Root Mean-Squared Error vs. Subarray Aperture.

Figure 18. Root Mean-Squared Error vs. Subarray Aperture.

Figure 19. Root Mean-Squared Error vs. Subarray Aperture.

Figure 20. Root Mean-Squared Error vs. Subarray Aperture.

Figure 21. Root Mean-Squared Error vs. Subarray Aperture.

Figure 22. Forward Only vs. Forward-backward Spatial Smoothing.

Figure 23. Forward Only vs. Forward-backward Spatial Smoothing.

Figure 24. Forward Only vs. Forward-backward Spatial Smoothing.

CHAPTER VII

CONCLUSION

We have presented a unified performance analysis of DOA estimation algorithms using physical parameters. These algorithms include MUSIC, Min-Norm, and Subspace Rotation. The analysis is applicable to these algorithms implemented on element space, beam space, forward-backward averaged and spatially smoothed data. Theoretical expressions for the estimation error are obtained by linking the statistical fluctuations of these quantities with singular vectors of signal and orthogonal subspaces, and then in turn with the DOA estimates themselves. Major results include the following

- provide unified expressions for statistics of DOA estimation of subspace-based algorithms implemented on different data structures.
- express all the statistics in terms of signal, noise and array characters (*physical* parameters) only. The relations that were previously masked by complex algebra are revealed.
- Draw and prove several important results from our simple formulas.

Simulation results were presented to demonstrate the validity of MSE expressions. In all the simulation examples, there was excellent agreement between the predicted and simulated DOA estimation errors.

REFERENCES

- [1] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Further unification, simplification, and observations," *to appear in IEEE Trans. Aerosp., Electron. Syst.*, vol. AES-29, October 1993.
- [2] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms using physical parameters," in *IEEE Proc. ICASSP'92*, (San Francisco, CA), pp. II-537-II-540, March 1992.
- [3] H. Liu and F. Li, "Statistical evaluation of beam-space direction-of-arrival estimators," in *IEEE Proc. ICASSP'92*, (San Francisco, CA), pp. V-389-V-392, March 1992.
- [4] F. Li and H. Liu, "Statistical analysis of of Beam-space estimation for Direction-Of-Arrival," *submitted to by IEEE Trans. Signal Processing*, 1992.
- [5] F. Li and H. Liu, "The effect of forward-backward data formulation to direction-of-arrival estimation," in *Proc. 25th Asilomar Conf. on Signals, Systems and Computers*, (Pacific Grove, CA), pp. 923-927, November 1991.
- [6] F. Li and H. Liu, "A high-resolution approach to simultaneously estimate seismic stacking velocity and zero-offset time," in *Proc. 6th IEEE-SP Workshop on Statistical Signal & Array Processing*, (Victoria, BC, Canada), October 1992.
- [7] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. RADC Spectral Estimation Workshop*, (Griffiss AFB, NY), pp. 243-258, 1979.
- [8] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerosp., Electron. Syst.*, vol. AES-19, pp. 134-139, January 1983.
- [9] P. Stoica and A. Nehorai, "Performance comparison of subspace rotation and MUSIC methods for direction estimation," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-39, pp. 446-453, February 1991.

- [10] S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao, "State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem," *J. Opt. Soc. Am.*, vol. 73, pp. 1799–1811, December 1983.
- [11] A. Paulraj, R. Roy, and T. Kailath, "Estimation of signal parameters via rotational invariance techniques – ESPRIT," in *Proc. 19th Asilomar Conf. on Signals, Systems and Computers*, (Pacific Grove, CA), pp. 83–89, November 1985.
- [12] F. Li and R. J. Vaccaro, "Analysis of MUSIC and Min-Norm for arbitrary array geometry," *IEEE Trans. Aerosp., Electron. Syst.*, vol. AES-26, pp. 976–985, November 1990.
- [13] F. Li, R. J. Vaccaro, and D. W. Tufts, "Min-Norm Linear Prediction for arbitrary sensor array," in *Proc. IEEE ICASSP'89*, (Glasgow, UK), pp. 2613–2616, May 1989.
- [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. Acoustics, Speech, and Signal Processing*, vol. ASSP-34, pp. 331–340, April 1986.
- [15] B. Porat and B. Friedlander, "Analysis of the asymptotic relative efficiency of the MUSIC algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-36, pp. 532–543, April 1988.
- [16] H. Clergeot, S. Tressens, and A. Ouamri, "Performance of high resolution frequencies estimation methods compared to the Cramer-Rao bounds," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 1703–1720, November 1989.
- [17] Y. Hua and T. K. Sarkar, "Perturbation analysis of TK method for harmonic retrieval problems," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-36, pp. 228–240, February 1988.
- [18] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood and Cramer-Rao bound," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 720–741, May 1989.

