ESPRIT based Joint Angle-Frequency Estimation

Algorithms and Simulations

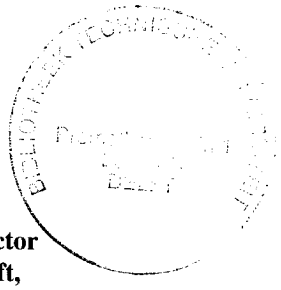
ESPRIT based Joint Angle-Frequency Estimation

Algorithms and Simulations

Aweke Negash Lemma

ESPRIT based Joint Angle-Frequency Estimation

Algorithms and Simulations



PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College voor Promoties aangewezen,
op dinsdag 11 Januari 2000 te 16.00 uur

door

Aweke Negash LEMMA

Bachelor of Science, Addis Ababa University
Master of Electronic Engineering, Technische Universiteit Eindhoven
Chartered Designer in Microelectronics, Technische Universiteit Delft

geboren te Arba Minch, Ethiopië

Dit proefschrift is goedgekeurd door de promotor:
Prof.dr.ir. P. M. Dewilde, Technische Universiteit Delft

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof.dr.ir. P.M. Dewilde, Technische Universiteit Delft, promotor
Prof.dr.ir. E.F. Deprettere, Technische Universiteit Delft, co-promotor
Prof.dr. J.A. Nossek, Technische Universität München
Prof.ir. P. van Genderen, Technische Universiteit Delft
Dr.ir. M. Moonen, Katholieke Universiteit Leuven
Dr.ir. P.C.W. Sommen, Technische Universiteit Eindhoven
Dr. I. Proudler, DERA, UK

ISBN: 90-5326-032-3

This work was conducted in cooperation with
TNO-FEL, The Hague, The Netherlands

Printed in The Netherlands

To my brother, Belete Negash Lemma

CONTENTS

LIST OF FIGURES

SUMMARY

Signals appear in our daily lives in various forms. The multitude of their facets is so overwhelming that many of them are usually unnoticed. In almost all instances, useful information is contained in certain signal parameters such as frequency, phase, amplitude, position and so on. To adequately extract the useful information, one must, therefore, be able to estimate these parameters from measured (acquired) data.

In this thesis, I consider estimation of direction of arrivals (DOAs) and center frequencies of narrowband far field radio signals from measurement data collected using arrays of sensors. The problem finds significant interest in a variety of fields, including radio astronomy, communications, radars, sonar, acoustic systems, etc. The resulting solution techniques are collectively referred to as joint angle-frequency estimation (JAFE) algorithms.

The oldest signal parameter estimation technique dates back to the war era, and involved a mere application of Fourier based spectral analysis on data collected using arrays of sensors. During the recent couple of decades, different high resolution signal parameter estimation techniques have emerged. Among these, the ESPRIT (estimation of signal parameters via rotational invariance technique) algorithm has been chosen as an underlying solution method for this work.

The ESPRIT algorithm falls into a class of estimators termed as *subspace fitting methods*. As is the case with the subspace fitting techniques, the ESPRIT algorithm estimates signal parameters by fitting the data to an underlying parameterized model. Since its derivation in 1985, the ESPRIT algorithm has evolved to become one of the dominant parameter estimation algorithms. The main attraction is that it is computationally much more efficient as compared to other high-resolution DOA estimation techniques.

A data matrix collected from a geometrically well defined antenna array impinged by narrowband far field signals has rich structure. In JAFE, the already

existing structure of the data matrix is further enriched by performing some data stacking techniques. The data stacking procedures are characterized by some parameters. These data stacking parameters affect the number of separable sources and the performance of the algorithm. Using perturbation analysis, I derive optimum values of these parameters. I show that, the algorithm attains the theoretically determined error level only when the data stacking parameters are chosen close to their optimum values. Choosing optimum data stacking parameters, however, can be computationally prohibitive.

To solve this complexity-accuracy tradeoff problem, I implement the so called multi-resolution ESPRIT (MR-ESPRIT) algorithm. In essence, the ESPRIT algorithm estimates signal parameters by exploiting structures in the data matrix induced by the underlying array of sensors. In most cases, the array geometries contain redundancies or can be designed to contain redundancies. In MR-ESPRIT, these redundancies are exploited to improve the parameter estimation accuracies. The behavioral analysis presents us with bounds on the attainable improvements.

One way of evaluating the robustness of an algorithm is to investigate its behavior when there exists a mismatch between the assumed data model and the actual behavior. To this end, I apply the JAFE algorithm to scenarios where the actual data structure significantly differs from the assumed model. For instance, the algorithm is elegantly used to detect and track frequency hopping sources and to estimate the bandwidths of modulated signals.

Finally, the derived algorithm is validated using actual field measurement data. The experimental results reveal that, in deriving the data model, it is important to appropriately consider the non-ideal array and front end system behaviors. These include, among others, antenna coupling, phase jitters in analog-to-digital converters and non-uniform antenna gain behaviors.

ACKNOWLEDGEMENT

This work has been supported by TNO Fysisch en Elektronisch Laboratorium (TNO-FEL), The Netherlands, under the projects Ondersteuning Digital Beamforming No. 015.28124.01.02.02 and Toepassing ADBF No. 015.28948.01.01.

First of all, I would like to thank all the people from the circuits and systems (CAS) research group, Delft University of Technology, and the TNO-FEL laboratories, who in one way or the another made my stay at the university quite pleasant and fruitful.

Outstandingly, I would like to thank my supervisor and co-promotor Prof. Ed. Deprettere, for presenting me with the opportunity to do my Ph.D. research at the CAS research group, for maintaining an environment to freely conduct my research, for providing me with the necessary guidances and challenges that made most of the results take their present shape and finally for the proof-reading of the original draft.

I specially would like to thank my promotor Prof. Patrick Dewilde for the careful proof-reading of the thesis and for coming up with quite useful comments that have helped in enhancing the quality of the thesis.

I would also like to thank Dr. Alle-Jan van der Veen, for sharing his expertise in this field and challenging me in several occasions, which has certainly contributed in sharpening the scientific value of the thesis.

Many thanks go to Laurens Bierens and John Scholtz for maintaining the cooperation from the TNO-FEL side, Kees van 't Wout (TNO-FEL), for helping me collect the experimental data, and Jean Paul van Assche (TNO-FEL) for supplying me with useful antenna simulation data.

I would also like to thank Marion de Vlieger and Corrie Boers-Boonekamp for their professional handling of the work, and for their kind assistance.

Finally, I would like to express my deepest appreciation to my family, specially my wife Gelila Cahalla, for their constant support and love.

Chapter 1

INTRODUCTION

Contents

Signals appear in our daily lives in various forms. The multitude of their facets is so overwhelming that many of them are usually unnoticed. In almost all instances, useful information is contained in certain signal parameters such as frequency, phase, amplitude, position and so on. To adequately extract the useful information, one must, therefore, be able to estimate these parameters from measured (acquired) data. If the parameter estimation method is good, we can handle many signals at the same time. Thus, signal parameters estimation can maximize the utilization of resources such as bandwidth. In the literature we find numerous accounts on signal parameter estimation techniques, a short history of these methods is presented in section 1.1.

In this thesis, I consider estimation of radio signal parameters from measurement data collected using arrays of sensors. In this context, it is useful to

classify signals as far field or near field depending on the transmitter-receiver separation distance compared to the physical dimensions of the sensor array, and as narrowband or wideband depending on the relative bandwidth of the signal compared to its center frequency. A typical scenario considered in the thesis is shown in Fig. 1.1. If a signal received is in the far field of its source, then the

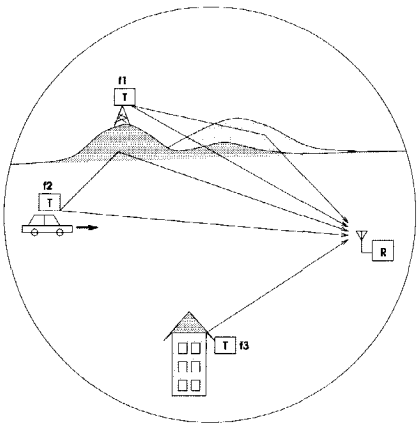


Fig. 1.1: Typical problem scenario

wavefront arriving at the receiver **R** is planar. This means that the time of arrival of the signal at two sensors separated by some distance Δ will differ by an amount proportional to the cosine of the angle between the line of propagation and the line connecting the two sensors. If the signal is narrowband, the time delay can be approximated with a phase shift. The direction of arrival (DOA) of the signal is obtained from the estimate of this phase shift. A summary of different techniques of estimating signal DOAs (and other signal parameters) is given in section 1.1. Among these, the ESPRIT algorithm (Estimation of Signal Parameters via Rotational Invariance Techniques) is chosen as an underlying solution method for my work. In the thesis, I principally addressed the problem of simultaneous estimation of DOAs and center frequencies of far field narrowband signals. Though, various related algorithms are proposed, we will collectively refer to them as the JAFE (Joint Angle-Frequency Estimation) algorithm(s).

In some applications, the angle and frequency estimates are taken as the end results. For instance, in military applications, DOA and frequency estimates are needed to detect and localize jamming sources. In mobile communication systems, this information is needed to localize emergency calls, fraud transmissions, etc. For JAFE based localization, the DOA estimation need to be conducted from two locations, as shown in Fig. 1.2. The area **A** is a measure

of the angle estimation accuracy. For good estimators, **A** is small.

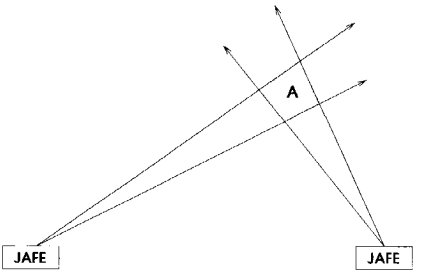


Fig. 1.2: Localizing a signal source

In other applications, the angle and frequency estimates are used for further processing. For example, to maximize bandwidth utilization via beamforming. Beamforming is a technique of maximizing reception in a certain direction (desired direction) while rejecting all other (interfering) sources. The JAFE based beamformer is slightly different from the classical beamformers in that it can discriminate between signals having the same DOA provided their center frequencies are different. In such case the basic beamforming system looks like the one shown in Fig. 1.3. The angle-frequency estimates obtained from the

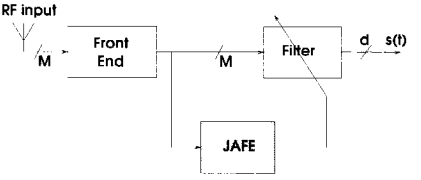


Fig. 1.3: JAFE based beamformer, where M is the number of antenna elements and d is the number of far field signals.

JAFE block are used to steer the filter to selectively extract a desired signal while rejecting the undesired ones. In the following, after presenting a brief background on the evolution of signal parameter estimation techniques, I discuss the motivation and focus of the work, and its main claims.

1.1 Short History of Signal Parameters Estimation Techniques

In many signal processing applications, it is required to estimate signal parameters from measurements collected by an array of sensors. Classic examples of such problems include direction-of-arrival (DOA) estimation, center frequency estimation, Doppler shift estimation, delay estimation and combinations thereof.

The oldest signal parameter estimation technique [4] involved an application of Fourier based spectral analysis on data collected using arrays of sensors. Later, adaptive beamforming techniques were used to improve the resolution. Among these is Capon's (1969) [13] method, known as the Minimum Variance Distortionless Response filter. The major drawback of this method is that its performance is directly dependent on the array size (aperture) irrespective of the time over which the data is collected.

An important breakthrough came in the field of geophysics by Burg (1967) [12], who applied a maximum entropy (ME) based spectral estimation technique for signal parameter estimation. Almost at the same time, the early versions of the maximum likelihood (ML) principle have been introduced by Schweppe [74] and MacDonald et. al. [54]. The ML [9, 81] approach is based on parametric models where the signal parameters are estimated by performing multi-dimensional optimization of the parametric model. While it is often successful, obtaining sufficiently accurate initial estimates is generally computationally too expensive. If these initial estimates are poor, the search procedure may converge to a local minimum and may never attain the global minimum.

A new set of parameter estimation techniques that are based on exploiting geometric structures of the data model induced by arrays of sensors were later introduced by different authors. Among the earliest works, it is worthwhile mentioning Pisarenko's (1973) [63] effort for the case of signals in additive noise. The first persons to correctly exploit the geometric information of an arbitrary (but known) array for signal parameter estimation were Schmidt (1977) [72,73] and independently Bienvenu (1979) [5,7]. Schmidt achieved this by first deriving a solution for a noise free data, and then applying the results to obtain approximate solutions for a noisy data. The resulting algorithm is known as MUSIC (MUltiple SIgnal Classification) in the signal processing literature and has been considerably developed by different authors (crf. [3,15,65,66,79,80,82]). Although MUSIC almost consistently attains performances that are close to that of the ML method, they are achieved at a considerable computational cost (Multi-dimensional search over parameter space).

A more efficient signal parameter estimation method that makes use of the concept of subspace based matrix shifting (viz. [42, 60]) was later introduced by Kung et. al. (1983) [42]. They used the subspace based matrix shift behavior for solving the harmonic retrieval problem. This concept was later developed to solve a wide range of other parameter estimation problems. The new algorithm was called the ESPRIT (Estimation of Signal Parameters via Rotational Invariant Techniques). The first full account of ESPRIT was published by Paulraj, Roy and Kailath in 1985 [60].

Since its introduction, the ESPRIT has been developed by different authors to become one of the widely used high resolution parameter estimation algorithms. In the following, I present a summary of the major developments of the ESPRIT Algorithm.

The first variation of the original ESPRIT was the incorporation of the total least squares (TLS) [20, 21, 35] principle aimed at mitigating its sensitivity to array and system imperfections. This led to the introduction of the TLS-ESPRIT by Roy and Kailath [67]. The TLS-ESPRIT algorithm was shown to be more robust to array and system errors. Later Zoltowski and Stavrinides [105] came up with another extension of ESPRIT, the PRO-ESPRIT. They incorporated a Procrustes rotation based eigenanalysis into the ESPRIT. The approach tried to solve the subspace fitting problem by splitting the $2M \times 2M$ eigenvalue problem of TLS-ESPRIT into two $M \times M$ sub-problems. The authors claim that the PRO-ESPRIT shows the same robustness as the TLS-ESPRIT, but with a reduced computational complexity.

Another version of ESPRIT aimed in improving its performance was presented by the name Multiple-Invariance ESPRIT (MI-ESPRIT) [84]. The original version of MI-ESPRIT exploited the multiple shift invariance property present in many array geometries to iteratively estimate the signal parameters. As compared to the direct ESPRIT, it provided a better performance in terms of accuracy and resolution.

Yet another form of the ESPRIT algorithm was introduced by the name *unitary*-ESPRIT by Haardt et. al. [24]. They exploited a special geometrical behavior in the so-called centro-symmetric antenna array [27, 43, 99] to obtain a factor 2 reduction in the computational complexity, without any loss of accuracy. In fact, the resulting real data matrix contained twice as many temporal samples, and hence provided a better performance and robustness.

Meanwhile, other groups of researchers were analyzing the applicability of ESPRIT for joint parameter estimation (multi-dimensional ESPRIT). In many cases, it is required to estimate more than a single signal parameter. In these cases, a joint estimation approach would save a lot of computation which

would be required to solve a non-trivial problem of grouping independently estimated parameters that belong to the same signal. In this aspect, the ESPRIT algorithm has shown tremendous advantages over other methods. It has been used to jointly estimate azimuth and elevation angles [86, 89], angle and delay [88], angle and frequency [27, 47], etc. Moreover, the joint parameter estimation principle was implemented in the so-called multi-resolution ESPRIT (MR-ESPRIT) algorithm [47,48,50] to improve the accuracy of ESPRIT. The MR-ESPRIT algorithm considers multiple shift invariance structure as the MI-ESPRIT, however, unlike the MI-ESPRIT, the solutions are presented in a closed (non-iterative) analytical form.

1.2 Motivation

This thesis was initiated by the TNO-FEL, the Netherlands. The primary goal is to localize a strong (jamming) source in order to reject the source using either some signal processing techniques or by physically eliminating it. The solution to the problem is required to meet the following properties,

- it should use minimal signal properties. The modulation types employed, the carrier frequencies and the nature of the signals (except for being far field narrowband) are unknown. Thus, the solution is expected to work without making use of these signal behaviors.
- it must handle the presence of more than a single source. I.e., it is required that the method continues to work in the presence of other intruding signals, and possibly be able to estimate the DOAs of these alien sources.
- it must have high resolution, and must be simple enough to be implemented using the state of the art signal processing systems without much sacrifice on its performance.
- it must achieve the required performance level for all signals in the frequency band of interest.

Originally, the problem was posed by the TNO-FEL as one that tries to improve the performance of the classical Adcock antenna array system using some digital signal processing techniques. The original Adcock system consists of four antenna elements uniformly spaced on a circle as shown in Fig. 1.4. In polar coordinate system, each voltage may be represented with magnitude and phase. If the antenna elements have identical gains, all the voltages have equal magnitudes v, and differ only by phase shifts. Let ϕ_x and ϕ_y be the phase shift

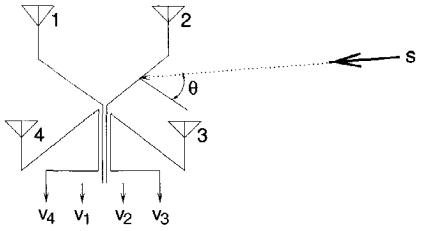


Fig. 1.4: The Adcock antenna array

between v_4 and v_2, and v_3 and v_1, respectively, d be the antenna spacing in meters and let λ be the wavelength of the radio signal in meters. Then,

$$\begin{aligned} \phi_x &= \frac{2\pi d}{\lambda}\sin(\theta) \\ \phi_y &= \frac{2\pi d}{\lambda}\cos(\theta) \end{aligned}$$

and

$$\begin{aligned} v_4 - v_2 &= 2v\sin(\frac{\phi_x}{2}) \\ v_3 - v_1 &= 2v\sin(\frac{\phi_y}{2}). \end{aligned}$$

If the antenna spacing (baseline) d is short compared to the signal wavelength λ, we may set $\sin(\frac{\phi_x}{2}) \approx \frac{\phi_x}{2}$ and $\sin(\frac{\phi_y}{2}) \approx \frac{\phi_y}{2}$. Thus, noting from Fig. 1.4 that ϕ_x and ϕ_y are proportional to the sine and cosine of θ, the DOA of the source is estimated using the approximation

$$\tan\theta \approx \frac{v_4 - v_2}{v_3 - v_1}. \tag{1.1}$$

When the short baseline condition is not met, that is when the spacing exceeds about 0.1λ, the DOA estimate begins to be biased. The bias related to the antenna spacing, commonly referred to as the octantal spacing error, is a function of DOA. Its maximum value is plotted in Fig. 1.5 as a function of antenna spacing. Apart from DOA, for a given baseline, the octantal error is also dependent on the elevation angle and the signal frequency. This makes it difficult to perform any error correction. To alleviate the problem, it has become common practice to employ the Adcock principle using more than four antenna elements. These approaches, while significantly reducing the sensitivity, fail to completely resolve the problem as they still rely on approximations similar to (1.1). Moreover, the requirement that the baseline be short means that

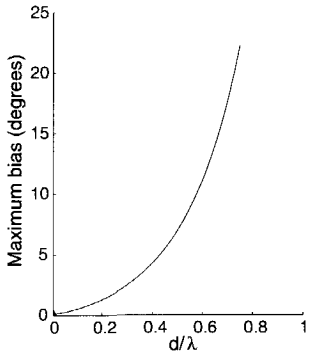


Fig. 1.5: Octantal spacing error as a function of antenna spacing

the phase shifts measured between the signals at the outputs of the antenna elements are relatively small, and hence, are quite sensitive to system noise. Apart from these, the Adcock system suffers from a number of problems including, failure to discriminate more than a single source and multipath propagation, and ambiguity between positive and negative angles.

Looking into the quality of service required and the weak points of the Adcock approach, I have chosen to depart from the initial proposal of performance improvement of the Adcock array (principle). Instead, I chose to relay on high resolution parameter estimation techniques, which are known to have superior behavior than any of the Adcock systems. Particularly, I use the ESPRIT algorithm as the underlying solution method. The preference is principally based on complexity-accuracy tradeoff (viz. section 1.1).

In the thesis, I give a complete description of an ESPRIT based solution to the problem, ranging from basic concepts to analysis from mathematical models to practical implementation issues. The work was successfully completed and the results have been validated using actual field measurement data.

1.3 Focus of the Work

This thesis focuses on two main issues: addressing the problem of jointly estimating the angle and frequency of narrowband plane wave signals (JAFE) and improving the accuracy of the ESPRIT algorithm via multi-resolution spatial and temporal sampling (the MR-ESPRIT).

In the literature, we find a few accounts on the JAFE problem. Particularly,

Haardt, et. al. [24] address the issue from the mobile communication perspectives. Their main contribution is the applicability of the unitary-ESPRIT in the JAFE context. They do not present any behavioral analysis of the algorithm. In this thesis, however, after briefly summarizing the underlying algorithm, I give an original behavioral analysis (in chapter 4). Moreover, in chapter 6, I discuss different practical implications, such as tracking of frequency hopping sources and estimating signal bandwidths.

As to the second problem, in chapter 5, I present a novel extension of the ESPRIT algorithm referred to as the MR-ESPRIT. Basically, the ESPRIT algorithm estimates signal parameters by exploiting structures in the data matrix induced by the underlying array of sensors. In most cases, the array geometries contain redundancies or can be designed to contain redundancies. In MR-ESPRIT, these redundancies are exploited to improve the parameter estimation accuracies.

The above two classes of problems are indeed based on the same multi dimensional ESPRIT principle. Consequently, they nicely fit into the same category. Moreover, in chapter 5, they are merged to develop an MR-ESPRIT based JAFE algorithm.

1.4 Summary of the Main Results

- *Performance analysis* [49]:- With a perturbation analysis of the JAFE algorithm, I am able to derive analytic expressions for the parameter estimation errors. The errors are shown to be dependent on certain *data stacking* parameters. The optimal values of these parameters are derived, and the results are shown to be consistent using simulation results.
- *MR-ESPRIT* [47, 50]:- I introduce a new variation of the ESPRIT algorithm referred to as the MR-ESPRIT algorithm. The MR-ESPRIT uses multi-resolution spatio-temporal sampling to significantly improve the performance of the ESPRIT algorithm without increasing the computational complexity. Using behavioral analysis, bounds on the attainable performance improvements are derived.
- *Practical considerations* [46, 51]:- One way of evaluating the robustness of an algorithm is to investigate its behavior when there exists a mismatch in the data model. To this end, I apply the JAFE algorithm to scenarios where the actual data structure significantly differs from the assumed model. For instance, the algorithm is elegantly used to detect and track frequency hopping sources and to estimate the bandwidths of modulated signals.

1.5 Notations

Throughout, column vectors are denoted by lower case bold faced letters and matrices with upper case bold faced letters. Row vectors are made distinct from column vectors by writing them as transposes $(\cdot)^T$ or Hermitian transposes $(\cdot)^H$ of column vectors.

For any positive integer p $\boldsymbol{I}_p$ denotes a $p \times p$ identity matrix and $\boldsymbol{\Pi}_p$ a $p \times p$ exchange matrix with ones on its anti-diagonal and zeros else were,

$$\boldsymbol{\Pi}_p = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}.$$

When these matrices are expressed without subscript, like $\boldsymbol{I}$ and $\boldsymbol{\Pi}$, they represent the identity and the exchange matrices with appropriate dimensions.

A diagonal matrix constructed by taking a sequence of entries like

$$\{\phi_0 \;\; \phi_1 \;\; \ldots \;\; \phi_N\}$$

or a sequence of function values like

$$\{\phi(0) \;\; \phi(1) \;\; \ldots \;\; \phi(N)\}$$

will often be written as

$$\boldsymbol{\Phi} = \mathrm{diag}\{\phi_i\}_{i=0}^{N} \quad \text{or} \quad \boldsymbol{\Phi} = \mathrm{diag}\{\phi(n)\}_{n=0}^{N},$$

respectively.