- [19] D. R. Farrier, D. J. Jeffries, and R. Mardani, "Theoretical performance prediction of the MUSIC algorithm," *IEE Proc., Pt. F*, vol. 135, pp. 216–225, June 1988.
- [20] D. W. Tufts, R. J. Vaccaro, and A. C. Kot, "Analysis of estimation of signal parameters by linear prediction at high SNR using matrix approximation," in *Proc. IEEE ICASSP'89*, (Glasgow, UK), May 1989.
- [21] D. V. Bhaskar Rao and K. V. S. Hari, "Performance analysis of Root-MUSIC," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 1939–1949, December 1989.
- [22] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model error - Part I: the MUSIC algorithm," *IEEE Trans on Signal Processing*, vol. SP-40, pp. 1758–1775, July 1992.
- [23] F. Li and R. J. Vaccaro, "Performance analysis of the State Space Realization (TAM) and ESPRIT algorithms for DOA estimation," *IEEE Trans. Antenna and Propagation*, vol. AP-39, pp. 418–423, March 1991.
- [24] F. Li and R. J. Vaccaro, *SVD and Signal Processing, II*, ch. Analytical Performance Prediction of the Subspace-Based Algorithms for DOA Estimation, pp. 243–260. Elsevier Science Publishers, North Holland, 1991.
- [25] F. Li, *A Unified Performance Analysis of Subspace-Based DOA Estimation Algorithms*. PhD thesis. University of Rhode Island, Kingston, RI, 1990.
- [26] F. Li and R. J. Vaccaro, "Sensitivity analysis of DOA estimation algorithms to sensor errors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-28, pp. 708–717, July 1992.
- [27] F. Li and R. J. Vaccaro, "Unified analysis for DOA estimation algorithms in array signal processing," *Signal Processing*, vol. 22, pp. 147–169, November 1991.
- [28] G. W. Stewart, "Error and perturbation bounds for subspaces associated with certain eigenvalue problems," *SIAM Review*, vol. 15-33, pp. 727–764, 1973.
- [29] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. New York, NY: Academic Press, second ed., 1978.

- [30] D. V. Bhaskar 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. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 1990–1995, December 1989.
- [31] F. Li and R. J. Vaccaro, "A unified performance analysis of subspace-based polynomial-rooting algorithms for DOA estimation," in *Proc. 1990 Conf. Information Sciences and Systems*, (Princeton, NJ), March 1990.
- [32] B. Veen and B. Williams, "Structures covariance matrices and dimensionally reduction in array processing," in *IEEE Fourth Annual ASSP Workshop on Spectral Estimation & Modeling*, (Minnesota, MN), pp. 168–171, August 1988.
- [33] G. Biennu and L. Kopp, "Decreasing high-resolution method sensitivity by conventional beamformer preprocessing," in *Proc. IEEE ICASSP'84*, 1984.
- [34] C. Byrne and A. Steele, "Sector-focused stability for high-resolution array processing," in *Proc. IEEE ICASSP'87*, (Dallas, TX), pp. 54.11.1–4, April 1987.
- [35] P. Forster and G. Vezzosi, "Application of spheroidal sequences to array processing," in *Proc. IEEE ICASSP'87*, (Dallas, TX), pp. 2267–2271, April 1987.
- [36] X.-L. Xu and K. M. Buckley, "Statistical performance comparison of MUSIC in Element-Space and Beam-Space," in *Proc. IEEE ICASSP'89*, (Glasgow, UK), pp. 2124–2127, May 1989.
- [37] H. R. Lee and M. S. Wengrovitz, "Resolution threshold of beam-space MUSIC for two closely spaced emitters," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-38, pp. 1545–1559, September 1990.
- [38] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Simultaneous sector processing via ROOT-MUSIC for large sensor arrays," in *Proc. 5th IEEE ASSP Workshop on Spectrum Estimation and Modeling*, (Rochester, NY), pp. 372–376, October 1990.
- [39] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Beam-space Root MUSIC," *Submitted to IEEE Trans. on SP*, 1991.
- [40] B. Friedlander, "Direction finding using an interpolated array," in *Proc. IEEE ICASSP'90*, (Albuquerque, NM), pp. 2951–2954, April 1990.

- [41] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 8–15, January 1989.
- [42] D. V. Bhaskar Rao and K. V. S. Hari, "Effect of spatial smoothing on the performance of MUSIC and the Minimum-Norm method," *IEE Proc., Pt. F*, pp. 449–458, December 1990.
- [43] D. V. Bhaskar Rao and K. V. S. Hari, "Effect of Spatial Smoothing on State-Space methods/ESPRIT," in *Proc. 5th IEEE ASSP Workshop on Spectrum Estimation and Modeling*, (Rochester, NY), pp. 377–381, October 1990.
- [44] S. U. Pillai and B. H. Kwon, "Performance analysis of MUSIC-type high resolution estimation for direction finding in correlated and coherent scenes," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 1176–1189, August 1989.
- [45] V.U.Reddy, A.Paulraj, and T.Kailath, "Performance analysis of the optimum beamformer in the presence of correlated sources and its behavior under spatial smoothing," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-35, pp. 927–936, July 1987.
- [46] J.E.Evans, J.R.Johnson, and D.F.Sun, "Applications of advanced signal processing techniques to angle of arrival estimation in ATC navigation and surveillance system," *M.I.T. Lincoln Lab.. Lexington. MA*, 1982.
- [47] T. J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation for coherent signals," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-33, pp. 806–811, August 1985.
- [48] D. V. Bhaskar Rao and K. V. S. Hari, "Effect of spatial smoothing on the performance on noise subspace based methods," in *Proc. IEEE ICASSP'90*, pp. 2687–2690, April 1990.
- [49] H.Krim, P.Forster, and J.G.Proakis, "Performance analysis of smoothed subspace-based estimation methods," in *Proc. IEEE ICASSP'91*, pp. 3053–3056, June 1991.
- [50] D. V. Bhaskar Rao and K. V. S. Hari, "Spatial smoothing and MUSIC: Further result," in *Proc. 2nd International Workshop on SVD and Signal Processing*, (Kingston, RI), pp. 128–135, June 1990.

- [51] D. V. Bhaskar Rao and K. V. S. Hari, "Weighted State-Space methods/ESPRIT and spatial smoothing," in *Proc. IEEE ICASSP'91*, pp. 3317–3320, June 1991.
- [52] A. Paulraj, T.J.Shan, and T. Kailath, "Performance analysis of the MUSIC algorithm with spatial smoothing in the presence of coherence sources," in *Proc. IEEE ICASSP'86*, (Tokyo, Japan), pp. 41.5.1–41.5.5, April 1986.
- [53] A. C. Kot, *Perturbation Analysis of SVD-Based Methods*. PhD thesis, University of Rhode Island, Kingston, RI, 1989.

APPENDIX A

PARAMETERS FOR SUBSPACE ROTATION METHOD

We first define a $\bar{\mathbf{L}}$ independent of $\mathbf{W}$ thus irrelevant to the algorithms

$$\bar{\mathbf{L}} = \mathbf{U}_s^H \mathbf{A} \quad (171)$$

and get

$$\bar{\mathbf{L}}^{-1} = \mathbf{R}_s \mathbf{A}^H \mathbf{U}_s \Sigma_s^{-1}. \quad (172)$$

Notice $\mathbf{u}_k$ is the k -th column of $\mathbf{L}$, we now define

$$\mathbf{e}_k = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}^H$$

in which only the k -th element is 1 so we have

$$\beta_k = \mathbf{V}_s \Sigma_s^{-1} \mathbf{W} \mathbf{L} \mathbf{e}_k = \mathbf{V}_s \Sigma_s^{-1} \bar{\mathbf{L}} \mathbf{e}_k \quad (173)$$

substitute (171) into (173) to obtain

$$\beta_k = \mathbf{V}_s \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{A} \mathbf{e}_k = \mathbf{V}_s \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{a}(\theta_k). \quad (174)$$

Notice $\mathbf{v}_k$ is the k -th row of $\mathbf{L}^{-1}$, we can express α_k^H as

$$\begin{aligned} \alpha_k^H &= C_k \mathbf{e}_k^H \mathbf{L}^{-1} \mathbf{W}^{-1} \mathbf{U}_s^{\dagger \dagger} (\mathbf{U}_o^{\dagger} \lambda_k^{-1} - \mathbf{U}_o^{\dagger}) \mathbf{U}_o^H \\ &= \mathbf{e}_k^H \bar{\mathbf{L}}^{-1} \mathbf{U}_s^{\dagger \dagger} (\mathbf{U}_o^{\dagger} \lambda_k^{-1} - \mathbf{U}_o^{\dagger}) \mathbf{U}_o^H. \end{aligned} \quad (175)$$

Using $\mathbf{L} = \mathbf{W}^{-1} \bar{\mathbf{L}}$, (48) becomes

$$\mathbf{A} = \mathbf{U}_s \bar{\mathbf{L}} \quad \text{or} \quad \mathbf{A}^{\dagger} = \mathbf{U}_s^{\dagger} \bar{\mathbf{L}}$$

hence

$$\mathbf{A}^{\dagger \dagger} = \bar{\mathbf{L}}^{-1} \mathbf{U}_s^{\dagger \dagger}. \quad (176)$$

If we further define $\mathbf{I}^{\dagger}$ and $\mathbf{I}^{\dagger}$ as the first $L-1$ and last $L-1$ columns, respectively, of the $L \times L$ identity matrix $\mathbf{I}$

$$(\mathbf{U}_o^{\dagger} \lambda_k^{-1} - \mathbf{U}_o^{\dagger}) \mathbf{U}_o^H = (\mathbf{I}^{\dagger} \lambda_k^{-1} - \mathbf{I}^{\dagger}) \mathbf{U}_o \mathbf{U}_o^H = (\mathbf{I}^{\dagger} \lambda_k^{-1} - \mathbf{I}^{\dagger}) \boldsymbol{\Omega}_o \quad (177)$$

where $\mathbf{\Omega}_o = \mathbf{U}_o \mathbf{U}_o^H$ is the projection matrix of orthogonal-subspace. Using (176) and (177), (175) becomes

$$\alpha_k^H = C_k \mathbf{e}_k^H \mathbf{A}^{\dagger\top} (\mathbf{I}^\top \lambda_k^{-1} - \mathbf{I}^\top) \mathbf{\Omega}_o. \quad (178)$$

Since we need preserve only k -th row of the $\mathbf{A}^{\dagger\top} (\mathbf{I}^\top \lambda_k^{-1} - \mathbf{I}^\top) \mathbf{\Omega}_o$ to calculate α_k^H , so we can rewrite (177) as

$$\alpha_k^H = C_k \mathbf{e}_k^H (\mathbf{D}^{-1} \mathbf{A}^{\dagger\top} \mathbf{I}^\top - \mathbf{A}^{\dagger\top} \mathbf{I}^\top) \mathbf{\Omega}_o = (\mathbf{A}^{\dagger\top} \mathbf{I}^\top - \mathbf{A}^{\dagger\top} \mathbf{I}^\top) \mathbf{\Omega}_o. \quad (179)$$