While the operator $\mathrm{conj}(\cdot)$ returns the complex conjugate of its input argument, the operators $\mathrm{real}(\cdot)$ and $\mathrm{imag}(\cdot)$ return the real and imaginary parts of their arguments, respectively. The set of complex numbers is denoted by $\mathbb{C}$ and the set of all real numbers by $\mathbb{R}$.

For three matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{m,n}$ and $\boldsymbol{C} \in \mathbb{C}^{p,q}$ the Kroneker product $\boldsymbol{A} \otimes \boldsymbol{C}$ and the entry-wise product $\boldsymbol{A} \odot \boldsymbol{B}$ are defined as

$$\boldsymbol{A} \otimes \boldsymbol{C} = \begin{bmatrix} a_{1,1}\boldsymbol{C} & a_{1,2}\boldsymbol{C} & \cdots & a_{1,n}\boldsymbol{C} \\ a_{2,1}\boldsymbol{C} & a_{2,2}\boldsymbol{C} & \cdots & a_{2,n}\boldsymbol{C} \\ \vdots & \vdots & & \vdots \\ a_{m,1}\boldsymbol{C} & a_{m,2}\boldsymbol{C} & \cdots & a_{m,n}\boldsymbol{C} \end{bmatrix} \in \mathbb{C}^{mp,nq}$$

and

$$\boldsymbol{A} \odot \boldsymbol{B} = \begin{bmatrix} a_{1,1}b_{1,1} & a_{1,2}b_{1,2} & \cdots & a_{1,n}b_{1,n} \\ a_{2,1}b_{2,1} & a_{2,2}b_{2,2} & \cdots & a_{2,n}b_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1}b_{m,1} & a_{m,2}b_{m,2} & \cdots & a_{m,n}b_{m,n} \end{bmatrix} \in \mathbb{C}^{m,n},$$

respectively. For a certain process or experiment, the statistical expectation is denoted with $E\{\cdot\}$. The symbols $:=$ and $=:$ are used to define terms or expressions. $l := r$ means that the left hand term or expression l is defined by the right hand term or expression r, and vice-versa for $l =: r$.

For an $M \times N$ matrix $\boldsymbol{A}$, the Moore-Penrose inverse (pseodo inverse) $\boldsymbol{A}^{\dagger}$ is defined as

$$\begin{aligned} \boldsymbol{A}^{\dagger} &= (\boldsymbol{A}^{H}\boldsymbol{A})^{-1}\boldsymbol{A}^{H}, \quad M \geq N \\ \boldsymbol{A}^{\dagger} &= \boldsymbol{A}^{H}(\boldsymbol{A}\boldsymbol{A}^{H})^{-1}, \quad M \leq N. \end{aligned}$$

1.6 Outline

All signal processing problems usually start with modeling. Signal modeling is a fairly complicated matter. However, in most cases, it is possible to come up with useful models by making some underlying assumptions. In chapter 2, I address the signal modeling issue and show how it fits into the context of antenna array for the implementation of the ESPRIT algorithm.

In chapter 3, I discuss some data manipulation technique that enriches the structure of the data matrix for the implementation of the Joint angle-frequency estimation (JAFE) algorithm. This is followed by the performance analysis of the algorithm in chapter 4. Here, I derive some optimal data stacking parameters and also derive bounds on the achievable parameter estimation accuracies.

The computational complexity of the JAFE algorithm implemented by taking the optimal data stacking parameters is usually prohibitive. This is mainly because of the presence of redundancy. In chapter 5, I use a sub-Nyquist spatio-temporal sampling rates to get rid of these redundancies. As a result, the computational complexity is reduced substantially without significant loss of accuracy.

In many applications, the considered signals do not perfectly fit into the assumed data model. In chapter 6, I address some practical examples where, despite some model mismatch, the JAFE algorithm may still be implemented by putting some additional restriction in the way the data is collected. Moreover, in the same chapter, I also discuss how some real life phenomena such as

antenna coupling, moving sources and receivers, etc, affect the performance of the algorithm.

It is mostly argued that high-resolution parameter estimation algorithms are far from implementation. To this end, I give some indications of the implementation challenges in chapter 7. Following this (chapter 8), the theoretical results discussed in the thesis are validated using actual field measurement data. Finally, the significance of this thesis, other related works and future directions are discussed in chapter 9.

Chapter 2

THE ESPRIT ALGORITHM

Contents

The ESPRIT algorithm falls into a class of estimators termed as *subspace fitting methods* [6, 69, 70, 73] (summaries of subspace based signal parameter estimation algorithms are presented in [41, 91, 92]). As is the case with the subspace fitting techniques, the ESPRIT algorithm estimates signal parameters by fitting the data to an underlying parameterized model. Since its derivation

in 1985 by Roy et. al. [60] [1], the ESPRIT algorithm has evolved to become one of the dominant parameter estimation algorithms. The main attraction is that it is computationally much more efficient as compared to other high-resolution DOA estimation techniques such as MUSIC.

In this chapter, I give a summary of the standard ESPRIT algorithm. The chapter is outlined as follows: first, in section 2.1, the data model and the array geometry required for the proper implementation of the ESPRIT algorithm is discussed. Then, in section 2.2, I introduce the notions of selection matrices and shift invariance properties. Following this, in section 2.3, the basic principles of the ESPRIT algorithm are presented. Finally, after considering an extension of ESPRIT to 2-dimensional parameter estimation (section 2.4.2), I conclude the chapter with some remarks.

2.1 The Data Model

Modeling the propagation of signals through a radio channel is a fairly complicated problem. A simplified and useful parametric model can be arrived at by making some assumptions. Following, I present a simple signal propagation model that is useful for our purpose, and show how it fits into the antenna array context for the implementation of the ESPRIT algorithm.

2.1.1 Narrowband Assumption

Throughout the thesis, I make the assumption that all signals satisfy a narrowband model. Thus, to make the rest of the discussions easy to follow, it is useful to clearly state the meaning the assumption in our context. Consider a certain

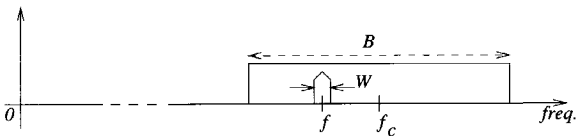


Fig. 2.1: Interpreting the narrowband assumption

frequency band B centered around a certain nominal carrier frequency f_c as shown in Fig. 2.1. Let $s(t)$ be a modulated signal with a bandwidth W and a

[1] Though the standard ESPRIT algorithm as presented in this work is due to Roy et. al., the concept of subspace based matrix shifting technique was first introduced by Kung et.al [42] in 1983.

center frequency f, where $f_c - \frac{B}{2} \leq f \leq f_c + \frac{B}{2}$. Then, in the sequel, the *narrowband assumption* is defined as

$$W \ll B \ll f_c.$$

In other words, the narrowband assumption states that the signal bandwidth W is much smaller than the observed band B, and secondly, it states that the observed band B has a band width much smaller than its center frequency f_c. The first assumption implies that a delay in the signal can be approximated by a phase shift in the complex envelope (see section 2.1.2). Whereas, the second assumption implies that, for a given time delay, all the signals in the band of interest experience equal phase shifts. The consequence of the second assumption will be clear later when we consider antenna array response.

2.1.2 Complex Signal Representation

Suppose that we are interested in the frequency band B centered at the nominal carrier frequency f_c as described in section 2.1.1. Let $s(t)$ be a narrowband signal with center frequency f, such that $f - f_c \in [-\frac{B}{2}, \frac{B}{2}]$. In the complex signal representation, let $\underline{s}(t)$ represent the complex envelope of $s(t)$, then we may write

$$s(t) = \underline{s}(t)e^{j2\pi ft} \tag{2.1}$$

Since ESPRIT-type algorithms do not require a priori exact carrier recovery,

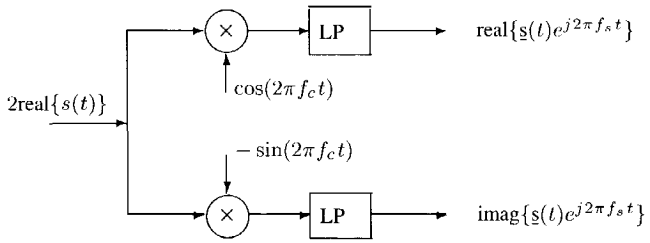


Fig. 2.2: Quadrature signal detector, where LP is a unity gain ideal low pass filter

a quadrature detector can be operated at the nominal carrier frequency f_c as indicated in Fig. 2.2 to generate complex data from the real input signal. The estimation of the center frequency f, then, reduces to estimating the offset (IF signal) frequency $f_s = f - f_c$. Assume at this point that $s(t)$ is delayed by a time τ, then

$$s(t-\tau) \quad = \quad \underline{s}(t-\tau)e^{j2\pi f(t-\tau)}$$

$$= \underline{s}(t-\tau)e^{-j2\pi f\tau}e^{j2\pi ft}. \tag{2.2}$$

The envelope of the delayed complex signal is, therefore,

$$\underline{s}_\tau(t) = \underline{s}(t-\tau)e^{-j2\pi f\tau}.$$

If the bandwidth W of the complex envelope is such that $|2\pi W\tau| \ll 1$, we can approximate

$$\underline{s}(t-\tau) \approx \underline{s}(t), \tag{2.3}$$

and thus $\underline{s}_\tau(t) \approx \underline{s}(t)e^{-j2\pi f\tau}$. In other words, the narrowband assumption allows us to approximate small time delays with phase shifts in the complex envelop. This is an important observation, as the ESPRIT algorithm (and many other signal parameter estimation algorithms) base their principles on this underlying property.

2.1.3 Antenna Array Response

Consider the scenario described in the original ESPRIT [60, 71] (Fig. 2.3), where M pairs of antenna doublets, with identical elements and translationally displaced in an arbitrary fashion are impinged by d narrowband plane wave signals. For $i = 1, \cdots, d$, let $g_i(t)$ represent the signal from the i-th source, and let f_i be its center frequency. Assume that the signal propagates in a single path (no multi-path) with the direction of arrival (DOA) α_i measured relative to the axes of the doublets as shown in Fig. 2.3. If the distances between the

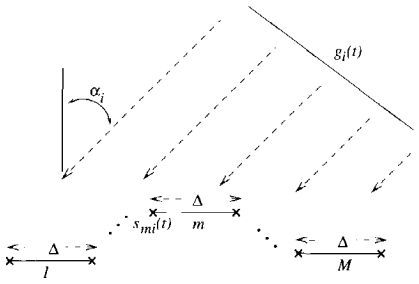


Fig. 2.3: An M doublets antenna array receiving a far field narrowband signal

array and the sources are large compared to the extent of the array, the incident waves on the array are approximately planar. Let the signal from the i-th source received at the input of the first antenna element of the m-th doublet be $s_{mi}(t)$ (it differs from $g_i(t)$ by a delay and some complex attenuation), then the corresponding signal received at the input of the second element is $s_{mi}(t-\tau_i)$,

where τ_i is the propagation time delay of the i-th signal between the doublet elements. If τ_i is small compared to the inverse bandwidth of the signal, as described in section 2.1.2, we get the approximation

$$s_{mi}(t-\tau_i) \approx s_{mi}(t)e^{-j2\pi f_i \tau_i} =: s_{mi}(t)\theta_i. \tag{2.4}$$

This shows that the signal received at the two elements of the doublet differ by a simple phase shift. Here, $\theta_i = e^{-j2\pi f_i \tau_i}$ is the parameterization of the signal DOA. In the sequel, it is referred to as the *parameterized DOA* of the i-th incoming signal. Let Δ be the antenna separation between the two elements of the antenna doublets measured in terms of fractions of the signal wavelength λ, and let c be the speed of propagation of the signal in free space, then

$$\tau_i = -\frac{\arg\theta_i}{2\pi f_i} = \frac{\Delta\lambda}{c}\sin\alpha_i. \tag{2.5}$$

Noting that $c = f_i\lambda$, it follows that $\tau_i = \frac{\Delta}{f_i}\sin\alpha_i$ and

$$\alpha_i = -\text{asin}\left(\frac{\arg\theta_i}{2\pi\Delta}\right) \tag{2.6}$$

If $\Delta < 1/2$, this transformation is unique. In the ESPRIT algorithm, first the parameterized DOAs are estimated, and then the actual angles of arrivals are computed from these estimates using the above transformation.

As the antenna spacing Δ is measured in terms of fractions of signal wavelength, it is frequency dependent. Let f_c be the nominal carrier frequency of the band of interest, and let $\mathring{\Delta}$ be the antenna spacing measured in terms of fractions of the signal wavelength corresponding to f_c. Then, the distance Δ measured in terms of the signal wavelength at a center frequency f is given by

$$\Delta = \mathring{\Delta}\frac{f}{f_c} = \mathring{\Delta}\,(1+\frac{f_s}{f_c}), \tag{2.7}$$

where $f_s = f - f_c$ is the IF signal frequency (see Fig. 2.1). If $|f_s| \ll f_c$ for all frequencies in the band of interest, Δ can be assumed to be constant. However, if f_s is comparable to f_c, the dependency of Δ on frequency cannot be neglected. In this case, the parameterized DOA θ is frequency dependent, and to note this, we write $\theta(f)$.

$$\theta(f) = \theta(f_c)^{\mathring{\Delta}\frac{f}{f_c}}. \tag{2.8}$$

From the parameter estimation point of view, the approximate representation of θ introduces a frequency dependent bias in the DOA estimates, which can

be corrected if the frequencies are known or determined using the Joint Angle-Frequency Estimation (JAFE) algorithm described in chapter 3. In the remainder of the sequel, we neglect the dependency of θ on f, and make corrections only when needed.

Let the signal due to the i-th source received at the input of the first element of the first antenna doublet be denoted by $s_i(t)$ (i.e., referring to (2.4), we set $s_{1i}(t) = s_i(t)$ to simplify subsequent notations). Moreover, let $b_1(\alpha_i)$ denote the gain of the antenna elements of the first doublet [2] to a signal from direction α_i. Without loss of generality, we assume that the antenna elements have flat frequency responses in the band of interest. Assuming that the channel is linear, and that the d narrowband signals are in the far field, the measured signals at the outputs of the first and second elements of the first doublet, respectively, are given by

$$x_1(t) = \sum_{i=1}^{d} b_1(\alpha_i)s_i(t) + n_{x1}(t) \quad \text{and} \quad y_1(t) = \sum_{i=1}^{d} b_1(\alpha_i)\theta_i s_i(t) + n_{y1}(t),$$

where $n_{x1}(t)$ and $n_{y1}(t)$ are noise terms, which are assumed to be white Gaussian. Let τ_{mi} be the additional delay taken by the i-th signal to propagate from the first to the m-th antenna doublet. If τ_{mi} is much smaller than the inverse bandwidth of $s_i(t)$, then

$$s_{mi}(t) = s_i(t)e^{-j2\pi\tau_{mi}f_i}. \tag{2.9}$$

Let $b_m(\alpha_i)$ be the gain of the antenna elements of the m-th doublet to the signal $s_i(t)$ from direction α_i, then the signal at the outputs of the first and second antenna elements of the m-th doublets, respectively, are

$$x_m(t) = \sum_{i=1}^{d} b_m(\alpha_i)s_i(t)e^{-j2\pi\tau_{mi}f_i} + n_{xm}(t),$$

and

$$y_m(t) = \sum_{i=1}^{d} b_m(\alpha_i)\theta_i s_i(t)e^{-j2\pi\tau_{mi}f_i} + n_{ym}(t),$$

where it should be noted that, since we are considering doublets with equal baseline separations and identical orientations (see Fig. 2.3), $\theta_i = e^{j2\pi f_i\tau_i}$ is the same for all the doublets. Let $\beta_m(\alpha_i)$ be defined as

$$\beta_m(\alpha_i) = b_m(\alpha_i)e^{-j2\pi\tau_{mi}f_i}$$

[2] Note that we have assumed that each doublet consists of identical antenna elements.

and let

$$\begin{aligned}
\boldsymbol{\beta}(\alpha_i) &= [\beta_1(\alpha_i)\ \beta_2(\alpha_i)\ \cdots\ \beta_M(\alpha_i)]^T, \\
\boldsymbol{n}_x(t) &= [n_{x1}(t)\ n_{x2}(t)\ \cdots\ n_{xM}(t)]^T, \\
\boldsymbol{n}_y(t) &= [n_{y1}(t)\ n_{y2}(t)\ \cdots\ n_{yM}(t)]^T, \\
\boldsymbol{x}(t) &= [x_1(t)\ x_2(t)\ \cdots\ x_M(t)]^T, \\
\boldsymbol{y}(t) &= [y_1(t)\ y_2(t)\ \cdots\ y_M(t)]^T,
\end{aligned}$$

then, from the above, it follows that

$$\begin{aligned}
\boldsymbol{x}(t) &= \sum_{i=1}^{d} \boldsymbol{\beta}(\alpha_i) s_i(t) + \boldsymbol{n}_x(t) \qquad (2.10) \\
\boldsymbol{y}(t) &= \sum_{i=1}^{d} \boldsymbol{\beta}(\alpha_i) \theta_i s_i(t) + \boldsymbol{n}_y(t).
\end{aligned}$$

Let the $d \times 1$ vector $\underline{s}(t) = [\underline{s}_1(t)\ \cdots\ \underline{s}_d(t)]^T$ be the stack of the envelops of the d impinging signals, the diagonal matrix $\boldsymbol{\Theta} = \text{diag}\{\theta_i\}_{i=1}^d$ contain the d parameterized DOAs of the d sources, and the $M \times d$ array gain matrix $\boldsymbol{B}$ be defined as

$$\boldsymbol{B} = [\boldsymbol{\beta}(\alpha_1)\ \cdots\ \boldsymbol{\beta}(\alpha_d)] \in \mathbb{C}^{M,d}.$$

Then, the data model of (2.10) can be written in a compact form as

$$\begin{aligned}
\boldsymbol{x}(t) &= \boldsymbol{B}\underline{s}(t) + \boldsymbol{n}_x(t) \qquad (2.11) \\
\boldsymbol{y}(t) &= \boldsymbol{B}\boldsymbol{\Theta}\underline{s}(t) + \boldsymbol{n}_y(t).
\end{aligned}$$

At this point, assume that the bandwidths of the narrowband signals are such that they can be sampled at a period T to satisfy the Nyquist rate. Moreover, assume that the band of interest has a bandwidth P times larger than that of the narrowband signals (i.e., referring to Fig. 2.1, we assume that $B = PW$). This means that, after demodulation to the IF, we have to sample at a period $\frac{T}{P}$. We normalize to $T = 1$.

Recall, from section 2.1.2, that the band of interest is assumed to be centered around a nominal carrier frequency f_c, and that the IF demodulation can be implemented using quadrature detector operating at the frequency f_c. Thus, if f_i is the center frequency of the i-th signal, after IF demodulation, its center frequency will be shifted to $f_{si} = f_i - f_c$. In the remainder of the thesis, to ease the notations, the IF signal frequency will be denoted with the term f_i instead of f_{si}. This means that, depending whether we are considering signals before or after the IF demodulation, f_i takes the meanings of RF (radio frequency) or IF signal frequency. The meaning should be clear from the context. Assume

now that we have collected N samples of the antenna outputs $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ at a rate P into the matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. Then, with $\phi_i = e^{j2\pi\frac{f_i}{P}}$ and $\boldsymbol{\Phi} = \mathrm{diag}\{\phi_i\}_{i=1}^{d}$,

$$\begin{aligned}\boldsymbol{X} &= \boldsymbol{B}\left[\ \underline{\boldsymbol{s}}(0) \quad \boldsymbol{\Phi}\underline{\boldsymbol{s}}(\tfrac{1}{P}) \quad \cdots \quad \boldsymbol{\Phi}^{N-1}\underline{\boldsymbol{s}}(\tfrac{N-1}{P})\ \right] + \boldsymbol{N}_x \in \mathbb{C}^{M,N} \\ \boldsymbol{Y} &= \boldsymbol{B}\boldsymbol{\Theta}\left[\ \underline{\boldsymbol{s}}(0) \quad \boldsymbol{\Phi}\underline{\boldsymbol{s}}(\tfrac{1}{P}) \quad \cdots \quad \boldsymbol{\Phi}^{N-1}\underline{\boldsymbol{s}}(\tfrac{N-1}{P})\ \right] + \boldsymbol{N}_y \in \mathbb{C}^{M,N}\end{aligned}$$

where the $M \times N$ matrices $\boldsymbol{N}_x$ and $\boldsymbol{N}_y$ are collections of N time samples of the noise vectors $\boldsymbol{n}_x(t)$ and $\boldsymbol{n}_y(t)$, respectively. The motivation for stacking the data into $M \times N$ matrices instead of MN-vectors is that, given the above signal model, collecting the data into matrix blocks makes the ranks of $\boldsymbol{X} - \boldsymbol{N}_x$ and $\boldsymbol{Y} - \boldsymbol{N}_y$ to be equal to d. This in turn makes it possible to estimate the d signal parameters via the subspace fitting technique. Let the $d \times N$ matrix $\boldsymbol{F}_s$ be defined as

$$\boldsymbol{F}_s = \left[\ \underline{\boldsymbol{s}}(0) \quad \boldsymbol{\Phi}\underline{\boldsymbol{s}}(\tfrac{1}{P}) \quad \cdots \quad \boldsymbol{\Phi}^{N-1}\underline{\boldsymbol{s}}(\tfrac{N-1}{P})\ \right] \in \mathbb{C}^{d,N}. \tag{2.12}$$

Then, $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be written more compactly as

$$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{F}_s + \boldsymbol{N}_x \quad \boldsymbol{Y} = \boldsymbol{B}\boldsymbol{\Theta}\boldsymbol{F}_s + \boldsymbol{N}_y. \tag{2.13}$$

It is seen that, under noise free conditions, the matrix $\boldsymbol{Y}\boldsymbol{F}_s^{\dagger}$ is a purely scaled version of $\boldsymbol{X}\boldsymbol{F}_s^{\dagger}$, and the scaling diagonal matrix $\boldsymbol{\Theta}$ contains the parameterized DOAs of the d incoming signals. The ESPRIT algorithm solves for $\boldsymbol{\Theta}$ by fitting the noisy measurement data to the above model. This leads to an elegant eigenvalue decomposition problem. [60, 70, 86, 88]

2.2 Selection Matrices

The doublet structure discussed in the previous subsection considers two subarrays with no overlap. That is, the matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ do not contain any common rows. Generally, however, it is possible to allow partial overlap of subarrays. In the latter case, the use of selection matrices simplifies the exposition tremendously. In this section, after considering a simple overlapping array structure, I shall introduce the notion of selection matrices.