We now post-multiplying $(\mathbf{A}^{\dagger\top} \mathbf{I}^\top - \mathbf{A}^{\dagger\top} \mathbf{I}^\top)$ by $\mathbf{A}$

$$(\mathbf{A}^{\dagger\top} \mathbf{I}^\top - \mathbf{A}^{\dagger\top} \mathbf{I}^\top) \mathbf{A} = (\mathbf{A}^{\dagger\top} \mathbf{A}^\top - \mathbf{A}^{\dagger\top} \mathbf{A}^\top) = (\mathbf{I} - \mathbf{I}) = \mathbf{0}. \quad (180)$$

(180) shows that $(\mathbf{A}^{\dagger\top} \mathbf{I}^\top - \mathbf{A}^{\dagger\top} \mathbf{I}^\top)$ is orthogonal to $\mathbf{A}$, or is already in $\mathbf{\Omega}_o$, we can drop $\mathbf{\Omega}_o$ from α_k

$$\alpha_k^H = C_k \mathbf{e}_k^H \mathbf{A}^{\dagger\top} (\mathbf{I}^\top \lambda_k^{-1} - \mathbf{I}^\top) \quad (181)$$

or

$$\alpha_k^H = C_k \mathbf{e}_k^H (\mathbf{A}^{\dagger\top} \mathbf{I}^\top - \mathbf{A}^{\dagger\top} \mathbf{I}^\top). \quad (182)$$

Further, since

$$\begin{aligned} & C_k \mathbf{e}_k^H \mathbf{A}^{\dagger\top} (\mathbf{I}^\top \lambda_k^{-1} - \mathbf{I}^\top) \mathbf{a}^{(1)}(\theta_k) \\ &= \mathbf{e}_k^H \mathbf{A}^{\dagger\top} \left[\begin{pmatrix} e^{j \frac{2\pi d}{\lambda_c} \sin \theta_k} \\ 2e^{j 2 \frac{2\pi d}{\lambda_c} \sin \theta_k} \\ \vdots \\ (L-1)e^{j \frac{(L-1)2\pi d}{\lambda_c} \sin \theta_k} \end{pmatrix} e^{-j \frac{2\pi d}{\lambda_c} \sin \theta_k} - \begin{pmatrix} 0 \\ e^{j \frac{2\pi d}{\lambda_c} \sin \theta_k} \\ \vdots \\ (L-2)e^{j \frac{(L-2)2\pi d}{\lambda_c} \sin \theta_k} \end{pmatrix} \right] \end{aligned}$$

$$= \mathbf{e}_k^H \mathbf{A}^{\dagger \dagger} \begin{pmatrix} 0 \\ e^{j \frac{2\pi d}{\lambda_c} \sin \theta_k} \\ \vdots \\ e^{j \frac{(L-2)2\pi d}{\lambda_c} \sin \theta_k} \end{pmatrix} = \mathbf{e}_k^H \mathbf{A}^{\dagger \dagger} \mathbf{a}(\theta_k)^{\dagger} = 1 \quad (183)$$

we have $\|\alpha_k \mathbf{a}^{(1)}(\theta_k)\|^2 = \|\alpha_k \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2 = 1$ we can express $\|\alpha_k\|$ as

$$\|\alpha_k\|^2 = \frac{\|\alpha_k\|^2}{\|\alpha_k \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2} = \frac{1}{\|\bar{\alpha}_k \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2} \quad (184)$$

where $\bar{\alpha}_k = \frac{\alpha_k}{\|\alpha_k\|}$. The expression of mean-squared error is

$$E_{\Delta \mathbf{Y}} (\Delta \theta_k)^2 = \frac{\|\beta_k\|^2 \sigma_n^2}{2 \|\bar{\alpha}_k \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)\|^2}. \quad (185)$$

APPENDIX B

FORWARD BACKWARD AVERAGING

Since

$$\begin{aligned}
 & [Im(\beta_{FBk}^H \mathbf{N}_{FB}^H \alpha_k)]^2 \\
 = & \left(\frac{\beta_{FBk}^H \mathbf{N}_{FB}^H \alpha_k - \alpha_k^H \mathbf{N}_{FB} \beta_{FBk}}{2j} \right)^2 \\
 = & \frac{1}{2} \beta_{FBk}^H \mathbf{N}_{FB}^H \alpha_k \alpha_k^H \mathbf{N}_{FB} \beta_{FBk} \\
 - & \frac{1}{2} Re(\alpha_k^H \mathbf{N}_{FB} \beta_{FBk} \alpha_k^H \mathbf{N}_{FB} \beta_{FBk}) \quad (186)
 \end{aligned}$$