2.2.1 The Uniform Linear Array

The simplest overlapping array geometry is the uniform linear array (ULA). In ULA, identical antenna elements are arranged uniformly on a line as shown in Fig. 2.4. Assume that an M element ULA is impinged by d narrowband

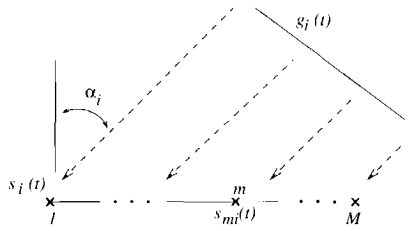


Fig. 2.4: An M element uniform linear antenna array

far field signals $g_i(t)$ with DOAs α_i. Let the corresponding signals received at the input of the first (reference) antenna element be denoted by $s_i(t)$ ($s_i(t)$ differs from $g_i(t)$ by a delay and complex attenuation). Then, the i-th signal received at the input of the m-th antenna element is $s_i(t - \tau_{mi})$, where τ_{mi} is the additional time taken by the i-th wavefront to propagate between the m-th and the reference antenna elements. From the ULA geometry, it can readily be shown that

$$\tau_{mi} = (m-1)\frac{\Delta\lambda}{c}\sin\alpha_i =: (m-1)\tau_i.$$

Let the gain of the antenna elements to the signal from direction α_i be given by $b(\alpha_i)$. Moreover, assume that the propagation delays between the antenna elements may be approximated by phase shifts as discussed in section 2.1.2. Then, with $\theta_i = e^{-2\pi\tau_i}$, the signal at the input of the m-th antenna element is

$$s_{mi}(t) = b(\alpha_i)\theta_i^{m-1}s_i(t).$$

To simplify notations, we assume that the antenna elements have omni-directional response, in which case we normalize to $b(\alpha_i) = 1$. Since the algorithm uses the phase information only, this does not impose any restriction. The measured signal at the output of the m-th antenna element is then given by

$$x_m(t) = \sum_{i=1}^{d} s_{mi}(t) + n_m(t) = \sum_{i=1}^{d} \theta_i^{m-1}s_i(t) + n_m(t),$$

where, $n_m(t)$ represents the noise at the m-th antenna element, which we assumed to be white Gaussian. Let the vector $\boldsymbol{a}(\theta_i)$ be defined as

$$\boldsymbol{a}(\theta_i) = \begin{bmatrix} 1 & \theta_i & \theta_i^2 & \dots & \theta_i^{M-1} \end{bmatrix}^T, \tag{2.14}$$

and let the M-vector $\boldsymbol{n}(t)$ be the stack of the noise terms of the M antenna elements, then collecting the outputs of the antenna elements into an M-vector

$\boldsymbol{x}(t)$, we obtain

$$\boldsymbol{x}(t) = \sum_{i=1}^{d} \boldsymbol{a}(\theta_i) s_i(t) + \boldsymbol{n}(t). \tag{2.15}$$

The vector $\boldsymbol{a}(\theta_i)$, referred to as the array steering vector, represents the response of the array to the signal with a parameterized DOA θ_i. Let the matrix $\boldsymbol{A}$ be defined as the collection of the d array steering vectors,

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}(\theta_1) & \dots & \boldsymbol{a}(\theta_d) \end{bmatrix} \in \mathbb{C}^{M,d}.$$

$\boldsymbol{A}$ is referred to as the array steering matrix. For a ULA, it has the following Vandermonde structure,

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \theta_1 & \theta_2 & \dots & \theta_d \\ \vdots & \vdots & & \vdots \\ \theta_1^{M-1} & \theta_2^{M-1} & \dots & \theta_d^{M-1} \end{bmatrix} \in \mathbb{C}^{M,d}. \tag{2.16}$$

With this definition, the summation in (2.15) can be written in a compact form as

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t),$$

where $\boldsymbol{s}(t) = [s_1(t) \ \cdots \ s_d(t)]^T$ is the stack of the d impinging signals. Let the bandwidth of $\boldsymbol{s}(t)$ be such that it can be sampled at a rate T to satisfy the Nyquist rate, we normalize to $T = 1$. Moreover, assume that we want to observe a frequency band that has a bandwidth P times larger than that of $\boldsymbol{s}(t)$. This means that, to avoid aliasing of this band, we have to sample at a rate P. Assume that we have collected N samples of the array output $\boldsymbol{x}(t)$ into the $M \times N$ data matrix $\boldsymbol{X}$, i.e.,

$$\boldsymbol{X} = \boldsymbol{A} \begin{bmatrix} \boldsymbol{s}(0) & \boldsymbol{s}(\frac{1}{P}) & \cdots & \boldsymbol{s}(\frac{N-1}{P}) \end{bmatrix} + \boldsymbol{N} =: \boldsymbol{A}\boldsymbol{F}_s + \boldsymbol{N} \in \mathbb{C}^{M,N}, \tag{2.17}$$

where the $d \times N$ matrix $\boldsymbol{F}_s$ is the collection of N samples of the d incoming signals as defined in (2.12). The above data model is rich in structure. As will be clear shortly, the ESPRIT algorithm exploits this structure to estimate the DOAs of the signals.

2.2.2 Selection Matrices and shift invariance

Given an array steering matrix $\boldsymbol{A} \in \mathbb{C}^{M,d}$, a selection matrix selects m rows of $\boldsymbol{A}$, where $m < M$. Thus, since each row of $\boldsymbol{A}$ is mapped onto a single element

of the antenna array, selection matrices can be used to specify a particular sub-array configuration. For example if we let $\boldsymbol{J}_x$ select the first $M-1$ rows and $\boldsymbol{J}_y$ select the last $M-1$ rows of $\boldsymbol{A}$, then these selection matrices are given by

$$\boldsymbol{J}_x = \begin{bmatrix} \boldsymbol{I}_{M-1} & 0 \end{bmatrix} \in \mathbb{R}^{M-1\times M} \text{ and } \boldsymbol{J}_y = \begin{bmatrix} 0 & \boldsymbol{I}_{M-1} \end{bmatrix} \in \mathbb{R}^{M-1\times M} \tag{2.18}$$

The structure of $\boldsymbol{A}$ in (2.16) is such that the sub-matrix $\boldsymbol{A}_y = \boldsymbol{J}_y\boldsymbol{A} \in \mathbb{C}^{M-1,d}$ is a scaled version of the sub-matrix $\boldsymbol{A}_x = \boldsymbol{J}_x\boldsymbol{A} \in \mathbb{C}^{M-1,d}$, that is,

$$\boldsymbol{A}_y = \boldsymbol{A}_x\boldsymbol{\Theta}, \quad \text{where } \boldsymbol{\Theta} = \text{diag}\{\theta_i\}_{i=1}^{d}. \tag{2.19}$$

This is referred to as the *shift invariance* property. The ESPRIT algorithm exploits this shift invariance structure of the array response matrix to estimate the parameterized DOAs. For a more complex array geometry, these selection matrices may be more complex.

2.3 The standard ESPRIT Algorithm

Now, we are in a position to discuss the basic computational stages involved in the ESPRIT algorithm. We will refer to the procedure as the *standard ESPRIT*. Though we derive the algorithm based on a ULA geometry with M antenna elements, the results can be extended readily to other array geometries.

Considering the data model given in (2.17), the initial step of the algorithm is to estimate the signal subspace. This, among other things, includes estimation of the number of active signal sources (estimating the rank of the data matrix). Though there have been several accounts to this problem [14, 40, 44, 77, 94, 98, 99], it remains to be an open research topic. The problem is, if we underestimate the size of the signal space, we throw away useful information, whereas if we overestimate it, we end up with an ill-conditioned estimation procedure. Thus, a reliable rank estimation is important for good performance. For our purpose, we assume that the number of sources is known or determined in some reliable way, and is equal to d. Once the size of the signal subspace is known, the next step is to estimate the subspace itself from the noise corrupted data, typically using the SVD (singular value decomposition). Note that to correctly estimate the signal subspace, we must have $d < M$. Let the data matrix $\boldsymbol{X}$ be decomposed using SVD, $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$. We partition this as

$$\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H = \begin{bmatrix} \boldsymbol{U}_s & \boldsymbol{U}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_s & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_s & \boldsymbol{V}_n \end{bmatrix}^H,$$

where the diagonal matrix $\boldsymbol{\Sigma}_s \in \mathbb{R}^{d,d}$ contains the d largest singular values, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are partitioned accordingly. The columns of $\boldsymbol{U}_s \in \mathbb{C}^{M,d}$ span

the estimated signal subspace, and the rank reduced data $\boldsymbol{X}_s$, defined as

$$\boldsymbol{X}_s = \boldsymbol{U}_s \boldsymbol{\Sigma}_s \boldsymbol{V}_s^H \in \mathbb{C}^{M,N}$$

is the estimate of the noise free data matrix, i.e., referring to (2.17) $\boldsymbol{X}_s \approx \boldsymbol{X} - \boldsymbol{N}_x = \boldsymbol{A}\boldsymbol{F}_s$. Since $\boldsymbol{U}_s$ and $\boldsymbol{A}$ span the same column space, they are related via a $d \times d$ nonsingular matrix $\boldsymbol{T}$ as

$$\boldsymbol{U}_s = \boldsymbol{A}\boldsymbol{T}^{-1}. \tag{2.20}$$

The second step in the algorithm is to form submatrices of $\boldsymbol{U}_s$ with the required shift invariance properties, using the proper selection matrices. To this end, let the selection matrices $\boldsymbol{J}_x$ and $\boldsymbol{J}_y$ be defined as in (2.18), and let the two matrices $\boldsymbol{U}_x$ and $\boldsymbol{U}_y$ be given by

$$\boldsymbol{U}_x = \boldsymbol{J}_x \boldsymbol{U}_s \in \mathbb{C}^{M-1,d}, \qquad \boldsymbol{U}_y = \boldsymbol{J}_y \boldsymbol{U}_s \in \mathbb{C}^{M-1,d}.$$

Putting (2.20) into the above expressions, and letting $\boldsymbol{A}_x = \boldsymbol{J}_x \boldsymbol{A} \in \mathbb{C}^{M-1,d}$ and $\boldsymbol{A}_y = \boldsymbol{J}_y \boldsymbol{A} \in \mathbb{C}^{M-1,d}$, we obtain

$$\boldsymbol{U}_x = \boldsymbol{A}_x \boldsymbol{T}^{-1}, \qquad \boldsymbol{U}_y = \boldsymbol{A}_y \boldsymbol{T}^{-1}.$$

The shift-invariance structure of the array steering matrix implies that

$$\boldsymbol{A}_y = \boldsymbol{A}_x \boldsymbol{\Theta} \Rightarrow \boldsymbol{U}_y = \boldsymbol{A}_x \boldsymbol{\Theta} \boldsymbol{T}^{-1} = \boldsymbol{U}_x \boldsymbol{T} \boldsymbol{\Theta} \boldsymbol{T}^{-1} =: \boldsymbol{U}_x \boldsymbol{E}_s,$$

where $\boldsymbol{\Theta} = \operatorname{diag}\{\theta_i\}_{i=1}^d$ contains the d shift invariant parameters of the d sources and $\boldsymbol{E}_s = \boldsymbol{T}\boldsymbol{\Theta}\boldsymbol{T}^{-1}$ is computed as

$$\boldsymbol{E}_s = \arg\min_{\boldsymbol{E}} \|\boldsymbol{U}_y - \boldsymbol{U}_x \boldsymbol{E}\|_F^2. \tag{2.21}$$

This minimization problem may be solved using Least Squares (LS), Total Least Squares (TLS) or Structured List Squares (SLS) approaches [25, 58, 65, 85]. It has been shown in [91] that the minimization problem has a unique solution and that as $N \to \infty$, the estimate of $\boldsymbol{E}_s$ converges to its true value. Finally, the parameters are estimated from the EVD of $\boldsymbol{E}_s$:

$$\boldsymbol{\Theta} = \boldsymbol{T}^{-1} \boldsymbol{E}_s \boldsymbol{T}.$$

Note that, for this to work, the submatrices $\boldsymbol{U}_x$ and $\boldsymbol{U}_y$ must both have at least d rows. Once the parameterized DOAs are computed, the transformation given in (2.6) may be used to compute the DOA in degrees. A summary of the standard ESPRIT algorithm is given in table 2.1.

Table 2.1: The standard ESPRIT algorithm

1. Collect tha sampled antenna output into the $M \times N$ data matrix $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{F}_s + \boldsymbol{N}_x$

2. Estimate signal subspace: i.e., compute $\boldsymbol{U}_s$ as the d dominant left singular vectors of $\boldsymbol{X}$, and apply the appropriate selection matrices to obtain submatrices with the required invariance property.

$$\begin{aligned} \boldsymbol{U}_x &= \boldsymbol{J}_x\boldsymbol{U}_s = \boldsymbol{A}_x\boldsymbol{T}^{-1} \\ \boldsymbol{U}_y &= \boldsymbol{J}_y\boldsymbol{U}_s = \boldsymbol{A}_x\boldsymbol{\Theta}\boldsymbol{T}^{-1} = \boldsymbol{U}_x\boldsymbol{E} \end{aligned}$$

3. Estimate $\boldsymbol{E}$ by solving the invariance equation (LS,TLS,SLS):

$$\boldsymbol{E} = \arg\min_{\boldsymbol{E}} \|\boldsymbol{U}_y - \boldsymbol{U}_x\boldsymbol{E}\|_F^2.$$

4. Determine the the DOAs from EVD of $\boldsymbol{E}$, i.e.,

$$\begin{aligned} \boldsymbol{\Theta} &= \boldsymbol{T}^{-1}\boldsymbol{E}\boldsymbol{T}, \\ \alpha_i &= -\text{asin}\left(\frac{\arg\theta_i}{2\pi\Delta}\right) \end{aligned}$$

2.4 Uniform Circular Antenna Array

Before concluding our discussion on antenna arrays, we will take a look at a circular antenna geometry (UCA), which has some interesting features that make it more attractive than a ULA. Consider a UCA consisting of M identical antenna elements spatially distributed on a circle, as shown in Fig. 2.5. The adjacent array elements are spaced by Δ units, and each antenna element is placed at an angular displacement of $\Delta\alpha = \frac{2\pi}{M}$ relative to the axis passing through the previous two antenna elements as shown in the figure. Let there be d point sources in the far field generating narrow band modulated signals and let α_i represent the angle of arrival of the i-th wave front measured with respect to the normal of the antenna axis $1-2$, and let the first antenna element

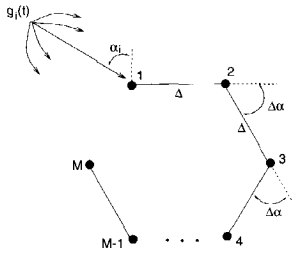


Fig. 2.5: Uniformly circular antenna array

be the reference antenna, then the relative propagation delays with respect to the reference antenna element are (see Fig. 2.5)

$$\begin{aligned}
\tau_{1i} &= 0 \\
\tau_{2i} &= \frac{\Delta}{f_i}\sin(\alpha_i) \\
\tau_{3i} &= \tau_{2i} + \frac{\Delta}{f_i}\sin(\alpha_i + \Delta\alpha) \\
\tau_{4i} &= \tau_{3i} + \frac{\Delta}{f_i}\sin(\alpha_i + 2\Delta\alpha) \\
&\vdots
\end{aligned}$$

where Δ is measured in terms of fractions of the signal wavelength, and is assumed to be constant over the band of interest (see section 2.1) and f_i is the RF signal carrier frequency. Using some trigonometric identities, for the m-th antenna element ($m \geq 2$), it can be shown that

$$\tau_{mi} = \frac{\Delta}{f_i}\sin(\alpha_i)\sum_{k=0}^{m-2}\cos(k\frac{2\pi}{M}) + \frac{\Delta}{f_i}\cos(\alpha_i)\sum_{k=0}^{m-2}\sin(k\frac{2\pi}{M}). \qquad (2.22)$$

With $\theta_i = e^{j2\pi f_i}$, the array steering vector is then given by

$$\boldsymbol{a}_i = \begin{bmatrix} 1 & \theta_i^{\tau_{2i}} & \ldots & \theta_i^{\tau_{Mi}} \end{bmatrix}^T \qquad (2.23)$$

The Fourier series expansion of the above power terms leads to a set of Bessel functions [56] with which an elegant formulation of a variant of the ESPRIT algorithm is possible. For our purpose however, we take the simple case of $M = 4$ to illustrate the strong points of the UCA geometry.

2.4.1 Uniform Circular Array with $M = 4$

In this section we give special attention to a four element UCA, referred to as 4-UCA. With reference to the discussions in the previous section, for $M = 4$,

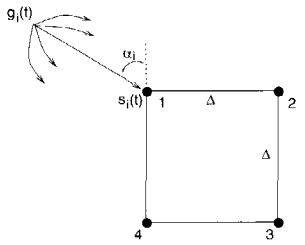


Fig. 2.6: A 4-UCA antenna array

we have $\Delta\alpha = \frac{2\pi}{M} = \frac{\pi}{2}$ and the resulting array geometry is as shown in Fig. 2.6. Let $\theta_i = e^{-j2\pi\Delta\sin(\alpha_i)}$ and $\gamma_i = e^{-j2\pi\Delta\cos(\alpha_i)}$. Setting $M = 4$ in (2.22) and using (2.23), it can be shown that the i-th steering vector of the 4-UCA is given by

$$\boldsymbol{a}_i = \begin{bmatrix} 1 & \theta_i & \theta_i\gamma_i & \gamma_i \end{bmatrix}^T, \tag{2.24}$$

The corresponding array steering matrix, thus, has the following structure,

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \theta_1 & \theta_2 & \dots & \theta_d \\ \theta_1\gamma_1 & \theta_2\gamma_2 & \dots & \theta_d\gamma_d \\ \gamma_1 & \gamma_2 & \dots & \gamma_d \end{bmatrix}. \tag{2.25}$$

Define the four selection matrices,

$$\begin{aligned} \boldsymbol{J}_{1\theta} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \boldsymbol{J}_{2\theta} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ \boldsymbol{J}_{1\gamma} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{J}_{2\gamma} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \end{aligned} \tag{2.26}$$

and let $\boldsymbol{\Gamma} = \text{diag}\{\gamma_i\}_{i=1}^d$, $\boldsymbol{\Theta} = \text{diag}\{\theta_i\}_{i=1}^d$. We then get the following factorizations,

$$\begin{aligned} \boldsymbol{J}_{2\theta}\boldsymbol{A} &= \boldsymbol{J}_{1\theta}\boldsymbol{A}\boldsymbol{\Theta} \\ \boldsymbol{J}_{2\gamma}\boldsymbol{A} &= \boldsymbol{J}_{1\gamma}\boldsymbol{A}\boldsymbol{\Gamma}. \end{aligned} \tag{2.27}$$

From these, it is seen that we can directly apply the procedure outlined in table 2.1 to estimate $\boldsymbol{\Theta}$ and $\boldsymbol{\Gamma}$. Once these two have been estimated, the DOAs are

computed using one of the following transformations

$$\alpha_i = \text{asin}\left(\frac{\arg\theta_i}{\Delta}\right) \tag{2.28}$$

$$\alpha_i = \text{acos}\left(\frac{\arg\gamma_i}{\Delta}\right) \tag{2.29}$$

$$\alpha_i = \text{atan}\left(\frac{\arg\theta_i}{\arg\gamma_i}\right). \tag{2.30}$$

Here, it seems that it is sufficient to estimate either $\boldsymbol{\Gamma}$ or $\boldsymbol{\Theta}$, but not both. However, as shown in appendix A.1, the independent transformations from θ_i to α_i and from γ_i to α_i are quite unstable around $\alpha_i = \pm\frac{\pi}{2}$ and $\alpha_i = 0$, respectively. On the other hand, the sensitivity of the last transformation is independent of α_i, making it more attractive than the others. This is one strong advantage of UCA over ULA geometry. To correctly use this advantage, however, the parameters θ_i and γ_i for $i = 1, \cdots, d$ must be correctly matched. To this end, as will be shown in the following section, a 2-dimensional (2-D) ESPRIT can be used to jointly solve for $\boldsymbol{\Theta}$ and $\boldsymbol{\Gamma}$, resulting in an automatic pairing of the parameters.

2.4.2 A 2-D ESPRIT Algorithm

Consider the 4-UCA array of Fig. 2.6. Let the $M \times N$ matrix $\boldsymbol{X}$ be constructed by collecting N samples of the $M = 4$ antenna elements. Let also the $M \times d$ unitary matrix $\boldsymbol{U}_s$ span the signal subspace. $\boldsymbol{U}_s$ is obtained by computing the SVD of the data matrix $\boldsymbol{X}$. As before, for an invertible $d \times d$ matrix $\boldsymbol{T}$, we have the relation

$$\boldsymbol{U}_s = \boldsymbol{A}\boldsymbol{T}^{-1},$$

where $\boldsymbol{A}$ is as given in (2.25). Using the pairs of selection matrices $(\boldsymbol{J}_{1\theta}, \boldsymbol{J}_{2\theta})$ and $(\boldsymbol{J}_{1\gamma}, \boldsymbol{J}_{2\gamma})$ defined in (2.26), and noting that $\boldsymbol{U}_s = \boldsymbol{A}\boldsymbol{T}^{-1}$, it follows that

$$\begin{aligned} \boldsymbol{J}_{2\theta}\boldsymbol{A} = \boldsymbol{J}_{1\theta}\boldsymbol{A}\boldsymbol{\Theta} \quad &\Rightarrow \quad \boldsymbol{J}_{2\theta}\boldsymbol{U}_s = \boldsymbol{J}_{1\theta}\boldsymbol{U}_s\boldsymbol{T}\boldsymbol{\Theta}\boldsymbol{T}^{-1} \\ \boldsymbol{J}_{2\gamma}\boldsymbol{A} = \boldsymbol{J}_{1\gamma}\boldsymbol{A}\boldsymbol{\Gamma} \quad &\Rightarrow \quad \boldsymbol{J}_{2\gamma}\boldsymbol{U}_s = \boldsymbol{J}_{1\gamma}\boldsymbol{U}_s\boldsymbol{T}\boldsymbol{\Gamma}\boldsymbol{T}^{-1}. \end{aligned}$$

Assuming that the above four matrices are full rank, it is seen that the two matrices $\boldsymbol{E}_\theta$ and $\boldsymbol{E}_\gamma$ defined as

$$\boldsymbol{E}_\theta = (\boldsymbol{J}_{1\theta}\boldsymbol{U}_s)^\dagger(\boldsymbol{J}_{2\theta}\boldsymbol{U}_s) = \boldsymbol{T}\boldsymbol{\Theta}\boldsymbol{T}^{-1}$$

and

$$\boldsymbol{E}_\gamma = (\boldsymbol{J}_{1\gamma}\boldsymbol{U}_s)^\dagger(\boldsymbol{J}_{2\gamma}\boldsymbol{U}_s) = \boldsymbol{T}\boldsymbol{\Gamma}\boldsymbol{T}^{-1}$$

are jointly diagonalizable by the same matrix $\boldsymbol{T}$. There are several algorithms to compute this joint diagonalization problem [17,23,28,86,88]. Some of these methods are summarized in chapter 3. The interesting outcome of the joint diagonalization process is that the parameters belonging to the same signal are on the same positions in the diagonal matrices $\boldsymbol{\Theta}$ and $\boldsymbol{\Gamma}$. Once these pairs are determined, one can use the relation given in (2.30) to compute the DOAs of the incoming signals.

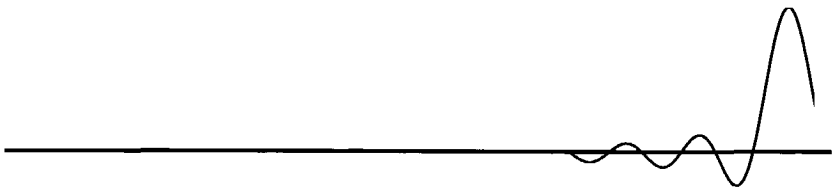
2.5 Discussion

In this chapter, a summary of the ESPRIT algorithm is presented. As shown, the signal parameters are estimated by fitting the measured data to an underlying parametric model. The resulting data structure is formulated into a matrix shifting problem, where the signal parameters are estimated from the shifting factors. The algorithm works if the matrices involved in the matrix shifting are full rank and have at least as many rows as number of sources.

Throughout the discussion, we have assumed that all the sources propagate in distinct paths and that they have unique DOAs. When this ideal behavior is violated, the data matrix becomes rank deficient and the algorithm fails to give reliable results. In such cases, some data manipulation techniques may be employed to restore the rank of the data matrix. These rank restoring techniques are considered in chapter 3 along with the joint angle-frequency estimation (JAFE) algorithm.

When deriving the data model, we have made the assumption that the signals under consideration are narrowband. The narrowband concept used here has double meanings. First, it means that the incoming signals are narrowband compared to the observed band B, and secondly, the bandwidth of the band of interest is much smaller than the nominal carrier frequency f_c. While the first narrowband assumption is important for the shift invariance property, the second assumption allows us to consider the array steering vectors to be frequency independent (or the parameterized DOAs to be frequency independent). If B is comparable to f_c, the DOAs can no more be assumed to be frequency independent. In this case, we need to compensate for the bias introduced by frequency. If the frequencies are unknown, this means that the center frequencies of the signals must be estimated in addition to their DOAs. This can be considered as one motivation for the JAFE algorithm, which is the main theme of chapter 3.