so

$$\begin{aligned}
 & E[Im(\beta_{FBk}^H \mathbf{N}_{FB}^H \alpha_k)]^2 \\
 = & \frac{1}{2} \beta_{FBk}^H E(\mathbf{N}_{FB}^H \alpha_k \alpha_k^H \mathbf{N}_{FB}) \beta_{FBk} \\
 - & \frac{1}{2} Re[\alpha_k^H E(\mathbf{N}_{FB} \beta_{FBk} \alpha_k^H \mathbf{N}_{FB}) \beta_{FBk}] \quad (187)
 \end{aligned}$$

where in the first term

$$\begin{aligned}
 & E(\mathbf{N}_{FB}^H \alpha_k \alpha_k^H \mathbf{N}_{FB}) \\
 = & E \left(\begin{pmatrix} \mathbf{N}^H \\ \mathbf{N}^T \mathbf{J} \end{pmatrix} \alpha_k \alpha_k^H (\mathbf{N} \mathbf{J} \mathbf{N}^*) \right) \\
 = & \begin{bmatrix} E(\mathbf{N}^H \alpha_k \alpha_k^H \mathbf{N}) & E(\mathbf{N}^H \alpha_k \alpha_k^H \mathbf{J} \mathbf{N}^*) \\ E(\mathbf{N}^T \mathbf{J} \alpha_k \alpha_k^H \mathbf{N}) & E(\mathbf{N}^T \mathbf{J} \alpha_k \alpha_k^H \mathbf{J} \mathbf{N}^*) \end{bmatrix} \\
 = & \sigma_n^2 ||\alpha_k||^2 \mathbf{I}_{2L \times 2L} \quad (188)
 \end{aligned}$$

the non-diagonal terms equal to zero because of circular property of $\mathbf{N}$. The first term of (187) becomes

$$\frac{\sigma_n^2 \mathbf{R}_{FB}^{-1}(k, k) ||\alpha_k||^2}{2\gamma_k^2}.$$

Unlike the case of forward-only formulation, the second term in (187) is not zero.

To calculate this term we first rewrite β_{FBk} as

$$\beta_{FBk}^H = [\mathbf{e}_k \mathbf{R}_{FB}^{-1} \mathbf{S} \quad \mathbf{e}_k \mathbf{R}_{FB}^{-1} \mathbf{D}^{-L+1} \mathbf{S}^*] = [\beta_{Fk}^H \quad \beta_{Bk}^H] \quad (189)$$

then

$$\begin{aligned} & E[\mathbf{N}_{FB} \beta_{FBk} \alpha_k^H \mathbf{N}_{FB}] \\ &= E[(\mathbf{N} \quad \mathbf{JN}^*) \begin{pmatrix} \beta_{Fk} \\ \beta_{Bk} \end{pmatrix} \alpha_k^H (\mathbf{N} \quad \mathbf{JN}^*)] \\ &= E(\mathbf{N} \beta_{Fk} \alpha_k^H \mathbf{N} + \mathbf{JN}^* \beta_{Bk} \alpha_k^H \mathbf{N} \\ & \quad \mathbf{N} \beta_{Fk} \alpha_k^H \mathbf{JN}^* + \mathbf{JN}^* \beta_{Bk} \alpha_k^H \mathbf{JN}^*) \end{aligned} \quad (190)$$

$$= E(\mathbf{JN}^* \beta_{Bk} \alpha_k^H \mathbf{N} \quad \mathbf{N} \beta_{Fk} \alpha_k^H \mathbf{JN}^*) \quad (191)$$

from (190) to (191) we used the circular property of $\mathbf{N}$ again. Now we first look at the i, j -th entry of the matrix product $E(\mathbf{JN}^* \beta_{Bk} \alpha_k^H \mathbf{N})$

$$\begin{aligned} & E(\mathbf{JN}^* \beta_{Bk} \alpha_k^H \mathbf{N})_{i,j} \\ &= E\left[\sum_{m=1}^M n_{L-i+1}(m)^* \beta_{Bk}(m) \left[\sum_{l=1}^L \alpha_k(l)^* n_l(j)\right]\right] \\ &= \sum_{m=1}^M \sum_{l=1}^L \beta_{Bk}(m) \alpha_k(l)^* E[n_{L-i+1}(m)^* n_l(j)] \\ &= \sum_{m=1}^M \sum_{l=1}^L \beta_{Bk}(m) \alpha_k(l)^* \sigma^2 \delta(L-i+1-l, j-m) \\ &= \sigma^2 \beta_{Bk}(j) \alpha_k(L-i+1)^* \end{aligned} \quad (192)$$

where the indices in subscript and parenthesis of noise element n in $\mathbf{N}$ are the indices

of the column and the row in $\mathbf{N}$. Now we obtain

$$\begin{aligned}
 E[\mathbf{N}_{FB} \beta_k \alpha_k^H \mathbf{N}_{FB}] &= \sigma_n^2 \begin{bmatrix} \alpha_k(L)^* \\ \vdots \\ \alpha_k(1)^* \end{bmatrix} \times \\
 &[\beta_{Bk}(1), \dots, \beta_{Bk}(M), \beta_{Fk}(1), \dots, \beta_{Fk}(M)] \\
 &= \sigma_n^2 \mathbf{J} \alpha_k^* (\beta_{Bk} \ \beta_{Fk})
 \end{aligned} \tag{193}$$