Chapter 3



JOINT ANGLE-FREQUENCY ESTIMATION

Contents

When ESPRIT is used to estimate multiple signal parameters, such as angle and frequency, the problem may be solved in one of the following two ways. In the first approach, the individual signal parameters are estimated independently, and only then (using some matching algorithm) the parameters that belong to the same signal are grouped together. Apart from the computational overhead needed for parameter matching, this leads to a numerically less conditioned set of solutions as it does not exploit the relation between the individual estimation problems.

A second method is to combine the individual estimation problems into a single *joint parameter estimation* problem. In our context, joint parameter estimation is discussed in a number of papers, including joint azimuth and elevation angle estimation [83, 86], joint frequency and 2-D angle estimation [26, 104], and joint angle and delay estimation [90]. Basically, these methods rely on the fact that each parameter is estimated from a certain eigenvalue problem, where all eigenvalue problems share the same eigenvectors (which are related to the beamforming vectors). This allows to pose the problem as a joint diagonalization problem of a collection of data matrices. The prime advantage of joint estimation is that the individual parameters are paired for free, and show a better robustness to signal and parameter disturbances.

In the literature, a number of ESPRIT-based joint angle and frequency estimation methods (JAFE) have been proposed. In particular, Zoltowski et al. [104] discuss this problem in the context of radar applications, and Haardt et al. [26] discuss it in the context of mobile communications for space division multiple access (SDMA) applications. Compared to these results, I have presented the algorithm in a more formal and elegant fashion, and have addressed some issues that have not been previously considered. The added values include the performance analysis and analysis of the behavior of the algorithm under model mismatches caused by some practical signal features.

In this chapter, I present a summary of the JAFE algorithm, and then in chapter 4, I present the performance analysis of the algorithm. Moreover, in chapter 6, it is shown that the JAFE algorithm may be used to track frequency hopping signals and to estimate bandwidths of modulated signals.

The chapter is outlined as follows: in section 3.1, a briefly summary of the signal model used is presented. Following, in section 3.2, some data stacking techniques that are useful for the implementation of the JAFE algorithm are

considered. In the same section I discuss how the data stacking techniques help in restoring the rank of the data matrix which would otherwise be rank deficient. Given the number of sensors and snapshots, there is a bound on the maximum number of resolvable sources. This issue is addressed in section 3.3. Following this, in section 3.5, I show how the JAFE algorithm may be used for blind signal reconstruction. One important step in the JAFE algorithm is the joint diagonalization (JD) of a set of eigenvalue decomposition (EVD) problems. To this end, in section 3.6, a few JD techniques relevant to our problem are summarized. Following this, in section 3.7, I discuss a 3-dimensional (3-D) extension of the JAFE algorithm for joint 2-D angle and frequency estimation. Finally, after considering a real-valued implementation of the algorithm in section 3.8, I give some concluding remarks in section 3.10.

3.1 Model

Suppose that we have an antenna array, observe a frequency band of interest (see Fig. 2.1), and want to separate and identify the directions and carrier frequencies of all sources that are present in this band. For frequency estimation to be meaningful, we assume that the sources are sufficiently narrow-band, typically with different carrier frequencies, but their spectra might be partly overlapping. The objective is to estimate the parameters and to construct a beamformer to separate the sources based on differences in arrival angles or carrier frequencies. We will assume that the sample rates are much higher than the data rates of each source, and that multi-path is negligible.

Suppose that there are d sources of interest, with complex baseband representations $s_i(t)$, where $i = 1, \cdots, d$. As in chapter 2, assume that the band of interest is centered around a nominal carrier frequency f_c, and that the i-th source has a center frequency of $f_c + f_i$ where $-\frac{B}{2} \leq f_i \leq \frac{B}{2}$ (see Fig. 2.1). This means that, after demodulation to intermediate frequency (IF) with a local oscillator operating at f_c, the IF signal due to the i-th source is $e^{j2\pi f_i t} s_i(t)$. Let θ_i be the parameterization of the direction of arrival of the i-th signal with respect to a common phase reference as described in section 2.3, let the M-vector $\boldsymbol{a}(\theta)$ be the response of the antenna array (with an arbitrary geometry) to a signal from direction θ, and $b_i \in \mathbb{R}^+$ be the amplitude of the i-th signal. If there is no multi-path, the signal received at the antenna array is

$$\boldsymbol{x}(t) = \sum_{i=1}^{d} \boldsymbol{a}(\theta_i) e^{j2\pi f_i t} b_i s_i(t) + \boldsymbol{n}(t),$$

where $\boldsymbol{n}(t) \in \mathbb{C}^{M,1}$ is a noise term. Further, as in chapter 2, assume that the narrow-band signals have bandwidths less than $\frac{1}{T}$, so that they can be sampled

with a period T to satisfy the Nyquist rate, and that the bandwidth of the band to be scanned is an integer number P times larger. I.e., referring to Fig. 2.1 we assume that $B = PW$. Normalizing to $T = 1$, this means that, after demodulation to IF we have to sample at a rate P. The data sample at the receiver is

$$\boldsymbol{x}(\tfrac{n}{P}) = \sum_{i=1}^{d} \boldsymbol{a}(\theta_i) b_i e^{j\frac{2\pi}{P} f_i n} s_i(\tfrac{n}{P}) + \boldsymbol{n}(\tfrac{n}{P}).$$

Let $\phi_i = e^{j\frac{2\pi}{P} f_i}$ be the parameterization of the i-th frequency, and let the diagonal matrix $\boldsymbol{\Phi}$ be defined as $\boldsymbol{\Phi} = \text{diag}\{\phi_i\}_{i=1}^{d}$. Moreover, let $\boldsymbol{B} = \text{diag}\{b_i\}_{i=1}^{d}$ be the signal gain matrix, $\boldsymbol{A}$ an $M \times d$ matrix collecting the d steering vectors, and finally, let the vector $\boldsymbol{s}(t)$ be a stack of the d signals. Then, the sample of the array output can be expressed in a matrix product form,

$$\boldsymbol{x}(\tfrac{n}{P}) = \boldsymbol{A}\boldsymbol{B}\boldsymbol{\Phi}^n \boldsymbol{s}(\tfrac{n}{P}) + \boldsymbol{n}(\tfrac{n}{P}). \tag{3.1}$$

Here each signal sample $s_i(\frac{n}{P})$ is assumed to have a unit amplitude. In the remainder of the chapter, unless it is necessary to write it explicitly, the diagonal matrix $\boldsymbol{B}$ in the data model is absorbed by $\boldsymbol{s}(\frac{n}{P}))$, in which case the amplitude of the i-th signal is equal to b_i instead of 1.

3.2 Data Extensions

In this section, I consider data stacking techniques that enrich the structures of the data matrix for the JAFE algorithm. These data stacking techniques, apart from structurally enriching the data matrix, affect the performance of the ESPRIT algorithm applied for the DOA estimation. Particularly, when two or more signals possess the same DOA or the same center frequency, the data matrix described in (2.17) becomes rank deficient. Under this situation, the data stacking procedures considered here may be used to restore the rank of the data matrix. Thus, as we proceed in the formulation of the JAFE algorithm, I shall also discuss how the data stacking procedures affect the DOA estimation.

3.2.1 Temporal smoothing

In this section, I consider a *temporal data stacking* technique that adds structure to the JAFE data model. Apart from structurally enriching the data matrix, temporal staking can restore the rank of a rank deficient data matrix. That is, assuming the narrowband assumption is valid, the data matrix in (2.17) is rank deficient when two or more signals have the same DOA. This is because the

array steering vectors corresponding to signals with the same DOAs are identical, and therefore the rank of the data matrix will be less than d, where d is the number of signals. It will be shown in this section that, under a certain condition, temporal stacking restores the rank of the data matrix. With some abuse of terminology, we refer to the temporally stacked data as a *temporally smoothed* data, though the smoothing is achieved later in the process. An m-*factor* temporally smoothed data matrix is constructed by stacking m temporally shifted versions of the original data matrix. This results in the following $mM \times N - m + 1$ matrix (viz. (2.17))

$$\boldsymbol{X}_m = \begin{bmatrix} \boldsymbol{A}\left[\, \boldsymbol{s}(0) \quad \boldsymbol{\Phi}\boldsymbol{s}(\frac{1}{P}) \quad \cdots \quad \boldsymbol{\Phi}^{N-m}\boldsymbol{s}(\frac{N-m}{P}) \,\right] \\ \boldsymbol{A}\boldsymbol{\Phi}\left[\, \boldsymbol{s}(\frac{1}{P}) \quad \boldsymbol{\Phi}\boldsymbol{s}(\frac{2}{P}) \quad \cdots \quad \boldsymbol{\Phi}^{N-m}\boldsymbol{s}(\frac{N-m+1}{P}) \,\right] \\ \vdots \qquad\qquad \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}^{m-1}\left[\, \boldsymbol{s}(\frac{m-1}{P}) \quad \boldsymbol{\Phi}\boldsymbol{s}(\frac{m}{P}) \quad \cdots \quad \boldsymbol{\Phi}^{N-m}\boldsymbol{s}(\frac{N-1}{P}) \,\right] \end{bmatrix} + \boldsymbol{N}_m \tag{3.2}$$

where $\boldsymbol{\Phi} = \text{diag}\{\phi_i\}_{i=1}^{d}$ contains the d parameterized frequencies [1] of the d sources, and $\boldsymbol{N}_m$ represents the noise term constructed from $\boldsymbol{N}$ in a similar way as $\boldsymbol{X}_m$ is obtained from $\boldsymbol{X}$. Assume that the signals are narrowband, i.e.,

$$\boldsymbol{s}(t) \approx \boldsymbol{s}(t + \tfrac{1}{P}) \approx \cdots \approx \boldsymbol{s}(t + \tfrac{m-1}{P}).$$

In this case, all the block rows in the right hand term of (3.2) are approximately equal, which means $\boldsymbol{X}_m$ has the factorization

$$\begin{aligned} \boldsymbol{X}_m &\approx \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{A}\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}^{m-1} \end{bmatrix} \left[\, \boldsymbol{s}(0) \quad \boldsymbol{\Phi}\boldsymbol{s}(\tfrac{1}{P}) \quad \cdots \quad \boldsymbol{\Phi}^{N-m}\boldsymbol{s}(\tfrac{N-m}{P}) \,\right] + \boldsymbol{N}_m \\ &=: \boldsymbol{A}_m \boldsymbol{F}_s + \boldsymbol{N}_m \in \mathbb{C}^{mM, N-m+1}, \end{aligned} \tag{3.3}$$

where $\boldsymbol{A}_m$, throughout the sequel, referred to as the extended array steering matrix, is given by

$$\boldsymbol{A}_m = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{A}\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}^{m-1} \end{bmatrix} \in \mathbb{C}^{mM,d} \tag{3.4}$$

[1] Note, the parameterized frequency ϕ_i is defined as $\phi_i = e^{j2\pi\frac{f_i}{P}}$ (See section 2.1).

and

$$\boldsymbol{F}_s = \left[\begin{array}{cccc} \boldsymbol{s}(0) & \boldsymbol{\Phi}\boldsymbol{s}(\frac{1}{P}) & \cdots & \boldsymbol{\Phi}^{N-m}\boldsymbol{s}(\frac{N-m}{P}) \end{array} \right] \in \mathbb{C}^{d,N-m+1} \tag{3.5}$$

is a matrix collecting $N - m + 1$ samples of the d sources.

Theorem 3.1 *Consider an M element antenna array impinged by $d < M$ narrowband far field signals. Assume that all the signals have distinct (different) center frequencies, but satisfy the narrowband assumption given in section 2.1.1. Suppose that the signals are divided into r groups, such that the signals from each group have the same DOA. Let, for $i = 1, \cdots, r$, p_i represent the number of sources in the i-th group. Then, the m-factor temporally smoothed data matrix $\boldsymbol{X}_m$ of (3.3) is full rank d if and only if $m \geq \max_i p_i$.*

Proof:- Since all the sources are assumed to have distinct frequencies, $\boldsymbol{F}_s$ in (3.3) has a full row rank. Thus, for the proof, it is sufficient to show that $\boldsymbol{A}_m$ is full column ranked matrix. To this end, let for $i = 1, \cdots, r$, θ_i represent the DOA of the signals from the i-th group, and for $j = 1, \cdots, p_i$, let $\phi_{i,j}$ be the center frequency of the j-th signal from the i-th group. More over let $\boldsymbol{A}_i$ and $\boldsymbol{\Phi}_i$ be defined as

$$\boldsymbol{A}_i = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \\ \theta_i & \theta_i & \cdots & \theta_i \\ \vdots & \vdots & & \vdots \\ \theta_i^{M-1} & \theta_i^{M-1} & \cdots & \theta_i^{M-1} \end{array} \right]$$

and $\boldsymbol{\Phi}_i = \text{diag}\{\phi_{i,j}\}_{j=1}^{p_i}$, then the extended steering matrix $\boldsymbol{A}_m$ can be expressed as

$$\boldsymbol{A}_m = \left[\begin{array}{cccc} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_r \\ \boldsymbol{A}_1\boldsymbol{\Phi}_1 & \boldsymbol{A}_2\boldsymbol{\Phi}_2 & \cdots & \boldsymbol{A}_r\boldsymbol{\Phi}_r \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_1\boldsymbol{\Phi}_1^{m-1} & \boldsymbol{A}_2\boldsymbol{\Phi}_2^{m-1} & \cdots & \boldsymbol{A}_r\boldsymbol{\Phi}_r^{m-1} \end{array} \right].$$

To prove the theorem, it is sufficient to show, for a given d-vector $\boldsymbol{x}$,

$$\boldsymbol{A}_m\boldsymbol{x} = 0 \Leftrightarrow \boldsymbol{x} = 0.$$

Let $\boldsymbol{x}$ be partitioned into r sub vectors $\boldsymbol{x}_1, \cdots \boldsymbol{x}_r$ with dimensions such that

$$\boldsymbol{A}_m\boldsymbol{x} = \left[\begin{array}{c} \boldsymbol{A}_1\boldsymbol{x}_1 + \boldsymbol{A}_2\boldsymbol{x}_2 + \cdots + \boldsymbol{A}_r\boldsymbol{x}_r \\ \boldsymbol{A}_1\boldsymbol{\Phi}_1\boldsymbol{x}_1 + \boldsymbol{A}_2\boldsymbol{\Phi}_2\boldsymbol{x}_2 + \cdots + \boldsymbol{A}_r\boldsymbol{\Phi}_r\boldsymbol{x}_r \\ \vdots \\ \boldsymbol{A}_1\boldsymbol{\Phi}_1^{m-1}\boldsymbol{x}_1 + \boldsymbol{A}_2\boldsymbol{\Phi}_2^{m-1}\boldsymbol{x}_2 + \cdots + \boldsymbol{A}_r\boldsymbol{\Phi}_r^{m-1}\boldsymbol{x}_r \end{array} \right] = 0. \tag{3.6}$$

If $\boldsymbol{A}_m$ is full column rank, the above will be satisfied if and only if $\boldsymbol{x}_1 = 0, \boldsymbol{x}_2 = 0, \cdots, \boldsymbol{x}_r = 0$. For a p-vector $\boldsymbol{v} = [v_1, \cdots, v_p]$, let the function $S(\boldsymbol{v})$ be given by

$$S(\boldsymbol{v}) = \sum_{k=1}^{p} v_i,$$

and let $\boldsymbol{A}_\theta$ be defined as

$$\boldsymbol{A}_\theta = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \theta_1 & \theta_2 & \cdots & \theta_r \\ \vdots & \vdots & & \vdots \\ \theta_1^{M-1} & \theta_2^{M-1} & \cdots & \theta_r^{M-1} \end{bmatrix}.$$

Since the θ_i are different and that we have assumed $M \geq d \geq r$, from the Vandermonde structure, it follows that $\boldsymbol{A}_\theta$ has a full column rank of r. Now, using the above definitions, (3.6) can equivalently be expressed as

$$\boldsymbol{A}_\theta \begin{bmatrix} S(\boldsymbol{x}_1) \\ S(\boldsymbol{x}_2) \\ \vdots \\ S(\boldsymbol{x}_r) \end{bmatrix} = \boldsymbol{A}_\theta \begin{bmatrix} S(\boldsymbol{\Phi}_1 \boldsymbol{x}_1) \\ S(\boldsymbol{\Phi}_2 \boldsymbol{x}_2) \\ \vdots \\ S(\boldsymbol{\Phi}_r \boldsymbol{x}_r) \end{bmatrix} = \cdots = \boldsymbol{A}_\theta \begin{bmatrix} S(\boldsymbol{\Phi}_1^{m-1} \boldsymbol{x}_1) \\ S(\boldsymbol{\Phi}_2^{m-1} \boldsymbol{x}_2) \\ \vdots \\ S(\boldsymbol{\Phi}_r^{m-1} \boldsymbol{x}_r) \end{bmatrix} = 0$$

Because $\boldsymbol{A}_\theta$ is a full rank matrix, it follows that $\boldsymbol{A}_m$ is also full rank if and only if there does not exists an $\boldsymbol{x}_i \neq 0$ such that

$$S(\boldsymbol{x}_i) = S(\boldsymbol{\Phi}_i \boldsymbol{x}_i) = \cdots = S(\boldsymbol{\Phi}_i^{m-1} \boldsymbol{x}_i) = 0, \quad i = 1, \cdots, r. \tag{3.7}$$

Let

$$\boldsymbol{A}_{\phi_i} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \phi_{i,1} & \phi_{i,2} & \cdots & \phi_{i,p_i} \\ \vdots & \vdots & & \vdots \\ \phi_{i,1}^{m-1} & \phi_{i,2}^{m-1} & \cdots & \phi_{i,p_i}^{m-1} \end{bmatrix},$$

then the conditions in (3.7) can be combined into the single expression,

$$\boldsymbol{A}_{\phi_i} \boldsymbol{x}_i = 0, \quad i = 1 \cdots, r. \tag{3.8}$$

Thus, $\boldsymbol{A}_m$ is full rank if and only if there does not exists an $\boldsymbol{x}_i \neq 0$ that satisfies (3.8). Or, in other words, $\boldsymbol{A}_m$ is full rank if and only if $\boldsymbol{A}_{\phi_i}$ is a full rank matrix. Observe that $\boldsymbol{A}_{\phi_i}$ has a Vandermonde structure. Since all the $\phi_{i,j}$ are assumed to be distinct, it will have full rank provided that $m \geq p_i$. From this, it follows

that the data matrix $\boldsymbol{X}_m$ is full rank if and only if, for $i = 1, \cdots, r$, all the $\boldsymbol{A}_{\phi_i}$ are all full rank matrices, which is satisfied if

$$m \geq \max_i p_i.$$

□

The above theorem shows that, with m-factor temporally smoothed data, the ESPRIT algorithm can resolve up to m signals having the same DOA. Apart from rank restoration, it is shown in chapter 4 that this process provides some performance improvement. Moreover, it enriches the structure of the data matrix resulting in some interesting properties. In section 3.2.2, these properties are exploited for joint angle and frequency estimation. For now, it suffices to note that temporal smoothing preserves the shift invariance structure needed for the DOA estimation. That is, the extended array steering matrix $\boldsymbol{A}_m$ has the required shift invariance structure. The DOA estimation now follows the same procedure summarized in table 2.1.

3.2.2 The JAFE algorithm

At this point, it is useful to note that the temporally smoothed data of (3.3) has much the same structure as in the classical ESPRIT algorithm (viz section 2.3), but with $\boldsymbol{A}$ replaced by the extended array steering matrix $\boldsymbol{A}_m$. However, as seen in (3.4), $\boldsymbol{A}_m$, has a dual (double) shift invariance structure. The estimation of the angle and frequency parameters and the construction of the beamformer follow by considering this dual shift invariance property. First note that, as stated in theorem 3.1, the rank of $\boldsymbol{X}_m$ is only d. Thus we compute the rank reduced (economical) SVD of $\boldsymbol{X}_m$, [2] i.e., $\boldsymbol{X}_m = \boldsymbol{U}_s \boldsymbol{\Sigma}_s \boldsymbol{V}_s$, where $\boldsymbol{U}_s$ has d columns spanning the column space of $\boldsymbol{X}_m$ ($\boldsymbol{U}_s$ corresponds to the d dominant left singular vectors of $\boldsymbol{X}_m$). As before we have

$$\boldsymbol{U}_s = \boldsymbol{A}_m \boldsymbol{T}^{-1}.$$

We begin the estimation of the parameters by defining two types of selection matrices: a pair to select submatrices for estimating the parameterized frequencies, $\boldsymbol{\Phi}$, and a pair for estimating the parameterized DOAs $\boldsymbol{\Theta}$. First, from (3.4), observe that $\boldsymbol{A}_m$ and, consequently, $\boldsymbol{U}_s$ has m blocks of M rowed matrices (referred to as block rows). As seen, these block rows form shift invariant structure via the diagonal matrix $\boldsymbol{\Phi}$. To estimate $\boldsymbol{\Phi}$, we consider this invariance structure, i.e., we take submatrices consisting of the first and, respectively, the last $m - 1$ block rows of $\boldsymbol{U}_s$. Let $\boldsymbol{J}_x(\phi)$ and $\boldsymbol{J}_y(\phi)$ be selection matrices that

[2] We assume at this stage that the number of sources is known, or determined using one of the techniques described in [14, 40, 44, 77, 94, 98, 99].

select the first and, respectively, the last $m-1$ block rows of $\boldsymbol{U}_s$, then the selection matrices are given by

$$\begin{array}{l} \boldsymbol{J}_x(\phi) := [\boldsymbol{I}_{m-1} \quad \boldsymbol{0}_1] \otimes \boldsymbol{I}_M \in \mathbb{R}^{(m-1)M,mM} \\ \boldsymbol{J}_y(\phi) := [\boldsymbol{0}_1 \quad \boldsymbol{I}_{m-1}] \otimes \boldsymbol{I}_M \in \mathbb{R}^{(m-1)M,mM}. \end{array} \tag{3.9}$$

To estimate $\boldsymbol{\Theta}$, define the selection matrices $\boldsymbol{J}_x(\theta)$ and $\boldsymbol{J}_y(\theta)$ such that $\boldsymbol{J}_x(\theta)$ selects the first $M-1$ rows from each block row of the matrix $\boldsymbol{U}_s$ and $\boldsymbol{J}_y(\theta)$ its last $M-1$ rows. Thus, selection matrices are given by

$$\begin{array}{l} \boldsymbol{J}_x(\theta) := \boldsymbol{I}_m \otimes [\boldsymbol{I}_{M-1} \quad \boldsymbol{0}_1] \in \mathbb{R}^{m(M-1),mM} \\ \boldsymbol{J}_y(\theta) := \boldsymbol{I}_m \otimes [\boldsymbol{0}_1 \quad \boldsymbol{I}_{M-1}] \in \mathbb{R}^{m(M-1),mM}. \end{array} \tag{3.10}$$

Theorem 3.2 *Let the four selection matrices $\boldsymbol{J}_x(\phi), \boldsymbol{J}_y(\phi), \boldsymbol{J}_x(\theta)$ and $\boldsymbol{J}_y(\theta)$ be as defined in (3.9) and (3.10), and let $\boldsymbol{A}_\phi = \boldsymbol{J}_x(\phi)\boldsymbol{A}_m \in \mathbb{C}^{(m-1)M,d}$ and $\boldsymbol{A}_\theta = \boldsymbol{J}_x(\theta)\boldsymbol{A}_m \in \mathbb{C}^{m(M-1),d}$. Then, the four submatrices of $\boldsymbol{U}_s$ defined as*

$$\left\{ \begin{array}{l} \boldsymbol{U}_{x,\phi} = \boldsymbol{J}_x(\phi)\boldsymbol{U}_s \in \mathbb{C}^{(m-1)M,d} \\ \boldsymbol{U}_{y,\phi} = \boldsymbol{J}_y(\phi)\boldsymbol{U}_s \in \mathbb{C}^{(m-1)M,d} \end{array} \right. \quad \left\{ \begin{array}{l} \boldsymbol{U}_{x,\theta} = \boldsymbol{J}_x(\theta)\boldsymbol{U}_s \in \mathbb{C}^{(m-1)M,d} \\ \boldsymbol{U}_{y,\theta} = \boldsymbol{J}_y(\theta)\boldsymbol{U}_s \in \mathbb{C}^{(m-1)M,d} \end{array} \right. \tag{3.11}$$

are such that

$$\left\{ \begin{array}{lcl} \boldsymbol{U}_{x,\phi} & = & \boldsymbol{A}_\phi \boldsymbol{T}^{-1} \\ \boldsymbol{U}_{y,\phi} & = & \boldsymbol{A}_\phi \boldsymbol{\Phi} \boldsymbol{T}^{-1} \end{array} \right. \quad \left\{ \begin{array}{lcl} \boldsymbol{U}_{x,\theta} & = & \boldsymbol{A}_\theta \boldsymbol{T}^{-1} \\ \boldsymbol{U}_{y,\theta} & = & \boldsymbol{A}_\theta \boldsymbol{\Theta} \boldsymbol{T}^{-1} \end{array} \right. , \tag{3.12}$$

Moreover, if these are full column ranked matrices, then there exist two jointly diagonalizable non-singular matrices $\boldsymbol{E}_\phi \in \mathbb{C}^{d,d}$ and $\boldsymbol{E}_\theta \in \mathbb{C}^{d,d}$ such that

$$\begin{array}{lcl} \boldsymbol{U}_{y,\phi} & = & \boldsymbol{U}_{x,\phi}\boldsymbol{E}_\phi \\ \boldsymbol{U}_{y,\theta} & = & \boldsymbol{U}_{x,\theta}\boldsymbol{E}_\theta \end{array} . \tag{3.13}$$

Proof:- Putting $\boldsymbol{U}_s = \boldsymbol{A}_m \boldsymbol{T}^{-1}$ into (3.11), it follows that

$$\left\{ \begin{array}{l} \boldsymbol{U}_{x,\phi} = \boldsymbol{J}_x(\phi)\boldsymbol{A}_m\boldsymbol{T}^{-1} \\ \boldsymbol{U}_{y,\phi} = \boldsymbol{J}_y(\phi)\boldsymbol{A}_m\boldsymbol{T}^{-1} \end{array} \right. \quad \left\{ \begin{array}{l} \boldsymbol{U}_{x,\theta} = \boldsymbol{J}_x(\theta)\boldsymbol{A}_m\boldsymbol{T}^{-1} \\ \boldsymbol{U}_{y,\theta} = \boldsymbol{J}_y(\theta)\boldsymbol{A}_m\boldsymbol{T}^{-1} \end{array} \right. . \tag{3.14}$$

However, from the shift invariance structure of the array steering matrix $\boldsymbol{A}_m$, it is seen that

$$\boldsymbol{J}_y(\phi)\boldsymbol{A}_m = \boldsymbol{J}_x(\phi)\boldsymbol{A}_m\boldsymbol{\Phi} \qquad \boldsymbol{J}_y(\theta)\boldsymbol{A}_m = \boldsymbol{J}_x(\theta)\boldsymbol{A}_m\boldsymbol{\Theta}$$

Now, putting these into (3.14) and letting $\boldsymbol{A}_\phi = \boldsymbol{J}_x(\phi)\boldsymbol{A}_m$ and $\boldsymbol{A}_\theta = \boldsymbol{J}_x(\theta)\boldsymbol{A}_m$ we obtain the invariance equations given in (3.12). This concludes the first part

of the proof. For the second part, note that it is assumed that all the submatrices given in (3.12) have full column ranks, hence, they have left-inverses. Substituting for $\boldsymbol{U}_{x,\phi}, \boldsymbol{U}_{y,\phi}, \boldsymbol{U}_{x,\theta}$ and $\boldsymbol{U}_{y,\theta}$, in (3.13) with the relations given in (3.12) and pre-multiplying the two equations of (3.13) with the left-inverses of $\boldsymbol{U}_{x,\phi}$ and $\boldsymbol{U}_{x,\theta}$, respectively, we get

$$\boldsymbol{E}_\phi = \boldsymbol{T}\boldsymbol{\Phi}\boldsymbol{T}^{-1} \qquad \boldsymbol{E}_\theta := \boldsymbol{T}\boldsymbol{\Theta}\boldsymbol{T}^{-1}. \tag{3.15}$$

From this, it is seen that $\boldsymbol{E}_\phi$ and $\boldsymbol{E}_\theta$ are, indeed, jointly diagonalizable by the same matrix $\boldsymbol{T}$. □

The JAFE algorithm is implemented by solving the joint diagonalization problem given in (3.15). To this end, there have been several algorithms, including the Jacobi method described in [86] and the QZ iteration methods of [17, 87]. A summary of joint diagonalization methods is given in section 3.6. Note that, for the joint diagonalization to work, it is necessary that each sub-matrix in (3.12) has at least d rows. After $\boldsymbol{T}$ has been found, we also have estimates of the pair (θ_i, ϕ_i) for each of the d sources. This provides us with angle and frequency estimates:

$$\alpha_i = \arcsin\left(\frac{\arg(\theta_i)}{2\pi\Delta}\right), \qquad f_i = \frac{P}{2\pi}\arg(\phi_i) \tag{3.16}$$

A summary of the JAFE algorithm is given in table 3.1. Note that, under noisy conditions, $\boldsymbol{E}_\phi$ and $\boldsymbol{E}_\theta$ are obtained by optimally fitting the left and right hand expressions in (3.13. This leads to the following minimization problems,

$$\begin{aligned} \boldsymbol{E}_\phi = \boldsymbol{T}\boldsymbol{\Phi}\boldsymbol{T}^{-1} &= \arg\min_{\boldsymbol{E}} \|\boldsymbol{U}_{x,\phi}\boldsymbol{E} - \boldsymbol{U}_{y,\phi}\|_F^2 \\ \boldsymbol{E}_\theta = \boldsymbol{T}\boldsymbol{\Theta}\boldsymbol{T}^{-1} &= \arg\min_{\boldsymbol{E}} \|\boldsymbol{U}_{x,\theta}\boldsymbol{E} - \boldsymbol{U}_{y,\theta}\|_F^2. \end{aligned}$$

These minimizations may be solved using the LS, TLS or SLS algorithms discussed in section 2.3.

3.2.3 Spatial Smoothing

As discussed above, temporal smoothing enables us to estimate the underlying parameters correctly even if the DOAs of more than one signal are the same. Employing a similar technique in the spatial domain, coherent signals can be separated. This is called *spatial smoothing* [61,64,78]. In spatial smoothing, an array of M sensors is sub divided into L sub-arrays. The number of elements in a sub-array depends on the way the division is made. For instance, in ULA,

Table 3.1: The JAFE algorithm

1. Collect tha sampled antenna output into the $M \times N$ data matrix $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{F}_s + \boldsymbol{N}$, and construct the $mM \times N - m + 1$, m-factor temporally smoothed data matrix $\boldsymbol{X}_m$.

2. Estimate signal subspace: i.e., compute $\boldsymbol{U}_s$ as the d dominant left singular vectors of $\boldsymbol{X}_m$, and apply the appropriate selection matrices, and solve the invariance equations (LS,TLS,SLS):

$$\begin{aligned} \boldsymbol{J}_y(\phi)\boldsymbol{U}_s &= \boldsymbol{J}_y(\phi)\boldsymbol{U}_s\boldsymbol{E}_\phi \\ \boldsymbol{J}_y(\theta)\boldsymbol{U}_s &= \boldsymbol{J}_y(\theta)\boldsymbol{U}_s\boldsymbol{E}_\theta \end{aligned}$$

3. Determine the the DOAs and frequencies by jointly solving the joint diagonalization problem,

$$\boldsymbol{\Theta} = \boldsymbol{T}^{-1}\boldsymbol{E}_\phi\boldsymbol{T}, \quad \text{and} \quad \boldsymbol{\Phi} = \boldsymbol{T}^{-1}\boldsymbol{E}_\theta\boldsymbol{T}$$

$$\Rightarrow f_i = \frac{P}{2\pi}\arg\phi_i \qquad \alpha_i = -\arcsin\left(\frac{\arg\theta_i}{2\pi\Delta}\right)$$

allowing a maximum overlap[3] as in Fig. 3.1, the number of elements per subarray is $M_L = M - L + 1$.

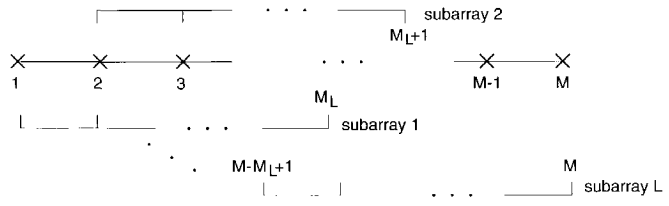


Fig. 3.1: Spatial smoothing

For $l = 1, \ldots, L$, let the $M_L \times M$ matrix $\boldsymbol{J}_l$ be a selection matrix that selects part of the $M \times N$ data matrix $\boldsymbol{X}$ that correspond to the l-th sub array. Then,

[3]Note that a maximum overlap of subarrays is obtained by shifting a selected window over a single antenna as in Fig. 3.1

a spatially smoothed $M_L \times LN$ data matrix $\boldsymbol{X}_L$ is constructed as

$$\boldsymbol{X}_L = \left[\begin{array}{cccc} \boldsymbol{J}_1\boldsymbol{X} & \boldsymbol{J}_2\boldsymbol{X} & \ldots & \boldsymbol{J}_L\boldsymbol{X} \end{array} \right] \in \mathbb{C}^{M_L, LM}. \tag{3.17}$$

Using the structure of $\boldsymbol{X}$ in (2.17), we can re-express (3.17) as

$$\boldsymbol{X}_L = \left[\begin{array}{cccc} \boldsymbol{J}_1\boldsymbol{A} & \boldsymbol{J}_2\boldsymbol{A} & \ldots & \boldsymbol{J}_L\boldsymbol{A} \end{array} \right] \left[\begin{array}{ccc} \boldsymbol{F}_s & & \\ & \ddots & \\ & & \boldsymbol{F}_s \end{array} \right] + \boldsymbol{N}_L,$$

where $\boldsymbol{N}_L$ is a noise term which has also been shuffled in a similar way as $\boldsymbol{X}_L$. Let $\boldsymbol{A}'$ contain the rows of $\boldsymbol{A}$ that correspond to the first subarray, then from the shift invariance property, we have the following relation for $k = 1, \ldots, L$

$$\boldsymbol{J}_k\boldsymbol{A} = \boldsymbol{J}_1\boldsymbol{A}\boldsymbol{\Theta}^{k-1} =: \boldsymbol{A}'\boldsymbol{\Theta}^{k-1}.$$

Using these properties, $\boldsymbol{X}_L$ can be written in a compact form as

$$\begin{aligned} \boldsymbol{X}_L &= \boldsymbol{A}' \left[\begin{array}{cccc} \boldsymbol{F}_s & \boldsymbol{\Theta}\boldsymbol{F}_s & \ddots & \boldsymbol{\Theta}^{L-1}\boldsymbol{F}_s \end{array} \right] + \boldsymbol{N}_L \\ &=: \boldsymbol{A}'\boldsymbol{F}_L + \boldsymbol{N}_L \in \mathbb{C}^{M_L, LN}. \end{aligned} \tag{3.18}$$

Theorem 3.3 *Consider an M element antenna array impinged by d narrowband far field signals. Assume that all the signals have distinct (different) DOAs. Suppose that the signals are divided into r groups, such that the signals from each group have the same center frequencies. Let, for $i = 1, \cdots, r$, q_i represent the number of sources in the i-th group. Then, the L-factor spatially smoothed data matrix $\boldsymbol{X}_L$, with $M_L > d$, (3.3) is full rank d if and only if $L \geq \max_i q_i$.*

Proof:- Consider the L-factor spatially smoothed data matrix of (3.18). As we have assumed that all the sources have different DOAs, the rank of $\boldsymbol{A}'$ is d. Thus, since $\boldsymbol{F}_L$ has only d rows, it is sufficient to show that these are linearly independent. The proof is similar to that given for theorem 3.1. First, note that ${\boldsymbol{F}_L}^T$ has the same structure as $\boldsymbol{A}_m$, with $\boldsymbol{\Theta}$ playing the role of $\boldsymbol{\Phi}$. Thus, with the same argument, it follows that $\boldsymbol{F}_L$ is full rank if

$$L \geq \max_i q_i.$$

□

3.2.4 Forward-Backward Averaging

Another way of extending the data matrix is termed as *forward-backward averaging* [2,33,52,62]. It uses the fact that the eigenvalues (θ_i, ϕ_i) lie on a unit circle and that the structure of $\boldsymbol{A}$ is *centro-symmetric* [27,43,99].

Definition 3.1 *An antenna array is said to be centro-symmetric if the element locations of the array are symmetric with respect to the centroid and the complex radiation characteristics of paired elements are the same.*

Let $\boldsymbol{\Pi}$ be an exchange matrix with appropriate dimensions and conj$(\cdot)$ be an entry-wise conjugation, then it can be shown that the steering matrix $\boldsymbol{A}$ of a centro-symmetric array satisfies

$$\text{conj}(\boldsymbol{\Pi A}) = \boldsymbol{A\Gamma}$$

for some unitary diagonal matrix $\boldsymbol{\Gamma}$. For ULA, for instance, $\boldsymbol{\Gamma} = \boldsymbol{\Theta}^{1-M}$, where the $d \times d$ diagonal matrix $\boldsymbol{\Theta}$ contains the d parameterized DOAs of the d sources as defined in section 2.3.

A forward-backward averaged data matrix $\boldsymbol{X}_{fb}$ is constructed from the data $\boldsymbol{X}$ given in (2.17) as

$$\boldsymbol{X}_{fb} = \left[\begin{array}{cc} \boldsymbol{X} & \text{conj}(\boldsymbol{\Pi X}) \end{array} \right] \in \mathbb{C}^{M,2N}. \tag{3.19}$$

It can be shown [27,90] that, if the centro-symmetric property is satisfied, the forward-backward averaged data $\boldsymbol{X}_{fb}$ has the required shift-invariant structure. We can, therefore, apply ESPRIT to solve for the underlying parameters. Note that, with this data extension, the number of available temporal samples per antenna element has essentially doubled from N to $2N$, which gives a significant improvement in accuracy. It also provides some protection against loss of rank in the case of coherent sources. i.e., even if $L = 1$ (see above), we can tolerate coherent signals with multiplicity 2.

3.2.5 Spatio-temporally smoothed and Forward-Backward Averaged Data Model

In this section, I derive a generalized data model that incorporates the above three data extension procedures. I start with the temporally smoothed data $\boldsymbol{X}_m$ given in (3.2). Let M_L be the number of antenna elements in the subarrays of the spatially smoothed data, and let for $l = 1, \ldots, L$, the selection matrix $\boldsymbol{J}_l \in \mathbb{R}^{mM_L, mM}$ select part of the data matrix $\boldsymbol{X}_m$ that corresponds to the

l-th subarray. Then, an (m, L) factor spatio-temporally smoothed data matrix $\boldsymbol{X}_{m,L}$ is constructed as

$$\boldsymbol{X}_{m,L} = \begin{bmatrix} \boldsymbol{J}_1\boldsymbol{X}_m & \boldsymbol{J}_2\boldsymbol{X}_m & \cdots & \boldsymbol{J}_l\boldsymbol{X}_m \end{bmatrix} \in \mathbb{C}^{mM_L, L(N-m+1)}. \quad (3.20)$$

Using the structure of $\boldsymbol{X}_m$ from (3.2), this can be factored as

$$\boldsymbol{X}_{m,L} = \begin{bmatrix} \boldsymbol{J}_1\boldsymbol{A}_m & \boldsymbol{J}_2\boldsymbol{A}_m & \cdots & \boldsymbol{J}_l\boldsymbol{A}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{F}_s & & \\ & \ddots & \\ & & \boldsymbol{F}_s \end{bmatrix} + \boldsymbol{N}_{m,L},$$

where $\boldsymbol{N}_{m,L}$ is a noise term which has also been shuffled in a similar way as $\boldsymbol{X}_{m,L}$. Let $\boldsymbol{A}'_m = \boldsymbol{J}_1\boldsymbol{A}_m \in \mathbb{C}^{mM_L, d}$, then from the shift invariance structure of $\boldsymbol{A}_m \in \mathbb{C}^{mM,d}$ it follows that, for $k = 1, \ldots, L$

$$\boldsymbol{J}_k\boldsymbol{A}_m = \boldsymbol{J}_1\boldsymbol{A}_m\boldsymbol{\Theta}^{k-1} =: \boldsymbol{A}'_m\boldsymbol{\Theta}^{k-1}.$$

Thus, $\boldsymbol{X}_{m,L}$ can be written in a compact form as

$$\begin{aligned} \boldsymbol{X}_{m,L} &= \boldsymbol{A}'_m \begin{bmatrix} \boldsymbol{F}_s & \boldsymbol{\Theta}\boldsymbol{F}_s & \cdots & \boldsymbol{\Theta}^{L-1}\boldsymbol{F}_s \end{bmatrix} + \boldsymbol{N}_{m,L} \\ &=: \boldsymbol{A}'_m\boldsymbol{F}_L + \boldsymbol{N}_{m,L} \end{aligned} \quad (3.21)$$

Let $\boldsymbol{\Pi}$ be an exchange matrix and $\mathrm{conj}(\cdot)$ denote entry-wise conjugation, then performing forward-backward averaging on the above (spatio-temporally smoothed) data, we get

$$\boldsymbol{X}_{m,L,fb} = \begin{bmatrix} \boldsymbol{X}_{m,L} & \mathrm{conj}(\boldsymbol{\Pi}\boldsymbol{X}_{m,L}) \end{bmatrix} \in \mathbb{C}^{mM_L, 2L(N-m+1)} \quad (3.22)$$

It can be shown [27, 90] that all the above data models contain the shift invariance properties needed by the JAFE algorithm. Thus, the angle-frequency pairs may be estimated using the procedure outlined in table 3.1.

3.3 Identifiability

The extended data $\boldsymbol{X}_{m,L,fb}$ of (3.22) is the generalized data model we want to work with. It incorporates three processes: 1) temporal smoothing 2) spatial smoothing and 3) forward-backward averaging. To derive identifiability conditions we assume that, initially, a total of N samples per antenna element are present. Thus, after temporal smoothing, spatial smoothing and forward-backward averaging, the extended data matrix $\boldsymbol{X}_{m,L,fb}$, has the dimensions $mM_L \times 2L(N-m+1)$. Let $\boldsymbol{U}_s$ be a full rank $mM_L \times d$ matrix that spans the column space of $\boldsymbol{X}_{m,L,fb}$.

Condition 1: Assuming via smoothing rank restoration has been achieved, to correctly estimate $\boldsymbol{U}_s$, $\boldsymbol{X}_{m,L,fb}$ must have at least d rows and d columns.

Once $\boldsymbol{U}_s$ is determined, the next step in JAFE is to construct submatrices with the required shift invariance properties using selection matrices. To this end, let the four selection matrices $\boldsymbol{J}_x(\phi)$, $\boldsymbol{J}_y(\phi)$, $\boldsymbol{J}_x(\theta)$ and $\boldsymbol{J}_y(\theta)$ be such that

$$\begin{cases} \boldsymbol{U}_{x,\phi} = \boldsymbol{J}_x(\phi)\boldsymbol{U}_s \\ \boldsymbol{U}_{y,\phi} = \boldsymbol{J}_y(\phi)\boldsymbol{U}_s \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{U}_{x,\theta} = \boldsymbol{J}_x(\theta)\boldsymbol{U}_s \\ \boldsymbol{U}_{y,\theta} = \boldsymbol{J}_y(\theta)\boldsymbol{U}_s \end{cases} \tag{3.23}$$

form shift invariant pairs.

Condition 2: To estimate the DOAs and frequencies properly, these matrices must have at least d rows.

The actual number of rows in these matrices depend on the way the selection matrices are defined. For a ULA for instance, with the subarrays chosen as shown in Fig. 3.1, the selection matrices are given by

$$\begin{cases} \boldsymbol{J}_x(\theta) = \boldsymbol{I}_m \otimes [\boldsymbol{I}_{M_L-1}\ \boldsymbol{0}_1], & \boldsymbol{J}_x(\phi) = \boldsymbol{I}_m \otimes [\boldsymbol{I}_{M_L-1}\ \boldsymbol{0}_1], \\ \boldsymbol{J}_y(\theta) = \boldsymbol{I}_m \otimes [\boldsymbol{0}_1\ \boldsymbol{I}_{M_L-1}]. & \boldsymbol{J}_y(\phi) = \boldsymbol{I}_m \otimes [\boldsymbol{0}_1\ \boldsymbol{I}_{M_L-1}]. \end{cases} \tag{3.24}$$

Putting these into (3.23) and noting that $M_L = M - L + 1$, it follows that $\boldsymbol{U}_{x,\phi}$ and $\boldsymbol{U}_{y,\phi}$ are both $(m-1)(M-L+1) \times d$ matrices, whereas $\boldsymbol{U}_{x,\theta}$ and $\boldsymbol{U}_{y,\theta}$ are both $m(M-L) \times d$ matrices. Thus, for ULA, combining *conditions 1 and 2*, we get the following identifiability criteria,

$$\begin{array}{ll} (a) & d \leq m(M-L) \\ (b) & d \leq (m-1)(M-L+1) \\ (c) & d \leq 2L(N-m+1). \end{array} \tag{3.25}$$

Given the number of sensors M and the number of snapshots N, we want to find the pair (m, L) that maximizes the number of signals that can be identified. In analogy to a similar problem considered in [90], we obtain as the solution to this maximization problem:

if $N \geq M + \frac{1}{\sqrt{2}}$

$$\begin{cases} d_{max} = M(N+1)(2-\sqrt{2})^2, \\ m_o = (N+1)(2-\sqrt{2}), \\ L_o = M(\sqrt{2}-1) \end{cases} \tag{3.26}$$

if $N \leq M - \frac{1}{\sqrt{2}}$

$$\begin{cases} d_{max} = N(M+1)(2-\sqrt{2})^2 \\ m_o = N(2-\sqrt{2}) + 1, \\ L_o = (M+1)(\sqrt{2}-1) \end{cases} \tag{3.27}$$

The first set of equations corresponds to a region where conditions (a) and (c) are satisfied with equality and the second set corresponds to a region where conditions (b) and (c) are satisfied with equality. The actual maxima are slightly smaller because m and L can take integer values only.

For identifiability, in addition to the above conditions, the sub-matrices $\boldsymbol{U}_{x\phi}$, $\boldsymbol{U}_{y\phi}$, $\boldsymbol{U}_{x\phi}$ and $\boldsymbol{U}_{y\phi}$ in (3.23) must also be full rank d. If the impinging wavefronts have distinct frequencies and DOAs, the Vandermonde structures of the array steering matrices ensure that this is the case. Under conditions where there are multiple DOAs or multiple center frequencies, the matrices will still be full ranked if, in addition to (3.26) and (3.27), the following are also satisfied (viz. [90])

$$\begin{cases} m & \geq & p \\ N & \geq & \frac{3}{2}p - 1 \end{cases}$$

and

$$\begin{cases} L & \geq & \frac{1}{2}q \\ M & \geq & \frac{3}{2}q, \end{cases}$$

where p and q are the multiplicity of the DOAs and center frequencies, respectively. These inequalities are derived by considering the results of theorem 3.1 and theorem 3.3.