so that

$$\begin{aligned}
 &\alpha_k^H E[\mathbf{N}_{FB} \beta_k \alpha_k^H \mathbf{N}_{FB}] \beta_k \\
 &= \sigma_n^2 (\alpha_k^H \mathbf{J} \alpha_k^*) (\beta_{Bk}^T \ \beta_{Fk}^T) \begin{pmatrix} \beta_{Fk} \\ \beta_{Bk} \end{pmatrix} \\
 &= 2\sigma_n^2 (\alpha_k^H \mathbf{J} \alpha_k^*) (\beta_{Fk}^T \beta_{Bk})
 \end{aligned} \tag{194}$$

notice that from (189)

$$\begin{aligned}
 \beta_{Fk}^T \beta_{Bk} &= [\mathbf{R}_{FB}^{-1*} \mathbf{S}^* \mathbf{S}^T \mathbf{D}^{L-1} \mathbf{R}_{FB}^{-1}](k, k) \\
 &= [\mathbf{R}_{FB}^{-1*} \mathbf{R}^* \mathbf{D}^{L-1} \mathbf{R}_{FB}^{-1}](k, k).
 \end{aligned} \tag{195}$$

Because we are only interested in k, k -th element of $[\mathbf{R}_{FB}^{-1*} \mathbf{R}^* \mathbf{D}^{L-1} \mathbf{R}_{FB}^{-1}]$, so $\mathbf{D}^{L-1}$ can be substituted by z_k^{L-1}

$$\begin{aligned}
 \beta_{Fk}^T \beta_{Bk} &= z_k^{L-1} [\mathbf{R}_{FB}^{-1*} \mathbf{R}^* \mathbf{R}_{FB}^{-1}](k, k) \\
 &\stackrel{\text{def}}{=} z_k^{L-1} [\bar{\mathbf{R}}_{FB}^{-1}](k, k).
 \end{aligned} \tag{196}$$

Now equation (187) can be expressed as

$$E(\triangle \theta_k)^2 = \frac{\sigma_n^2 \|\alpha_k\|^2}{2\gamma_k^2} [\mathbf{R}_{FB}^{-1} + 2\text{Re}(\phi_k \bar{\mathbf{R}}_{FB}^{-1})](k, k) \tag{197}$$

or using (79)

$$E(\Delta\theta_k)^2 = \frac{\sigma_n^2[\mathbf{R}_{FB}^{-1} + 2\text{Re}(\phi_k \bar{\mathbf{R}}_{FB}^{-1})](k, k)}{2[\mathbf{a}^{(1)}(\theta_k)^H \boldsymbol{\Omega}_o \mathbf{W} \boldsymbol{\Omega}_o \mathbf{a}^{(1)}(\theta_k)]} \quad (198)$$

where we define

$$\phi_k = -z_k^{L-1} \frac{(\alpha_k^H \mathbf{J} \alpha_k^*)}{||\alpha_k||^2}. \quad (199)$$

APPENDIX C

FORWARD ONLY SPATIAL SMOOTHING

Starting from (155), we first rewrite β_k^H as:

$$\beta_k^H = \mathbf{e}_k \mathbf{R}_{sF}^{-1} [\mathbf{S} \mathbf{D} \mathbf{S} \dots \mathbf{D}^{L-Q} \mathbf{S}] \quad (200)$$

and

$$\beta_k^H \mathbf{N}^H = \mathbf{e}_k \mathbf{R}_s^{-1} [\mathbf{S} \mathbf{D} \mathbf{S} \dots \mathbf{D}^{L-Q} \mathbf{S}] [\mathbf{N}_1 \mathbf{N}_2 \dots \mathbf{N}_{L-Q+1}]^H. \quad (201)$$

Notice that we can reorder the above multiplication so that it becomes a summation of product of signal and noise from different snapshot.

$$[\mathbf{S} \mathbf{D} \mathbf{S} \dots \mathbf{D}^{L-Q} \mathbf{S}] \mathbf{N}^H = \sum_{i=1}^M \bar{\mathbf{S}}_{Fi} \bar{\mathbf{N}}_{Fi}^H \quad (202)$$

where each $\bar{\mathbf{S}}_{Fi}$ and $\bar{\mathbf{N}}_{Fi}$ are smoothed signal and noise matrix from the i^{th} snapshot.

$$\bar{\mathbf{S}}_{Fi} \stackrel{\text{def}}{=} [\mathbf{s}_i \mathbf{D} \mathbf{s}_i \dots \mathbf{D}^{L-Q} \mathbf{s}_i]$$

$$\mathbf{s}_i = [s_{1,i} \ s_{2,i} \ \dots \ s_{p,i}]^T$$

and

$$\bar{\mathbf{N}}_{Fi} \stackrel{\text{def}}{=} \begin{pmatrix} n_{1,i} & n_{2,i} & \dots & n_{L-Q+1,i} \\ \vdots & \vdots & \ddots & \vdots \\ n_{Q,i} & n_{Q+1,i} & \dots & n_{L,i} \end{pmatrix}.$$

The first term in (155) can now be written as

$$E[\mathbf{e}_k \mathbf{R}_{sF}^{-1} \sum_{i=1}^M (\bar{\mathbf{S}}_{Fi} \bar{\mathbf{N}}_{Fi}^H) \alpha_k \alpha_k^H \sum_{i=1}^M (\bar{\mathbf{N}}_{Fi} \bar{\mathbf{S}}_{Fi}^H) \mathbf{R}_{Fs}^{-1} \mathbf{e}_k^H]. \quad (203)$$