3.4 Whitening as the JAFE Processing Stage

The spatio-temporal smoothing procedure introduces correlation between the noise terms in the different rows of the data matrix. In many cases this correlation causes degradation as it tends to reduce the degree of averaging that could have been obtained had the noise been white. In this context the JAFE algorithm can be preceded with a whitening filter as shown in Fig. 3.2. Consider

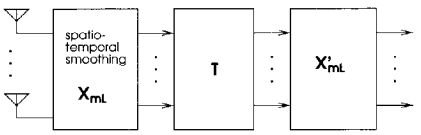


Fig. 3.2: Whitening the spatio-temporally smoothed data

the noise part of spatio-temporally smoothed data matrix given in 3.20. Let the the singular value decomposition of the noise covariance matrix $\boldsymbol{R}_{nn} = \boldsymbol{N}_{m,L}\boldsymbol{N}_{m,L}^H$ be given by

$$\boldsymbol{R}_{nn} = \boldsymbol{U}_n\boldsymbol{\Sigma}_n\boldsymbol{V}_n^H.$$

Then, the Whitened data matrix is derived as (viz. [34])

$$\boldsymbol{X}'_{m,L} = \boldsymbol{U}_n \boldsymbol{\Sigma}_n^{-1/2} \boldsymbol{U}_n^H \boldsymbol{X}_{m,L}. \tag{3.28}$$

Thus, in Fig. 3.2, the transformation matrix $\boldsymbol{T}$ is equal to $\boldsymbol{U}_n \boldsymbol{\Sigma}_n^{-1/2} \boldsymbol{U}_n^H$. Let the SVD of $\boldsymbol{X}'_{m,L}$ be given by

$$\boldsymbol{X}'_{m,L} = \boldsymbol{U}' \boldsymbol{\Sigma}' \boldsymbol{V}',$$

and let $\boldsymbol{U}'_s$ be the d dominant columns of $\boldsymbol{U}$ corresponding to the d largest singularvalues, then the JAFE algorithm is implemented using the procedure outlined in table 3.1 by putting $\boldsymbol{U}_s = \boldsymbol{U}_n \boldsymbol{\Sigma}_n^{1/2} \boldsymbol{U}_n^H \boldsymbol{U}'_s$. In section 3.9, I present simulation results comparing the performances of the JAFE algorithm implemented with and without a whitening.

3.5 Blind Signal Reconstruction

In many applications, in addition to DOAs and center frequencies, the signals themselves are of interest. Estimating the signals is usually termed as *signal reconstruction* or *signal copy* [57, 68], or beamforming. In this section, I shall briefly consider how this can be achieved under the considered data model.

Consider the temporally smoothed data model given in (3.3). [4] As the rows of $\boldsymbol{X}_m$ are constructed from the linear combinations of the rows of $\boldsymbol{F}_s$, generally, it follows that the row span of $\boldsymbol{X}_m$, (row($\boldsymbol{X}_m$)), is a subset of the row span of $\boldsymbol{F}_s$, (row($\boldsymbol{F}_s$)), i.e.,

$$\text{row}(\boldsymbol{X}) \subseteq \text{row}(\boldsymbol{F}_s) \tag{3.29}$$

The equality holds if the extended array steering matrix $\boldsymbol{A}_m$ (the row mixing matrix) in (3.3) has full column rank. Let us assume that all the sources have unique center frequencies such that $\boldsymbol{A}_m$ is made full rank by performing enough temporal smoothing as discussed in theorem 3.1. Then, since row($\boldsymbol{X}_m$) = row($\boldsymbol{F}_s$), the estimate $\hat{\boldsymbol{F}}_s$ of $\boldsymbol{F}_s$ can be obtained from the linear combinations of the rows of the data matrix $\boldsymbol{X}_m$,

$$\hat{\boldsymbol{F}}_s = \boldsymbol{W}^H \boldsymbol{X}_m \tag{3.30}$$

where $\boldsymbol{W}^H$ is a weight matrix. It is interesting to see that, in our context, the meaning of the weight matrix $\boldsymbol{W}$ is slightly different from the well understood conventional beamforming matrices [27, 100]. The main distinction is

[4] Though we are considering a temporally smoothed data with no spatial smoothing or forward-backward averaging, the results are also applicable for the general cases.

that, since $\boldsymbol{X}_m$ is a temporally smoothed data matrix, the beamforming matrix can be chosen in such a way that it can separate signals with the same DOA, provided that they have different center frequencies. To be specific, assume that q of the existing d sources (where $q < d$) come from the same direction. It has been shown in theorem 3.3 that the columns of $\boldsymbol{A}_m$ are distinct ($\boldsymbol{A}_m$ is full rank) if $m > q$ and the signals have distinct center frequencies. Let $\boldsymbol{A}'_m$ be estimate of $\boldsymbol{A}_m$ constructed from the angle-frequency estimates. Then, under full rank condition, if we choose the beamforming matrix as $\boldsymbol{W}^H = \boldsymbol{A}'^{\dagger}_m$, then it separates all the signals including those coming from the same direction. This is because, apart from the DOAs, the columns of $\boldsymbol{A}_m$ are now frequency dependent as well. Under this conditions, the reconstructed signal $\hat{\boldsymbol{F}}_s$ minimizes the error $\|\boldsymbol{X}_m - \boldsymbol{A}'_m \boldsymbol{F}_s\|_F$ over all possible $\boldsymbol{F}_s$. i.e., the impinging wavefronts are constructed as the least square solutions.

$$\hat{\boldsymbol{F}}_s = (\boldsymbol{A}'^{H}_m \boldsymbol{A}'_m)^{-1} \boldsymbol{A}'^{H}_m \boldsymbol{X} \tag{3.31}$$

3.6 Joint Diagonalization Techniques

As discussed in section 3.2.2, the JAFE algorithm is implemented by solving a joint diagonalization (JD) problem. Apart from JAFE, JD of a set of equal dimensioned matrices appears in numerous problems [11, 27, 83, 86, 88, 104]. In this section, I give a summary of some of the JD techniques. Consider a set of $d \times d$ eigenvalue decomposition (EVD) problems,

$$\boldsymbol{E}_i = \boldsymbol{T}^{-1} \boldsymbol{\Lambda}_i \boldsymbol{T} + \boldsymbol{N}_i, \qquad i = 1, \ldots, q,$$

where $\boldsymbol{N}_i$ represents a small perturbation (noise factor). Here the only assumption we make is that the unperturbed matrices $\boldsymbol{E}_i - \boldsymbol{N}_i$ are all full rank d. There are two approaches for this problem. First, one can independently solve each of the EVD problems or second, one may combine them into a single JD problem. The advantage of the second approach is that, by optimally combining the independent EVD problems, a numerically much more attractive solution method is derived. Particularly, it has been shown in [17] that, when the perturbation $\boldsymbol{N}_i$ in the data causes the eigenvalues to cross each other, the independent approaches may give errors orders of magnitudes higher than that of the joint approaches.

Let $\mathcal{U}(\cdot)$ denote an operator that extracts the off diagonal entries of its matrix argument. Then, the joint diagonalization problem can then be described as the one that tries to minimize the cost function

$$\phi(\boldsymbol{T}) = \sum_{i=1}^{q} \|\mathcal{U}(\boldsymbol{T} \boldsymbol{E}_i \boldsymbol{T}^{-1})\|_F^2, \tag{3.32}$$

where $\|.\|_F^2$ denotes a Frobenius norm. When only the $\boldsymbol{\Lambda}_i$ matrices are of interest, joint triangularization (Schur decomposition) is implemented. In this case, the cost function is defined as

$$\phi(\boldsymbol{Q}) = \sum_{i=1}^{q} \|\mathcal{L}(\boldsymbol{Q}\boldsymbol{E}_i\boldsymbol{Q}^T)\|_F^2,$$

where $\mathcal{L}(\cdot)$ is an operator that extracts the strictly lower triangular part of its matrix argument, and the upper triangular matrix $\boldsymbol{R}_i = \boldsymbol{Q}\boldsymbol{E}_i\boldsymbol{Q}^T$ contains the eigenvalues on its main diagonal, i.e., $\text{diag}\{\boldsymbol{R}_i\} = \text{diag}\{\boldsymbol{\Lambda}_i\}$.

In the literature, there have been several approaches to this JD problem. All of the approaches are suboptimal with fairly acceptable accuracies. In the following, I present summaries of some of these techniques that are relevant to the thesis.

3.6.1 Schur-Jacobi iteration

A simple and straight forward "joint" triangularization scheme is discussed in [86]. The idea is, first, to compute the Schur decomposition $\boldsymbol{E}_I = \boldsymbol{Q}_I^T\boldsymbol{R}_I\boldsymbol{Q}_I$, for some $I \in [1, q]$ and then for $i = 1 \ldots q,\ i \neq I$, to compute the decomposition $\boldsymbol{R}_i' = \boldsymbol{Q}_I\boldsymbol{E}_i\boldsymbol{Q}_I^T$. Note that when $\boldsymbol{N}_i$ in (3.32) is zero, $\boldsymbol{R}_i' = \boldsymbol{R}_i$ and the computation is complete. In noisy conditions, however, $\boldsymbol{R}_i' \neq \boldsymbol{R}_i$. Nevertheless, it is "almost upper triangular" and its diagonal entries are rough estimates of the true eigenvalues of $\boldsymbol{R}_i$. This means that only "small rotations" in the Jacobi [89] iteration algorithm are needed to make $\boldsymbol{R}_i'$ upper triangular. In particular, for each i, there exists a close to identity unitary matrix $\boldsymbol{U}_i$ such that,

$$\boldsymbol{R}_i' = \boldsymbol{U}_i^H\boldsymbol{R}_i\boldsymbol{U}_i. \tag{3.33}$$

This means that $\boldsymbol{U}_i$ is a minimal rotational perturbation on $\boldsymbol{Q}_I$ such that $\boldsymbol{Q}_i = \boldsymbol{U}_i\boldsymbol{Q}_I$ triangularizes $\boldsymbol{E}_i$ without affecting the ordering of the diagonal entries. The correct eigenvalue pairs are thus the entries at the same positions on the main diagonals of $\boldsymbol{R}_i, i = 1 \ldots q$.

From the discussion, it is clear that this algorithm is essentially a non-joint method because it first solves an isolated EVD problem $\boldsymbol{E}_I = \boldsymbol{Q}_I^T\boldsymbol{R}_I\boldsymbol{Q}_I$ and only then applies the estimated $\boldsymbol{Q}_I$ to all other matrices. This means that, with this method, a true joint-triangularization of the matrices is not achieved.

3.6.2 The T approach

When the matrices are real valued matrices and $q = 2$, the method proposed in [28,55,102] can be used for the joint diagonalization. To this end, the problem is described as

$$\boldsymbol{E}_1 + j\boldsymbol{E}_2 = \boldsymbol{T}^{-1}(\boldsymbol{\Lambda}_1 + j\boldsymbol{\Lambda}_2)\boldsymbol{T}.$$

(If the $\boldsymbol{T}$ matrix is not of interest, we can alternatively work with the Schur decomposition.) This method works fine most of the cases. However, it does not guarantee that the $\boldsymbol{T}$ matrix is real, which means that the pair $(\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2)$ may not be extracted correctly. A second problem is that the method cannot be used for more than two matrices and it cannot be modified to solve a generalized eigenvalue problem.

3.6.3 Simultaneous Schur Decomposition

In [23, 26], a simultaneous Schur decomposition (SSD) which works in conjunction with the real valued ESPRIT, commonly referred to as the *unitary ESPRIT* (see section 3.8), is discussed. The algorithm tries to minimize the cost function defined as

$$\phi(\boldsymbol{Q}) = \sum_{i=1}^{q} \|\mathcal{L}(\boldsymbol{Q}^T \boldsymbol{E}_i \boldsymbol{Q})\|_F^2$$

where $\mathcal{L}(\cdot)$ is the operator defined earlier. Let $\boldsymbol{Q}_{qp}$ be an elementary (Jacobi) rotation, i.e, $\boldsymbol{Q}_{pq}$ is an orthonormal matrix constructed from the identity matrix where the pp-th and qq-th entries are set to $c = \cos(\varphi)$, and the pq-th and qp-th entries are set to $s = \sin(\varphi)$ and $-s$, respectively. φ is called the rotation angle. At a particular iteration, the SSD method tries to find the rotation angle φ such that the cost function $\phi(\boldsymbol{Q})$ monotonically decreased. The problem associated with this approach is that the choice of the rotation angle at each iteration step does not guarantee convergence to the global minimum. In fact, it is very likely that the algorithm may stall in a local minimum.

3.6.4 The Alternating QZ and The Simultaneous QZ Algorithms

In [87] a JD problem was put as *supper-generalized Schur problem*. There, it was tried to find a joint generalized Schur decomposition of a set of matrices. The idea is, at each iteration step k, to find the set of elementary Jacobi

rotations $(\boldsymbol{Q}^{(k)}, \boldsymbol{Z}^{(k)})$ such that the cost function

$$\phi(\{\boldsymbol{Q}^{(k)}, \boldsymbol{Z}^{(k)}) = \sum_{i=1}^{q} \|\mathcal{L}(\boldsymbol{Q}^{(k)} \boldsymbol{E}_i \boldsymbol{Z}^{(k)})\|_F^2$$

is minimized, where $\mathcal{L}(\cdot)$ is as defined earlier. In this approach, $\boldsymbol{Q}$ and $\boldsymbol{Z}$ are updated alternatively, and thus we refer to it as the *Alternating QZ* algorithm. This is summarized in table 3.2. The alternating QZ gives good estimates of the

Table 3.2: The QZ Joint Diagonalization Procedure

Initialize $\boldsymbol{Q}^{(0)} = \boldsymbol{I}$ and $\boldsymbol{Z}^{(0)} = \boldsymbol{I}$
for $k = 1, 2, \ldots$ do

1. Find the unitary matrix $\boldsymbol{Q}^{(k)}$ to minimize

$$\sum_{i=1}^{q} \|\mathcal{L}\boldsymbol{Q}^{(k)}(\boldsymbol{E}_i \boldsymbol{Z}^{(k-1)})\|_F^2$$

2. Find the unitary matrix $\boldsymbol{Z}^{(k)}$ to minimize

$$\sum_{i=1}^{q} \|\mathcal{L}(\boldsymbol{Q}^{(k)} \boldsymbol{E}_i)\boldsymbol{Z}^{(k)}\|_F^2$$

optimum values of $\boldsymbol{Q}$ and $\boldsymbol{Z}$. However, convergence to the global minimum is not guaranteed and the cost function may not decrease monotonically. In [17] an extension of this method (referred to as the *simultaneous QZ* algorithm) that mitigates some of these problems is discussed. Unlike the alternating QZ, here, $\boldsymbol{Z}^{(k)}$ and $\boldsymbol{Q}^{(k)}$ are updated simultaneously. At each iteration step, a two-sided Jacobi rotation is performed, where the rotation angles are obtained by solving a set of biquadratic equations. The simultaneous QZ can be seen as a generalization of the SSD method described in section 3.6.3. In fact, when $\boldsymbol{Z} = \boldsymbol{Q}$ the simultaneous QZ method reduces to SSD. It, therefore shares the same problem of possible convergence to a local minimum. In [17], however, the author claims that the simultaneous QZ method converges to the true minimum 99% of the times. For determining the global minimum, a heuristic rule was used. According to the rule, the necessary and sufficient condition for attaining the global minimum is that the cost function cannot be further reduced by reinitial-

izing the algorithm with the best outer rotation for any possible rotation pairs (i, j).

Remark

In this thesis, the QZ algorithm (alternating or simultaneous QZ) is used to solve the JD problems. This choice is based mainly on the ease with which it is applied to complex valued data. Thus, unless stated otherwise, all the JD problems are solved using the QZ algorithm.

3.7 3-D ESPRIT (Joint Azimuth, Elevation and Carrier Frequency Estimation)

The joint parameter estimation process can easily be extended to estimating more than two signal parameters [27, 50, 104]. In this section, I take a look how this can be done for the case where we want to jointly estimate azimuth angles, elevation angles and center frequencies of the incoming signals.

The data Model

Consider a uniform circular array consisting of M antenna elements in space as shown in Fig. 3.3. The data model for the 3-D signal representation can easily be generated by appropriately describing the array steering vector $\boldsymbol{a}_i$ for a uniform circular array (UCA). This was derived in section 2.4 considering a two dimensional model. There, assuming the first antenna element to be the reference, the propagation time delay τ_{mi} of the i-th wavefront between the m-th and the reference antenna elements is shown to be as given in (2.22). With the 3-D data model of Fig. 3.3, this is generalized to obtain the following expression, $\tau_{m1} = 0$, and for $i \geq 2$,

$$\tau_{mi} = \left(\frac{\Delta}{f_i} \sin(\alpha_i) \sum_{k=0}^{m-2} \cos(k\frac{2\pi}{M}) + \frac{\Delta}{f_i} \cos(\alpha_i) \sum_{k=0}^{m-2} \sin(k\frac{2\pi}{M}) \right) \sin(\sigma_i),$$

where α_i and σ_i are the azimuth and elevation angles of the i-th incident wavefront, f_i is its center frequency and Δ is the antenna spacing measured in fractions of signal wavelength. For $M = 4$ (4-UCA antenna array), the array steering matrix of the 3-D data model has the same structure as that given in (2.25), but with $\theta_i = e^{-j2\pi\Delta \sin(\alpha_i)\sin(\sigma_i)}$ and $\gamma_i = e^{-j2\pi\Delta \cos(\alpha_i)\sin(\sigma_i)}$,

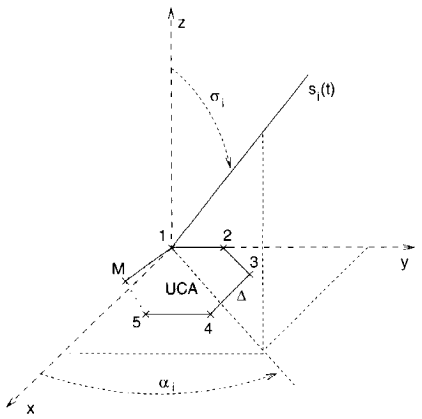


Fig. 3.3: 3-D signal propagation model

Let us now assume that we have collected N samples of the array output at a rate P (normalized w.r.t. signal bandwidth) into a data matrix $\boldsymbol{X}$. After temporal smoothing and forward backward averaging, the extended data matrix $\boldsymbol{X}_m$ is given by (cf. (3.3))

$$\boldsymbol{X}_m = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{A}\boldsymbol{\Phi} \\ \boldsymbol{A}\boldsymbol{\Phi}^2 \\ \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}^{m-1} \end{bmatrix} \boldsymbol{F}_s =: \boldsymbol{A}_m \boldsymbol{F}_s, \tag{3.34}$$

where $\boldsymbol{F}_s$ is as defined in (3.5) and the diagonal matrix $\boldsymbol{\Phi}$ contains the d parameterized frequencies of the d sources. Let $\boldsymbol{J}_{1\phi} = [\boldsymbol{I}_{m-1} \ \ \boldsymbol{0}]$ and $\boldsymbol{J}_{2\phi} = [\boldsymbol{0} \ \ \boldsymbol{I}_{m-1}]$, and $\boldsymbol{J}_{1\theta}, \boldsymbol{J}_{2\theta}, \boldsymbol{J}_{1\gamma}$ and $\boldsymbol{J}_{2\gamma}$ be as given in (2.26). Moreover, let the matrix $\boldsymbol{U}_s$ represent the d dominant left singular vectors of $\boldsymbol{X}_m$. Then, using the selection matrices

$$\begin{aligned} \boldsymbol{J}_{1m}(\theta) &= \boldsymbol{I}_m \otimes \boldsymbol{J}_{1\theta} & \boldsymbol{J}_{2m}(\theta) &= \boldsymbol{I}_m \otimes \boldsymbol{J}_{2\theta} \\ \boldsymbol{J}_{1m}(\gamma) &= \boldsymbol{I}_m \otimes \boldsymbol{J}_{1\gamma} & \boldsymbol{J}_{2m}(\gamma) &= \boldsymbol{I}_m \otimes \boldsymbol{J}_{2\gamma} \\ \boldsymbol{J}_{1m}(\phi) &= J_{1\phi} \otimes \boldsymbol{I}_M & \boldsymbol{J}_{2m}(\phi) &= J_{2\phi} \otimes \boldsymbol{I}_M, \end{aligned} \tag{3.35}$$

we can estimate the triplet $(\gamma_i, \theta_i, \phi_i)$ for all the wave fronts by considering the following eigenvalue decomposition (EVD) problems

$$\begin{aligned} \boldsymbol{E}_\theta &= (\boldsymbol{J}_{1m}(\theta)\boldsymbol{U}_s)^\dagger \boldsymbol{J}_{2m}(\theta)\boldsymbol{U}_s = \boldsymbol{T}\boldsymbol{\Theta}\boldsymbol{T}^{-1} \\ \boldsymbol{E}_\gamma &= (\boldsymbol{J}_{1m}(\gamma)\boldsymbol{U}_s)^\dagger \boldsymbol{J}_{2m}(\gamma)\boldsymbol{U}_s = \boldsymbol{T}\boldsymbol{\Gamma}\boldsymbol{T}^{-1} \\ \boldsymbol{E}_\phi &= (\boldsymbol{J}_{1m}(\phi)\boldsymbol{U}_s)^\dagger \boldsymbol{J}_{2m}(\phi)\boldsymbol{U}_s = \boldsymbol{T}\boldsymbol{\Phi}\boldsymbol{T}^{-1}. \end{aligned} \tag{3.36}$$

Automatic pairing of the parameters is obtained by employing one of the joint diagonalization procedure discussed in section 3.6. The elevation and azimuth angles, and the center frequencies are obtained form γ_i, θ_i and ϕ_i using the following relations,

$$\begin{aligned} \alpha_i &= \operatorname{atan}\left(\frac{\arg(\theta_i)}{\arg(\gamma_i)}\right) \\ \sigma_i &= \operatorname{asin}\left(\frac{|\arg(\theta_i) + j\arg(\gamma_i)|}{2\pi\Delta}\right) \\ f_i &= \frac{P}{2\pi}\arg(\phi_i) \end{aligned}$$

3.8 JAFE and the Unitary ESPRIT

Unitary ESPRIT [24,27] is an extension of the ESPRIT algorithm that achieves a substantial reduction in complexity by mapping the so-called *centro-Hermitian* [43] complex matrices to real matrices. This complex to real transformation has been used by a number of authors [24, 32, 99] to reduce the computational overhead of array processing algorithms. In this section, I give a brief summary of the unitary ESPRIT algorithm, [5] and show how it can be applied in the JAFE context. Since I make repeated use of the operators $\mathbf{\Pi}_p$ and conj$(\cdot)$, at this point, it is useful to recall that $\mathbf{\Pi}_p$ is a $p \times p$ exchange matrix that contains ones in its anti-diagonal and zeros elsewhere and conj$(\cdot)$ denotes entry-wise conjugation.

Definition 3.2 *(*Centro-Hermitian matrix *[30, 43]) A complex $p \times q$ matrix $\boldsymbol{X}$ is said to be centro-Hermitian if it satisfies the condition*

$$conj(\mathbf{\Pi}_p \boldsymbol{X} \mathbf{\Pi}_q) = \boldsymbol{X}$$

Definition 3.3 *(*Left Π-real Matrix *[43]) A $p \times q$ complex matrix $\boldsymbol{Q}$ is said to be left Π-real if it satisfies the relation*

$$conj(\mathbf{\Pi}_p \boldsymbol{Q}) = \boldsymbol{Q} \Leftrightarrow \mathbf{\Pi}_p \boldsymbol{Q} = conj(\boldsymbol{Q}). \tag{3.37}$$

Frequently [52, 102, 103], Π-real matrices are referred to as conjugate centro-symmetric or column conjugate symmetric.

[5] My exposition fully follows that of [27].