Take the expectation with respect to noise and use the snapshot independency property of signal and noise, the above expression becomes

$$\begin{aligned} & E \left(\sum_{i=1}^M \bar{\mathbf{S}}_{Fi} \bar{\mathbf{N}}_{Fi}^H \alpha_k \alpha_k^H \sum_{i=1}^M \bar{\mathbf{S}}_{Fi} \bar{\mathbf{N}}_{Fi}^H \right) \\ &= \sum_{i=1}^M E \left(\bar{\mathbf{S}}_{Fi} \bar{\mathbf{N}}_{Fi}^H \alpha_k \alpha_k^H \bar{\mathbf{N}}_{Fi} \bar{\mathbf{S}}_{Fi}^H \right). \end{aligned} \quad (204)$$

Similar to the derivation in [53], [17]

$$E[\bar{\mathbf{N}}^H \alpha_k \alpha_k^H \bar{\mathbf{N}}] = \sigma_n^2 \mathbf{B} \mathbf{B}^H \quad (205)$$

where

$$\mathbf{B}_{L-Q+1 \times L} = \begin{pmatrix} \alpha_{k1} & \alpha_{k2} & \dots & \alpha_{kQ} & & \\ & \alpha_{k1} & \alpha_{k2} & \dots & \alpha_{kQ} & 0 \\ 0 & & \ddots & \ddots & \ddots & \\ & & & \alpha_{k1} & \alpha_{k2} \dots & \alpha_{kQ} \end{pmatrix}.$$

So

$$E(\mathbf{S} \mathbf{N}^H \alpha_k \alpha_k^H \mathbf{N} \mathbf{S}^H) = \sigma_n^2 \sum_{i=1}^M \bar{\mathbf{S}}_i \mathbf{B} \mathbf{B}^H \bar{\mathbf{S}}_i^H \stackrel{\text{def}}{=} \sigma_n^2 \mathbf{R}_W \quad (206)$$

where we define the *Weighted* signal correlation matrix $\mathbf{R}_W$. The weighting matrix is only a function of the array manifold $\mathbf{A}$ (Since α_k is only a function of $\mathbf{A}$). Now we have the mean-squared error:

$$E(\Delta \theta_k)^2 = \frac{\sigma_n^2}{2} \frac{\mathbf{R}_{sF}^{-1} \mathbf{R}_W \mathbf{R}_{sF}^{-1}(k, k)}{\gamma_k^2}. \quad (207)$$

APPENDIX D

FORWARD BACKWARD SPATIAL SMOOTHING

Expending

$$E[\Im(\beta_k^H \mathbf{N}_{FB}^H \alpha_k)]^2 \quad (208)$$

$$= \frac{1}{2} E(\beta_k^H \mathbf{N}_{FB}^H \alpha_k)^2 - \frac{1}{4} E(\beta_k^H \mathbf{N}_{FB}^H \alpha_k \beta_k^H \mathbf{N}_{FB}^H \alpha_k) - \frac{1}{4} E(\alpha_k^H \mathbf{N}_{FB} \beta_k \alpha_k^H \mathbf{N}_{FB} \beta_k). \quad (209)$$

Using the result of section VI.2.2, we have

$$\begin{aligned} & E[\mathbf{S}_{FB} \mathbf{N}_{FB}^H \alpha_k \alpha_k^H \mathbf{N}_{FB} \mathbf{S}_{FB}^H] \\ &= \sigma_n^2 \sum_{i=1}^M (\bar{\mathbf{S}}_{Fi} \mathbf{D}^{1-L} \bar{\mathbf{S}}_{Fi}^*) \begin{pmatrix} \mathbf{B} \mathbf{B}^H & 0 \\ 0 & \mathbf{B}^* \mathbf{B}^T \end{pmatrix} \begin{pmatrix} \bar{\mathbf{S}}_{Fi}^H \\ \bar{\mathbf{S}}_{Fi}^T \mathbf{D}^{L-1} \end{pmatrix} \\ &= \mathbf{R}_W + \mathbf{D}^{1-L} \mathbf{R}_W^* \mathbf{D}^{L-1} \stackrel{\text{def}}{=} \mathbf{R}_{WFB}. \end{aligned} \quad (210)$$

The first term in (209) becomes

$$\frac{\sigma_n^2}{2} \mathbf{R}_{sFB}^{-1} \mathbf{R}_{WFB} \mathbf{R}_{sFB}^{-1}. \quad (211)$$

Now we derive second and third terms (they are conjugate). Rewrite

$$\beta_k^H \mathbf{N}_{FB}^H = \sum_{i=1}^M \beta_{ki}^H \bar{\mathbf{N}}_{iFB}^H$$

where β_{ki}^H and $\bar{\mathbf{N}}_{iFB}^H$ are the i th snapshot term:

$$\beta_{ki}^H = \mathbf{e}_k \mathbf{R}_{sFB}^{-1} [\bar{\mathbf{S}}_{Fi} \mathbf{D}^{1-L} \bar{\mathbf{S}}_{Fi}^*] \stackrel{\text{def}}{=} [\beta_{ki,1}^H \ \beta_{ki,2}^H]$$

$$\bar{\mathbf{N}}_{iFB}^H \stackrel{\text{def}}{=} [\bar{\mathbf{N}}_{Fi} \ \mathbf{J} \bar{\mathbf{N}}_{Fi}^*].$$

Remember the noise are independent from snapshot to snapshot. Taking the expectation, we can eliminate the cross term from third term in (209)

$$E(\alpha_k^H \mathbf{N}_{FB} \beta_k \alpha_k^H \mathbf{N}_{FB} \beta_k) = E \sum_{i=1}^M \alpha_k^H \bar{\mathbf{N}}_{iFB} \beta_{ki} \alpha_k^H \bar{\mathbf{N}}_{iFB} \beta_{ki}. \quad (212)$$

The middle part of the above expression can be further expended as

$$\begin{aligned}
& E[\bar{\mathbf{N}}_{iFB} \beta_{ki} \alpha_k^H \bar{\mathbf{N}}_{iFB}] \\
&= E(\bar{\mathbf{N}}_{Fi} \mathbf{J} \bar{\mathbf{N}}_{Fi}^*) \begin{pmatrix} \beta_{ki,1} \\ \beta_{ki,2} \end{pmatrix} \alpha_k^H (\bar{\mathbf{N}}_{Fi} \mathbf{J} \bar{\mathbf{N}}_{Fi}^*) \\
&= E(\mathbf{J} \bar{\mathbf{N}}_i^* \beta_{ki,2} \alpha_k^H \bar{\mathbf{N}}_i \bar{\mathbf{N}}_i \beta_{ki,1} \alpha_k^H \mathbf{J} \bar{\mathbf{N}}_i^*). \tag{213}
\end{aligned}$$

So equation (212) becomes

$$\begin{aligned}
& E \sum_{i=1}^M \alpha_k^H \bar{\mathbf{N}}_{iFB} \beta_{ki} \alpha_k^H \bar{\mathbf{N}}_{iFB} \beta_{ki} \\
&= \sum_{i=1}^M E(\alpha_k^H \mathbf{J} \bar{\mathbf{N}}_{Fi}^* \beta_{ki,2} \alpha_k^H \bar{\mathbf{N}}_{Fi} \beta_{ki,1} + \alpha_k^H \bar{\mathbf{N}}_{Fi} \beta_{ki,1} \alpha_k^H \mathbf{J} \bar{\mathbf{N}}_{Fi}^* \beta_{ki,2}) \tag{214} \\
&= \sum_{i=1}^M (\beta_{ki,2}^T \bar{\mathbf{B}}^* \mathbf{B}^H \beta_{ki,1} + \beta_{ki,1}^T \mathbf{B}^* \bar{\mathbf{B}}^H \beta_{ki,2}) \\
&= \mathbf{e}_k \mathbf{R}_{sFB}^{-1} \sum_{i=1}^M (\bar{\mathbf{S}}_{Fi} \mathbf{D}^{L-1} \bar{\mathbf{B}}^* \mathbf{B}^H \bar{\mathbf{S}}_{Fi}^H + \bar{\mathbf{S}}_{Fi}^* \mathbf{B}^* \bar{\mathbf{B}}^H \mathbf{D}^{L-1} \bar{\mathbf{S}}_{Fi}^T) \mathbf{R}_{sFB}^{-1} \mathbf{e}_K^H \\
&= \mathbf{R}_{sFB}^{-1} \bar{\mathbf{R}}_{WFB} \mathbf{R}_{sFB}^{-1}(k, k) \tag{215}
\end{aligned}$$

where

$$\bar{\mathbf{R}}_{WFB} \stackrel{\text{def}}{=} \sum_{i=1}^M \bar{\mathbf{S}}_{Fi} \mathbf{D}^{L-1} \bar{\mathbf{B}}^* \mathbf{B}^H \bar{\mathbf{S}}_{Fi}^H + \bar{\mathbf{S}}_{Fi}^* \mathbf{B}^* \bar{\mathbf{B}}^H \mathbf{D}^{L-1} \bar{\mathbf{S}}_{Fi}^T. \tag{216}$$

This is a *Weighted* sources correlation matrix because of the forward-backward spatial smoothing. Let

$$\bar{\mathbf{B}}_{L-Q+1 \times L} = \begin{bmatrix} \alpha_{kQ} & \alpha_{kQ-1} & \dots & \alpha_{k1} & & \\ & \alpha_{kQ} & \alpha_{kQ-1} & \dots & \alpha_{k1} & 0 \\ 0 & & \ddots & \ddots & \ddots & \\ & & & \alpha_{kQ} & \alpha_{kQ-1} & \dots & \alpha_{k1} \end{bmatrix}.$$

We now have MSE of DOA estimation algorithms

$$E(\Delta\theta_k)^2 = \frac{\sigma_n^2}{2\gamma_k^2}(\mathbf{R}_{sFB}^{-1}\mathbf{R}_{WFB}\mathbf{R}_{sFB}^{-1} - \Re(\mathbf{R}_{sFB}^{-1}\mathbf{\bar{R}}_{WFB}\mathbf{R}_{sFB}^{-1}))(k, k). \quad (217)$$