3.8.1 Real Valued Invariance Equation

Theorem 3.4 *([27, 43]) Given two left* Π*-real matrices* $\boldsymbol{Q}_p \in \mathbb{C}^{p,p}$ *and* $\boldsymbol{Q}_q \in \mathbb{C}^{q,q}$*, the transformation*

$$\mathcal{T}(\boldsymbol{X}) = \boldsymbol{Q}_p^{-1} \boldsymbol{X} \boldsymbol{Q}_q \tag{3.38}$$

maps any $p \times q$ *centro-Hermitian matrix onto a real matrix of the same size.*

This is a fundamental result that enables us to transform the shift invariance relations from the complex domain to the real domain. In [27], it has been shown that a data matrix collected by a *centro-symmetric* antenna array (see definition 3.1) is centro-Hermitian. Moreover, a centro-Hermitian matrix retains its property after temporal and spatial smoothing. Let us assume now that the antenna array under consideration is centro-symmetric. This means that, the $mM \times N - m + 1$ temporally smoothed data $\boldsymbol{X}_m$ of (3.3) is centro-Hermitian. Thus, for $i = 1, \ldots, N - m + 1$, if $\boldsymbol{X}_i$ represents the i-th column of $\boldsymbol{X}_m$, then there exist two left Π-real matrices $\boldsymbol{Q}_{mM} \in \mathbb{C}^{mM,mM}$ and $\boldsymbol{Q}_2 \in \mathbb{C}^{2,2}$, such that the $mM \times 2$ matrix $\boldsymbol{Z}_i$ constructed as

$$\boldsymbol{Z}_i = \mathcal{T}(\boldsymbol{X}_i) = \boldsymbol{Q}_{mM}^{-1} \left[\begin{array}{cc} \boldsymbol{X}_i & \mathrm{conj}(\boldsymbol{\Pi}_{mM} \boldsymbol{X}_i) \end{array} \right] \boldsymbol{Q}_2 \tag{3.39}$$

is real.

Lemma 3.1 *([27]) Let* $\boldsymbol{Z}_i$ *be as given in (3.39), and the* $mM \times 2(N - m + 1)$ *matrix* $\boldsymbol{Z}$ *be a stack of* $\boldsymbol{Z}_i$*, i.e.,* $\boldsymbol{Z} = \left[\{\boldsymbol{Z}_i\}_{i=1}^{N-m+1} \right]$*. Moreover, let the* d *dominant left singular vectors of* $\boldsymbol{Z}$ *be given by* $\boldsymbol{U}'_s \in \mathbb{R}^{mM,d}$*, then the left* Π*-real matrix*

$$\boldsymbol{U}_s = \boldsymbol{Q}_{mM} \boldsymbol{U}'_s \tag{3.40}$$

is the basis for the estimated signal subspace. More precisely, $\boldsymbol{U}_s$ *corresponds to the* d *dominant left singular vectors of* $\boldsymbol{X}_m$*.*

Theorem 3.5 *([27]) Consider an* $M \times N$ *data matrix* $\boldsymbol{X}$ *collected by a centro-symmetric antenna array. Let* $\boldsymbol{X}_m$ *be a temporally smoothed* $mM \times N - m + 1$ *complex data matrix constructed from* $\boldsymbol{X}$ *as described in section 3.2.1, and the* $mM \times d$ *matrix* $\boldsymbol{U}_s$ *correspond to the* d *dominant left singular vectors of* $\boldsymbol{X}_m$*. Also, let the* $mM \times 2(N - m + 1)$ *real matrix* $\boldsymbol{Z}$ *and the* $mM \times d$ *matrix* $\boldsymbol{U}'_s$ *be as defined in lemma 3.1. Then, the complex-valued invariance equations given in (3.13) can be replaced with their real-valued counter parts*

$$\begin{array}{rcl} \boldsymbol{K}_{1\phi} \boldsymbol{U}'_s & = & \boldsymbol{K}_{2\phi} \boldsymbol{U}'_s \boldsymbol{E}'_\phi \\ \boldsymbol{K}_{1\theta} \boldsymbol{U}'_s & = & \boldsymbol{K}_{2\theta} \boldsymbol{U}'_s \boldsymbol{E}'_\theta \end{array} \tag{3.41}$$

where $\boldsymbol{E}'_\phi$ *and* $\boldsymbol{E}'_\theta$ *are jointly diagonalizable by the same* $d \times d$ *matrix* $\boldsymbol{P}$ [6]*, i.e.,*

$$\begin{aligned} \boldsymbol{E}'_\phi &= \boldsymbol{P}\boldsymbol{\Phi}'\boldsymbol{P}^{-1} \\ \boldsymbol{E}'_\theta &= \boldsymbol{P}\boldsymbol{\Theta}'\boldsymbol{P}^{-1} \end{aligned}$$

and the transformed selection matrices are

$$\begin{aligned} \boldsymbol{K}_{1\theta} &= \boldsymbol{I}_M \otimes \boldsymbol{Q}_m^H(\boldsymbol{J}_{1\theta} + \boldsymbol{J}_{2\theta})\boldsymbol{Q}_M & \boldsymbol{K}_{2\theta} &= \boldsymbol{I}_M \otimes \boldsymbol{Q}_m^H(\boldsymbol{J}_{1\theta} - \boldsymbol{J}_{2\theta})\boldsymbol{Q}_M \\ \boldsymbol{K}_{1\phi} &= \boldsymbol{Q}_m^H(\boldsymbol{J}_{1\phi} + \boldsymbol{J}_{2\phi})\boldsymbol{Q}_M \otimes \boldsymbol{I}_M & \boldsymbol{K}_{1\phi} &= \boldsymbol{Q}_m^H(\boldsymbol{J}_{1\phi} - \boldsymbol{J}_{2\phi})\boldsymbol{Q}_M \otimes \boldsymbol{I}_M \end{aligned} \tag{3.42}$$

Moreover, let the bilinear transformation function $f(x)$ *be defined as*

$$f(x) = \frac{x - j}{x + j},$$

then the least squares solutions of the complex-valued invariance equations $\boldsymbol{\Phi}$ *and* $\boldsymbol{\Theta}$*, and that of the real-valued invariance equation* $\boldsymbol{\Phi}'$ *and* $\boldsymbol{\Theta}'$ *are related via the bilinear transformations*

$$\begin{aligned} \boldsymbol{\Phi} &= f(\boldsymbol{\Phi}') \\ \boldsymbol{\Theta} &= f(\boldsymbol{\Theta}') \end{aligned} \tag{3.43}$$

Putting (3.43) into (3.16), it can be shown that the center frequency f_i and the DOA α_i are obtained from ϕ'_i (the ii-th entry of $\boldsymbol{\Phi}'$) and the θ'_i (the ii-th entry of $\boldsymbol{\Theta}'$), respectively using the following relations.

$$\begin{aligned} \alpha_i &= \operatorname{asin}\left(\frac{\operatorname{atan}(2\theta'_i)}{2\pi\Delta}\right) \\ f_i &= \frac{P}{2\pi}\operatorname{atan}(2\phi'_i). \end{aligned}$$

Since all computations in unitary ESPRIT are conducted on real numbers, it gives approximately a factor four reduction in complexity as compared to the estimation based on complex forward-backward averaged data. The computational efficiency of unitary ESPRIT over the estimation based on the direct complex-valued data is summarized in chapter 7.

3.9 Simulations

In this section I give simulation results comparing the performances of the JAFE algorithm implemented with and without a pre-whitening filter (viz. sec-

[6] Since $\boldsymbol{E}'_\phi$ and $\boldsymbol{E}'_\theta$ are real-valued, the columns of $\boldsymbol{P}$ are either real or appear in complex conjugate. In the latter case, the algorithm is said to be failed the *unitary ESPRIT reliability test [27]* and the results are discarded.

tion 3.4). I consider two equal amplitude, narrowband signals (25 kHz), with center frequencies $(f_1, f_2) = (200, 400)$kHz and DOAs $(\theta_1, \theta_2) = (55, 10)$ degrees, respectively. The sampling rate is 1 MHz (i.e., the IF band has a bandwidth of 1 MHz). The data is collected using a ULA with $M = 8$ antenna elements and the number of temporal samples per data block is $N = 32$. The temporal and spatial smoothing factors are chosen to be $m = 3$ and $L = 3$, respectively. The frequency and DOA estimation errors as functions of SNR are plotted in Fig. 3.4 and Fig. 3.5 (For clarity, only behaviors corresponding to the first source are shown. Similar situations are observed for the second source as well.). From the results one sees that the whitening has very minor effect on the estimation errors. At low SNR regions, a slight performance improvement is observed, (particularly in the frequency estimation). At high SNR, the two behaviors are essentially indistinguishable.

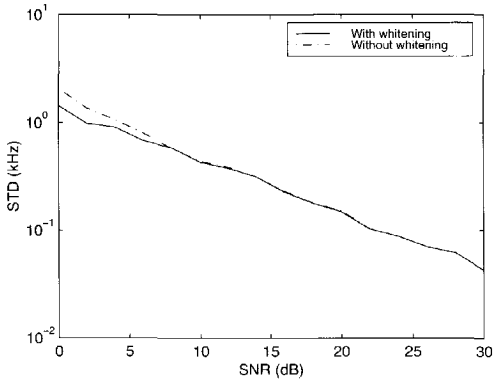


Fig. 3.4: Effect of whitening on the frequency estimation error

3.10 Discussion

In this chapter I have presented the JAFE algorithm. The algorithm exploits the rich data structure resulting from the stacking of the data matrix for the simultaneous estimation of angles and frequencies of narrowband far field signals. I have shown that the data stacking techniques, apart from providing added structures useful for the JAFE, successfully restore the rank of the data matrix when multiple DOAs and multiple frequencies occur mutually exclusively. However, these techniques fail to achieve rank restoration when the multiplicities occur simultaneously. This is because, when the DOAs and center frequencies of two signals coincide, their temporal and spatial behaviors in the context of

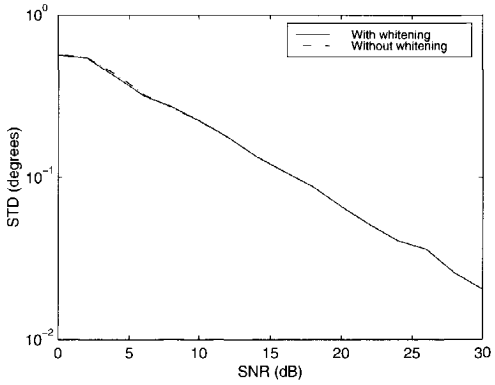


Fig. 3.5: Effect of whitening on the DOA estimation error

the considered data model become indistinguishable. In such cases, additional behaviors, such as constant modulus [87], finite alphabet [53], etc need to be exploited.

I have also shown that the spatio-temporal smoothing parameters m and L can be chosen to maximize the number of estimated signal parameters. However, from the results of the error analysis in chapter 4, it is seen that the optimum data stacking parameters derived here do not minimize the parameter estimation errors. In fact, in chapter4, other optimum values of the data stacking parameters are derived in reference to estimation errors.

It has been shown that the JAFE algorithm is implemented by jointly solving a pair of EVD problems generated by considering a dual shift-invariance structure of a temporally smoothed data matrix. In section 5.1, the multiple-shift invariance property of the data matrix is extended for implementing the multi-resolution ESPRIT algorithm. There, multiple estimates of each signal parameter are systematically combined to obtain a better estimation accuracy.

Chapter 4

A PERFORMANCE ANALYSIS

Contents

In this chapter, I give a behavioral analysis of the JAFE algorithm. As described in section 2.3, the ESPRIT algorithm estimates the parameterized DOAs by solving an eigenvalue problem generated by the shift invariance property. The

algorithm involves three main steps, namely, SVD (singularvalue decomposition) of the data matrix, diagonalization of a set of EVD (eigenvalue decomposition) problems and the transformation of the eigenvalues into signal parameters. The first step, which is equivalent to finding the EVD of the data covariance matrix, is well studied in the literature [1, 10, 36, 38, 44], for the case of white Gaussian noise contaminated data model. In my case, however, since some data stacking techniques have been employed, the noise is no longer white. Thus, first I will derive the eigenvalue estimation error for the JAFE data model and then proceed to the other two JAFE stages. Analysis of the last two steps, in the context of white Gaussian noise, have been presented in [65] and [80]. The results obtained here could be seen as the generalization of these results.

When we assume that pre-whitening has been applied to the data before the application of the JAFE algorithm as discussed in section 3.4, the analysis reduces to the forms similar to those described in [65] and [80]. However, the results in [65] are derived considering a ULA only. Their final result does not give explicit relations between the parameter estimation errors and the noise. In [80] the results of [65] are extended to more general array geometries, and the analysis there is fairly complete. However, the results are derived for DOA estimation only, and they consider a data model without any extension or stacking. Here, I give derivations for both angle and frequency estimations, and I also show how the different data extension procedures affect the estimation performances.

I begin by deriving expressions for the errors on the estimation of the eigenvectors of the covariance matrix (section 4.1). In section 4.2, these results are used in computing the estimation errors of the shift invariance parameters. Finally, in section 4.4, I show how the errors in the actual signal parameters are derived from the errors of the shift invariance parameters.

4.1 Eigenvectors of the Data Covariance Matrix

In this section, I derive the expression for eigenvectors estimation error. Later in section 4.2, I will use these results in deriving the errors in estimating the shift invariance parameters.

The following theorem whose proof is given in appendix A.2, gives the eigenvectors estimation errors for the JAFE data model given in (3.21).

Theorem 4.1 *Consider an M-element antenna array impinged by* d *far field narrowband signals. Let the* (m,L) *factor spatio-temporally smoothed data ma-*

trix $\boldsymbol{X}_{m,L} \in \mathbb{C}^{mM_L \times L(N-m)}$ *be as given in (3.21),* $\boldsymbol{R}_{m,L} = \frac{1}{N}\boldsymbol{X}_{m,L}\boldsymbol{X}_{m,L}^H$ *be the finite sample data covariance matrix, and* $\boldsymbol{U} \in \mathbb{C}^{mM_L, mM_L}$ *and* $\boldsymbol{\Sigma} \in \mathbb{C}^{mM_L, mM_L}$ *be such that the eigenvalue decomposition of* $\bar{\boldsymbol{R}}_{m,L} = E\{\boldsymbol{R}_{m,L}\}$ *is given by*

$$\bar{\boldsymbol{R}}_{m,L} = \boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{U}^H.$$

Let $\boldsymbol{U}_s$ *be the first* d *columns of* $\boldsymbol{U}$*. In chapter 2, it has been shown that for some invertible matrix* $\boldsymbol{T}$*,*

$$\boldsymbol{U}_s = \boldsymbol{A}'_m\boldsymbol{T}^{-1}, \tag{4.1}$$

where $\boldsymbol{A}'_m$ *is as given in (3.21). Now, let* $\boldsymbol{u}_j$ *be the* j*-th eigenvector of* $\bar{\boldsymbol{R}}_{m,L}$*, where* $j \leq d$*. Moreover, let* $\Delta\boldsymbol{u}_j$ *represent a noise caused perturbation on* $\boldsymbol{u}_j$*,* $\sigma_k^2 = \bar{\sigma}_k^2 + \sigma_n^2$ *be the* k*-th eigenvalue of* $\bar{\boldsymbol{R}}_{m,L}$*, where* $\bar{\sigma}_i^2$ *is the* k*-th noise free eigenvalue and* σ_n^2 *is the noise contribution. Let* $\boldsymbol{\Phi} = diag\{\phi_i\}_{i=1}^d$*, and* $\boldsymbol{\Theta} = diag\{\theta_i\}_{i=1}^d$ *be the parameterizations of the center frequencies and direction of arrivals of the* d *signals. Moreover, let*

- $\boldsymbol{Z}^x$ *be a Toeplitz matrix with all the elements equal to zero, except for those unity valued entries on the* x*-th parallel to the main diagonal,*
- $m_o := \min(m, N-m)$ *and* $L_o := \min(M_L, L)$
- $\boldsymbol{Q} := \boldsymbol{T}^{-1}$*,* $\boldsymbol{q}_x^H$ *and* $\boldsymbol{t}_x$ *be the* x*-th row and column of the matrices* $\boldsymbol{Q}^H$ *and* $\boldsymbol{T}^H$*, respectively. Note that* $\boldsymbol{q}_x^H\boldsymbol{t}_y = 0$*, for* $x \neq y$ *and* $\boldsymbol{q}_x^H\boldsymbol{t}_y = 1$ *for* $x = y$*.*
- $\Upsilon(x,y,v,w) := \sum_{h=-m_o}^{m_o}\sum_{r=-L_o}^{L_o} \boldsymbol{u}_x^H\boldsymbol{Z}^{hM_L+r}\boldsymbol{u}_y\boldsymbol{q}_v^H\boldsymbol{\Phi}^h\boldsymbol{\Theta}^r\boldsymbol{t}_w$
- $\Omega(x,y,v,w) := \sum_{h=-m_o}^{m_o}\sum_{r=-L_o}^{L_o} \boldsymbol{u}_x^H\boldsymbol{Z}^{-hM_L-r}\boldsymbol{u}_y\boldsymbol{q}_v^H\boldsymbol{\Phi}^{-h}\boldsymbol{\Theta}^{-r}\boldsymbol{t}_w$
- $\eta(x,y,v,w) := \sum_{h=-m_o}^{m_o}\sum_{r=-L_o}^{L_o} \boldsymbol{u}_x^H\boldsymbol{z}^{hM_L+r}\boldsymbol{u}_y\boldsymbol{u}_v^H\boldsymbol{z}^{-hM_L-r}\boldsymbol{u}_w$

Assuming that $\bar{\boldsymbol{R}}_{m,L}$ *has distinct eigenvalues, the covariance of the eigenvector estimation error* $E\{\Delta\boldsymbol{u}_j\Delta\boldsymbol{u}_p^H\}$ *(to the first order approximation) is given by*

$$E\{\Delta \boldsymbol{u}_v \Delta \boldsymbol{u}_w^H\}_{m,L} \approx \frac{\sigma_n^2}{L(N-m+1)} \times \sum_{\substack{j=1 \\ j\neq v}}^{mM_L} \sum_{\substack{k=1 \\ k\neq w}}^{mM_L} \frac{\Omega(j,k,v,w)\bar{\sigma}_v^2 + \eta(j,k,v,w)\sigma_n^2}{(\sigma_v^2 - \sigma_j^2)(\sigma_w^2 - \sigma_k^2)} \boldsymbol{u}_j \boldsymbol{u}_k^H + \sum_{\substack{j=1 \\ j\neq v}}^{d} \sum_{\substack{k=1 \\ k\neq w}}^{d} \frac{\Upsilon(v,w,j,k)\bar{\sigma}_j^2}{(\sigma_v^2 - \sigma_j^2)(\sigma_w^2 - \sigma_k^2)} \boldsymbol{u}_j \boldsymbol{u}_k^H \tag{4.2}$$

Lemma 4.1 *For whitened spatio-temporally smoothed data matrix, the covariance of the eigenvector estimation error $E\{\Delta \boldsymbol{u}_v \Delta \boldsymbol{u}_w^H\}$ (to the first order approximation) is given by*

$$E\{\Delta \boldsymbol{u}_v \Delta \boldsymbol{u}_w^H\} = 0, \quad \textit{for } w \neq v,$$

and

$$E\{\Delta \boldsymbol{u}_v \Delta \boldsymbol{u}_v^H\}_{m,L} \approx \frac{1}{L(N-m+1)} \times \left(\sum_{\substack{j=1 \\ j\neq v}}^{d} \frac{\sigma_v^2\sigma_j^2 - \bar{\sigma}_v^2\bar{\sigma}_j^2}{(\sigma_v^2 - \sigma_j^2)^2} \boldsymbol{u}_j \boldsymbol{u}_j^H + \sigma_n^2\sigma_v^2 \sum_{j=d+1}^{mM_L} \frac{\boldsymbol{u}_j \boldsymbol{u}_j^H}{(\sigma_v^2 - \sigma_j^2)^2} \right). \tag{4.3}$$

for $w = v$.

Proof:- See Appendix A.3. □

4.2 The Shift Invariance Parameters

Let $\boldsymbol{X}_{m,L} \in \mathbb{C}^{mM_L, L(N-m)}$ and $\boldsymbol{U}_s = \boldsymbol{A}_m \boldsymbol{T}^{-1} \in \mathbb{C}^{mM_L, d}$ be as defined in theorem 4.1. In section 3.2.1 it has been shown that the angle and frequency estimation is attained by considering the dual shift invariance structure present in $\boldsymbol{U}_s$. In this section we take a close look at the behavior of this computation. To this end, as before define two selection matrices $\boldsymbol{J}_x$ and $\boldsymbol{J}_y$, such that the two full column ranked matrices $\boldsymbol{U}_x$ and $\boldsymbol{U}_y$ defined as

$$\begin{aligned} \boldsymbol{U}_x &= \boldsymbol{J}_x \boldsymbol{U}_s \\ \boldsymbol{U}_y &= \boldsymbol{J}_y \boldsymbol{U}_s \end{aligned} \tag{4.4}$$

are such that the $d \times d$ matrix $\boldsymbol{E} = \boldsymbol{U}_x^{\dagger}\boldsymbol{U}_y$ has the eigenvalue decomposition

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{bmatrix} = \boldsymbol{QET} \tag{4.5}$$

$$= \begin{bmatrix} \boldsymbol{q}_1^H \\ \vdots \\ \boldsymbol{q}_d^H \end{bmatrix} \boldsymbol{E} \begin{bmatrix} \boldsymbol{t}_1 & \cdots & \boldsymbol{t}_d \end{bmatrix}, \tag{4.6}$$

where $\boldsymbol{Q} = \boldsymbol{T}^{-1}$ and $\boldsymbol{T}$ is as in (4.1). In the above model, coinciding eigenvalues are allowed, as long as the eigenvectors (the corresponding columns of $\boldsymbol{T}$) are linearly independent. In fact, we require that the eigenvectors be sufficiently distinct,[1] such that after small perturbation they remain to be linearly independent. This assumption is needed because $\boldsymbol{T}^{-1}$ appears in the derivations, and thus we want $\boldsymbol{T}$ to be invertible. From (4.6), it follows that

$$\lambda_i = \boldsymbol{q}_i^H \boldsymbol{E} \boldsymbol{t}_i, \tag{4.7}$$

where $\boldsymbol{q}_i$ and $\boldsymbol{t}_i$ are the left and right eigenvectors of $\boldsymbol{E}$, respectively, i.e., $\boldsymbol{q}_i^H \boldsymbol{E} = \lambda_i \boldsymbol{q}_i^H$ and $\boldsymbol{E}\boldsymbol{t}_i = \lambda_i \boldsymbol{t}_i$. Moreover, for $i, j = 1, 2, \cdots, d$, $\boldsymbol{q}_i$ and $\boldsymbol{t}_j$ satisfy

$$\boldsymbol{q}_i^H \boldsymbol{t}_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}. \tag{4.8}$$

Let $\Delta\lambda_i, \Delta\boldsymbol{q}_i, \Delta\boldsymbol{t}_i$ and $\Delta\boldsymbol{E}$ represent noise caused perturbations on $\lambda_i, \boldsymbol{q}_i, \boldsymbol{t}_i$ and $\boldsymbol{E}$, respectively. We assume that the eigenvectors are sufficiently separated and remain distinct after these small variations. Thus, under noisy situations, (4.7) may be re-written as

$$\lambda_i + \Delta\lambda_i = (\boldsymbol{q}_i^H + \Delta\boldsymbol{q}_i^H)(\boldsymbol{E} + \Delta\boldsymbol{E})(\boldsymbol{t}_i + \Delta\boldsymbol{t}_i).$$

Taking only the linear terms in the above equation, and after some rearrangement of the terms, we obtain

$$\begin{aligned} \Delta\lambda_i &\approx \boldsymbol{q}_i^H \Delta\boldsymbol{E}\boldsymbol{t}_i + \Delta\boldsymbol{q}_i^H \boldsymbol{E}\boldsymbol{t}_i + \boldsymbol{q}_i^H \boldsymbol{E}\Delta\boldsymbol{t}_i \\ &= \boldsymbol{q}_i^H \Delta\boldsymbol{E}\boldsymbol{t}_i + \Delta\boldsymbol{q}_i^H \lambda_i \boldsymbol{t}_i + \lambda_i \boldsymbol{q}_i^H \Delta\boldsymbol{t}_i \\ &= \boldsymbol{q}_i^H \Delta\boldsymbol{E}\boldsymbol{t}_i + \lambda_i(\Delta\boldsymbol{q}_i^H \boldsymbol{t}_i + \boldsymbol{q}_i^H \Delta\boldsymbol{t}_i). \end{aligned} \tag{4.9}$$

Since $\boldsymbol{Q} = \boldsymbol{T}^{-1}$, the noise terms $\Delta\boldsymbol{q}_i$ and $\Delta\boldsymbol{t}_i$ are not independent. Their relation is derived by noting that (4.8) is valid under noisy conditions as well, i.e.,

$$(\boldsymbol{q}_i^H + \Delta\boldsymbol{q}_i^H)(\boldsymbol{t}_j + \Delta\boldsymbol{t}_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases},$$

[1] Note that the eigensubspaces belonging to such a multiple eigenvalue do not have unique eigenvectors. Thus, orthogonal basis of eigenvectors in that subspace can be chosen.

Taking the first order terms only, this simplifies to

$$\begin{aligned} \boldsymbol{q}_i^H \boldsymbol{t}_i + \boldsymbol{q}_i^H \Delta\boldsymbol{t}_i + \Delta\boldsymbol{q}_i^H \boldsymbol{t}_i &\approx 1 \\ \boldsymbol{q}_i^H \Delta\boldsymbol{t}_i + \Delta\boldsymbol{q}_i^H \boldsymbol{t}_i &\approx 0. \end{aligned}$$

This means that the second term in (4.9) is approximately zero, and the first order approximations of the errors on the eigenvalues are given by (viz. [65])

$$\Delta\lambda_i \approx \boldsymbol{q}_i^H \Delta\boldsymbol{E}\boldsymbol{t}_i. \tag{4.10}$$

Let $\Delta\boldsymbol{U}_x$ represent the noise perturbation on $\boldsymbol{U}_x$, and similarly for $\Delta\boldsymbol{U}_y$ and $\boldsymbol{U}_y$, then an expression for $\Delta\boldsymbol{E}$ is derived by noting that

$$\boldsymbol{E} + \Delta\boldsymbol{E} = (\boldsymbol{U}_x + \Delta\boldsymbol{U}_x)^\dagger (\boldsymbol{U}_y + \Delta\boldsymbol{U}_y). \tag{4.11}$$

If the perturbation is small enough, the first term in the above equation can be approximated (up to first order) as

$$(\boldsymbol{U}_x + \Delta\boldsymbol{U}_x)^\dagger \approx (\boldsymbol{I} - \boldsymbol{U}_x^\dagger \Delta\boldsymbol{U}_x)\boldsymbol{U}_x^\dagger.$$

Putting this into (4.11) and taking the linear terms only, we get the following approximation for $\Delta\boldsymbol{E}$

$$\Delta\boldsymbol{E} \approx \boldsymbol{U}_x^\dagger (\Delta\boldsymbol{U}_y - \Delta\boldsymbol{U}_x \boldsymbol{E}).$$

Using this relation, and noting that $\boldsymbol{E}\boldsymbol{t}_i = \lambda_i \boldsymbol{t}_i$, $\Delta\boldsymbol{U}_x = \boldsymbol{J}_x \Delta\boldsymbol{U}_s$ and $\Delta\boldsymbol{U}_y = \boldsymbol{J}_y \Delta\boldsymbol{U}_s$, the expression for $\Delta\lambda_i$ in (4.10) becomes

$$\Delta\lambda_i \approx \boldsymbol{q}_i^H \boldsymbol{U}_x^\dagger (\boldsymbol{J}_y - \lambda_i \boldsymbol{J}_x) \Delta\boldsymbol{U}_s \boldsymbol{t}_i$$

Note that, $\boldsymbol{U}_s$ and the selection matrices have dimensions that are functions of the spatio-temporal smoothing factors m and L. In the following, these dependencies are made explicit using indexed references. Thus, put $\boldsymbol{r}_{m,L}^H(\lambda_i) = \boldsymbol{q}_i^H \boldsymbol{U}_x^\dagger (\boldsymbol{J}_y - \lambda_i \boldsymbol{J}_x)$, and let $\boldsymbol{u}_j$ and $\Delta\boldsymbol{u}_j$ be such that $\boldsymbol{U}_s = [\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_d]$ and $\Delta\boldsymbol{U}_s = [\Delta\boldsymbol{u}_1 \ \cdots \ \Delta\boldsymbol{u}_d]$. With these definitions, the error $\Delta\lambda_i$ may be written as $\Delta\lambda_i \approx \boldsymbol{r}_{m,L}^H(\lambda_i) \Delta\boldsymbol{U}_s \boldsymbol{t}_i$ and its mean square value as

$$\begin{aligned} \sigma_{m,L}(\lambda_i) := E\{|\Delta\lambda_i|^2\} &\approx \boldsymbol{r}_{m,L}^H(\lambda_i) E\{\Delta\boldsymbol{U}_s \boldsymbol{t}_i \boldsymbol{t}_i^H \Delta\boldsymbol{U}_s^H\} \boldsymbol{r}_{m,L}(\lambda_i) \\ &= \boldsymbol{r}_{m,L}^H(\lambda_i) E\left\{ \sum_{j=1}^{d} \sum_{k=1}^{d} \Delta\boldsymbol{u}_j t_{ji} t_{ki}^H \Delta\boldsymbol{u}_k^H \right\} \boldsymbol{r}_{m,L}(\lambda_i), \end{aligned}$$

where t_{ji} is the j-th entry of the column vector $\boldsymbol{t}_i$ defined earlier. Noting that t_{ji} is noise independent, the above can be re-written as

$$\sigma_{m,L}(\lambda_i) = \boldsymbol{r}_{m,L}^H(\lambda_i) \left(\sum_{j=1}^{d} \sum_{k=1}^{d} t_{ji} t_{ki}^H E\{\Delta\boldsymbol{u}_j \Delta\boldsymbol{u}_k^H\} \right) \boldsymbol{r}_{m,L}(\lambda_i). \tag{4.12}$$

Further simplification of the above expression is conveniently obtained for a whitened data matrix. In the remainder of this chapter, I will assume that a whitening has been applied to the data matrix prior to the application of the algorithm. In this case the eigenvector estimation error $E\{\Delta \boldsymbol{u}_j \Delta \boldsymbol{u}_k^H\}$ is as given in (4.3). Putting this into the above expression, it follows that

$$\sigma_{m,L}(\lambda_i) \approx \boldsymbol{r}_{m,L}^H(\lambda_i) \sum_{j=1}^{d} |t_{ji}|^2 \frac{1}{L(N-m+1)} \times \left(\sum_{\substack{k=1 \\ k \neq j}}^{d} \frac{\sigma_j^2 \sigma_k^2 - \bar{\sigma}_j^2 \bar{\sigma}_k^2}{(\sigma_j^2 - \sigma_k^2)^2} \boldsymbol{u}_k \boldsymbol{u}_k^H + (\sigma_j^2 \sigma_n^2) \sum_{k=d+1}^{mM_L} \frac{\boldsymbol{u}_k \boldsymbol{u}_k^H}{(\sigma_j^2 - \sigma_n^2)^2} \right) \boldsymbol{r}_{m,L}(\lambda_i). \tag{4.13}$$

Let the left eigenvectors of the data covariance matrix be partitioned into $\boldsymbol{U}_s$ and $\boldsymbol{U}_n$, such that $\boldsymbol{U}_s$ spans the signal subspace and $\boldsymbol{U}_n$ spans the noise subspace, and let the (punctured) diagonal matrix $\boldsymbol{\Gamma}_j$, with a zero at the jj-th position, be defined as

$$\boldsymbol{\Gamma}_j = \begin{bmatrix} \operatorname{diag}\left\{ \frac{\sigma_j^2 \sigma_k^2 - \bar{\sigma}_j^2 \bar{\sigma}_k^2}{(\sigma_j^2 - \sigma_k^2)} \right\}_{k=1}^{j-1} & & \\ & 0 & \\ & & \operatorname{diag}\left\{ \frac{\sigma_j^2 \sigma_k^2 - \sigma_j^2 \bar{\sigma}_k^2}{(\sigma_j^2 - \sigma_k^2)} \right\}_{k=j+1}^{d} \end{bmatrix}.$$

Then, we can re-write (4.13) in a simplified way as

$$\sigma_{m,L}(\lambda_i) \approx \boldsymbol{r}_{m,L}^H(\lambda_i) \left(\sum_{j=1}^{d} |t_{ji}|^2 \frac{1}{L(N-m+1)} \times \left[\boldsymbol{U}_s \boldsymbol{\Gamma}_j \boldsymbol{U}_s^H + \frac{\sigma_j^2 \sigma_n^2}{(\sigma_j^2 - \sigma_n^2)^2} \boldsymbol{U}_n \boldsymbol{U}_n^H \right] \right) \boldsymbol{r}_{m,L}(\lambda_i). \tag{4.14}$$

Note that in the above discussions we have not made any assumption on the array geometry. The geometry information is contained in $\boldsymbol{r}_{m,L}^H(\lambda_i) = \boldsymbol{q}_i^H \boldsymbol{U}_x^\dagger (\boldsymbol{J}_y - \lambda_i \boldsymbol{J}_x)$, referred to as the *array geometry parameter*. Let $\boldsymbol{A}_{m,x} = \boldsymbol{J}_x \boldsymbol{A}'_m$ and $\boldsymbol{A}_{m,y} = \boldsymbol{J}_y \boldsymbol{A}'_m$, where $\boldsymbol{A}'_m$ is as given in (3.21), and let $\boldsymbol{w}_{m,i}^H$ be the i-th row of $\boldsymbol{A}_{m,x}^\dagger$. Then, noting that $\boldsymbol{U}_x = \boldsymbol{A}_{m,x} \boldsymbol{T}^{-1}$, where $\boldsymbol{T}$ is as given in (4.5), the array geometry parameter may be expressed as

$$\begin{aligned} \boldsymbol{r}_{m,L}^H(\lambda_i) &= \boldsymbol{q}_i^H \boldsymbol{T} \boldsymbol{A}_{m,x}^\dagger (\boldsymbol{J}_y - \lambda_i \boldsymbol{J}_x) \\ &= \boldsymbol{w}_{m,i}^H (\boldsymbol{J}_y - \lambda_i \boldsymbol{J}_x), \end{aligned} \tag{4.15}$$

A further simplification of (4.14) is obtained by noting $\boldsymbol{A}_{m,y} = \boldsymbol{A}_{m,x}\boldsymbol{\Lambda}$ and the following fact (viz. [80]):

$$\begin{aligned}
\boldsymbol{r}_{m,L}^H(\lambda_i)\boldsymbol{U}_s &= \boldsymbol{r}_{m,L}^H(\lambda_i)\boldsymbol{A}_m'\boldsymbol{T}^{-1} \\
&\approx \boldsymbol{q}_i^H\boldsymbol{T}\boldsymbol{A}_{m,x}^\dagger(\boldsymbol{J}_y - \lambda_i\boldsymbol{J}_x)\boldsymbol{A}_m'\boldsymbol{T}^{-1} \\
&= \boldsymbol{q}_i^H\boldsymbol{T}\boldsymbol{A}_{m,x}^\dagger(\boldsymbol{A}_{m,y} - \lambda_i\boldsymbol{A}_{m,x})\boldsymbol{T}^{-1} \\
&= \boldsymbol{q}_i^H\boldsymbol{T}\boldsymbol{A}_{m,x}^\dagger(\boldsymbol{A}_{m,x}\boldsymbol{\Lambda} - \lambda_i\boldsymbol{A}_{m,x})\boldsymbol{T}^{-1} \\
&= \boldsymbol{q}_i^H\boldsymbol{T}\boldsymbol{A}_{m,x}^\dagger\boldsymbol{A}_{m,x}(\boldsymbol{\Lambda} - \lambda_i\boldsymbol{I})\boldsymbol{T}^{-1} \\
&= \boldsymbol{q}_i^H\boldsymbol{T}(\boldsymbol{\Lambda} - \lambda_i\boldsymbol{I})\boldsymbol{T}^{-1} = [0\ \cdots\ 0],
\end{aligned}$$

From this, it follows that the first term in (4.14) vanishes and, with the parameter η_i defined as

$$\frac{1}{\eta_i} = \sum_{j=1}^{d} |t_{ji}|^2 \frac{\sigma_j^2}{(\sigma_j^2 - \sigma_n^2)^2}, \tag{4.16}$$

we get the following expression for $\sigma_{m,L}(\lambda_i)$:

$$\sigma_{m,L}(\lambda_i) \approx \frac{\sigma_n^2/\eta_i}{L(N-m+1)} \boldsymbol{r}_{m,L}^H(\lambda_i)\boldsymbol{U}_n\boldsymbol{U}_n^H\boldsymbol{r}_{m,L}(\lambda_i). \tag{4.17}$$

Let $\rho_{m,L}(\lambda_i)$ be defined as

$$\rho_{m,L}(\lambda_i) = \boldsymbol{r}_{m,L}^H(\lambda_i)\boldsymbol{r}_{m,L}(\lambda_i), \tag{4.18}$$

then, noting that $\boldsymbol{r}_{m,L}^H(\lambda_i)\boldsymbol{U}_n\boldsymbol{U}_n^H = \boldsymbol{r}_{m,L}^H(\lambda_i)(\boldsymbol{I} - \boldsymbol{U}_s\boldsymbol{U}_s^H) = \boldsymbol{r}_{m,L}^H(\lambda_i)$, (4.17) simplifies to

$$\sigma_{m,L}(\lambda_i) \approx \frac{1}{L(N-m+1)} \frac{\sigma_n^2}{\eta_i} \rho_{m,L}(\lambda_i). \tag{4.19}$$

Theorem 4.2 *Suppose that an M element antenna array, that is split into L subarrays each with M_L elements, is impinged by $d < M_L$ narrowband signals. Let $\boldsymbol{X} \in \mathbb{C}^{mM_L, L(N-m)}$ be an (m, L)-factor spatio-temporally smoothed data matrix collected at the output of the array, σ_n^2 be the power of the noise, and $\boldsymbol{U}_s \in \mathbb{C}^{mM_L, d}$ be the d dominant eigenvectors of the data covariance matrix. Now, let $\boldsymbol{U}_x$ and $\boldsymbol{U}_y$ be two submatrices of $\boldsymbol{U}_s$ such that, for some non singular $d \times d$ matrix $\boldsymbol{T}$ and a $d \times d$ non singular diagonal matrix $\boldsymbol{\Lambda}$, $\boldsymbol{U}_x^\dagger\boldsymbol{U}_y = \boldsymbol{T}^{-1}\boldsymbol{\Lambda}\boldsymbol{T}$. Let λ_i be the i-th entry of $\boldsymbol{\Lambda}$ and $\sigma_{m,L}(\lambda_i)$ be the variance of the estimation error on λ_i. Then the dependence of $\sigma_{m,L}(\lambda_i)$ on the array geometry and the spatio-temporal smoothing factors is completely described by the factor*

$$G(m, L) := \frac{1}{L(N-m)} \rho_{m,L}(\lambda_i).$$

Proof:- For the condition stated in the theorem, it has already been shown that $\sigma_{m,L}(\lambda_i)$ is given by (4.19). Thus, for the proof, it is sufficient to show that η_i defined in (4.16) is independent of the array geometry, and the factors m and L. To this end, let the $Mm \times d$ matrix $\boldsymbol{A}_m$ be the *extended* array steering matrix associated with a data matrix with a temporal smoothing factor of m, $\boldsymbol{U}_s$ be a unitary matrix that spans the column space of the extended data matrix, and $\boldsymbol{\Sigma}_s$ be a diagonal matrix that contains the d largest singular values of the extended data matrix. Then there exists a non singular $d \times d$ matrix $\boldsymbol{T}$, such that the data covariance matrix $\boldsymbol{R}_x = \frac{1}{L(N-m+1)} \boldsymbol{X}\boldsymbol{X}^H$ may be expressed as

$$\begin{aligned} \boldsymbol{U}_s \boldsymbol{\Sigma}_s^2 \boldsymbol{U}_s^H = \boldsymbol{R}_x &= \boldsymbol{A}_m \boldsymbol{R}_s \boldsymbol{A}_m^H + \sigma_n^2 \boldsymbol{I} \\ &= \boldsymbol{U}_s \boldsymbol{T} \boldsymbol{R}_s \boldsymbol{T}^H \boldsymbol{U}_s^H + \sigma_n^2 \boldsymbol{I}, \end{aligned}$$

where $\boldsymbol{R}_s$ is the signal covariance matrix. Solving for $\boldsymbol{R}_s^{-1}$, we obtain

$$\boldsymbol{R}_s^{-1} = \boldsymbol{T}^H (\boldsymbol{\Sigma}_m^2 - \sigma_n^2 \boldsymbol{I})^{-1} \boldsymbol{T}.$$

and thus, with t_{ji} equal to the ji-th entry of $\boldsymbol{T}$,

$$\{\boldsymbol{R}_s^{-1}\}_{ii} = \sum_{j=1}^{d} |t_{ji}|^2 \frac{1}{\sigma_j^2 - \sigma_n^2}.$$

Note that the signal covariance matrix (the left hand expression) is independent of the array geometry. This means that the right hand summation, and therefore, η_i is also independent of the array geometry and of m. In a similar way, one can show that η_i is also independent of the spatial smoothing factor L. Note that for good SNR, $\sigma_j \gg \sigma_n$ for all $j \leq d$, in which case $\eta_i \approx 1/\{\boldsymbol{R}_s^{-1}\}_{ii}$, and hence is independent of m and L. □

This result is useful when we consider the effects of m and L on the estimation errors because the dependency of $\sigma_{m,L}(\lambda_i)$ on these parameters is completely described by the less complex factor $\rho_{m,L}(\lambda_i)$. To emphasize the fact that the geometry information is fully described by $\rho_{m,L}(\lambda_i)$, in the sequel, it will be referred to as the *geometric factor*. Moreover, since η_i is dependent purely on the signal covariance matrix and the SNR, it is referred to as the *signal factor*, and the ratio $\hat{\text{SNR}} = \eta_i/\sigma_n^2$ is termed as the *effective SNR*.

In the foregoing discussions, we have not made any reference to the parameter to be estimated. The analysis up to now, therefore, applies for both the parameterized DOA and frequency estimations alike. The distinction comes in the way the selection matrices are defined.

4.2.1 The Parameterized DOA Estimation

In line with the discussions in the above sub-section, the parameterized DOA estimation error $\sigma_{m,L}(\theta_i)$ is obtained from (4.19), by replacing λ_i with θ_i and $\rho_{m,L}(\lambda_i)$ with $\rho_{m,L}(\theta_i)$:

$$\sigma_{m,L}(\theta_i) \approx \frac{1}{L(N-m+1)} \frac{\sigma_n^2}{\eta_i} \rho_{m,L}(\theta_i) \tag{4.20}$$

Here θ_i is the i-th parameterized DOA defined in section 2.1 and $\rho_{m,L}(\theta_i)$ is the corresponding array factor constructed using the selection matrices $\boldsymbol{J}_x(\theta)$ and $\boldsymbol{J}_y(\theta)$. In this section, I present the analysis where $m = 2$,[2] and $L = 1$, in which case the column span of the data matrix is given by

$$\boldsymbol{A}_m = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{A}\boldsymbol{\Phi} \end{bmatrix}. \tag{4.21}$$

Here, $\boldsymbol{A}$ represents the array steering matrix. For a ULA and $(m, L) = (2, 1)$, the selection matrices $\boldsymbol{J}_x(\theta)$ and $\boldsymbol{J}_y(\theta)$ select the first and respectively the last $M-1$ rows from each of the two block entries of $\boldsymbol{A}_m$. Thus,

$$\boldsymbol{J}_y(\theta) - \theta_i \boldsymbol{J}_x(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \boldsymbol{\Upsilon}(\theta_i),$$

where

$$\boldsymbol{\Upsilon}(\theta_i) = \begin{bmatrix} -\theta_i & 1 & & & \\ & -\theta_i & 1 & & \\ & & \ddots & \ddots & \\ & & & -\theta_i & 1 \end{bmatrix}. \tag{4.22}$$

Let $\boldsymbol{A}_x$ denote the first $M-1$ rows of $\boldsymbol{A}$ and $\boldsymbol{A}_{m,x}$ be defined as

$$\boldsymbol{A}_{m,x} = \begin{bmatrix} \boldsymbol{A}_x \\ \boldsymbol{A}_x\boldsymbol{\Phi} \end{bmatrix}.$$

It follows that $\boldsymbol{W}_{m,x} = \frac{1}{\sqrt{2}}[\boldsymbol{A}_x^\dagger \;\; \boldsymbol{\Phi}^{-1}\boldsymbol{A}_x^\dagger]$ is a left inverse of $\boldsymbol{A}_{m,x}$ and, if we let $\boldsymbol{w}_i^H = [w_{i,1} \;\cdots\; w_{i,M-1}]$ represent the i-th row of $\boldsymbol{A}_x^\dagger$, then the i-th row of $\boldsymbol{W}_{m,x}$ is given by $\boldsymbol{w}_{m,i}^H = \frac{1}{\sqrt{2}}[\boldsymbol{w}_i^H \;\; \phi_i^{-1}\boldsymbol{w}_i^H]$. Thus, using (4.15) and (4.18), it follows that

$$\rho_{2,1}(\theta_i) = \frac{1}{2}\left(|w_{i,1}|^2 + |w_{i,M-1}|^2 + \sum_{k=2}^{M-1} |w_{i,k-1} - \theta_i w_{i,k}|^2 \right). \tag{4.23}$$

[2]Note that for JAFE, the minimum possible value of m is 2 and that of L is 1. Behaviors corresponding to larger m and L values are considered in section 4.3.

For a *single* source, $w_{1k} = \frac{1}{M-1}\theta_1^{-k}$ and thus $\rho_{2,1}(\theta_1) = 1/(M-1)^2$. Moreover, $\hat{\text{SNR}}_1 = \eta_1/\sigma_n^2 = \text{SNR}_1$. Combining these and using (4.20), we get the following approximation for the estimation error

$$\sigma_{2,1}(\theta_1) \approx \frac{1}{(M-1)^2(N-1)} \frac{1}{\text{SNR}_1}. \tag{4.24}$$

If we have started by setting $m = 1$, we get the expression,

$$\sigma_{2,1}(\theta_1) \approx \frac{2}{(M-1)^2 N} \frac{1}{\text{SNR}_1}, \tag{4.25}$$

which agrees with those described in [65] and [80] for a single source scenario. It is seen that the DOA estimation error is proportional to the inverse of the square of the number of antennas. This means that the algorithm, for large M, fails to achieve the CRB (Cramer-Rao Lower Bound) [3] (i.e., it is inefficient). However, in section 4.3, it will be shown that, using data extension procedures, the algorithm can be made efficient.

4.2.2 The Parameterized Frequency Estimation

For the frequency estimation, with reference to the discussions in section 4.2, the parameterized frequency estimation error $\sigma_{m,L}(\phi_i)$ is obtained from (4.19) by replacing λ_i with ϕ_i and the geometric factor $\rho_{m,L}(\lambda_i)$ with $\rho_{m,L}(\phi_i)$:

$$\sigma_{m,L}(\phi_i) \approx \frac{1}{L(N-m+1)} \frac{\sigma_n^2}{\eta_i} \rho_{m,L}(\phi_i), \tag{4.26}$$

where $\rho_{m,L}(\phi_i)$ is defined using the selection matrices $\boldsymbol{J}_x(\phi)$ and $\boldsymbol{J}_y(\phi)$. For a ULA, these selection matrices are given in (3.9). In the following, I give a performance analysis for the case $m = 2$ and $L = 1$. The effects of other (m, L) values is considered in section 4.3. For $(m, L) = (2, 1)$, the column span of the data matrix is given by (4.21). For the parameterized frequency estimation, the selection matrix $\boldsymbol{J}_x(\phi)$ selects the first M rows of $\boldsymbol{A}_m$ (which is $\boldsymbol{A}$) and $\boldsymbol{J}_y(\phi)$ its last M rows. Thus,

$$\boldsymbol{J}_y(\phi) - \phi_i \boldsymbol{J}_x(\phi) = \begin{bmatrix} -\phi_i \boldsymbol{I} & \boldsymbol{I} \end{bmatrix}.$$

Let $\boldsymbol{w}_i^H$ be the i-th row of $\boldsymbol{A}^\dagger$, then using (4.15), it follows that $\boldsymbol{r}_{2,1}^H(\phi_i) = [-\phi_i \boldsymbol{w}_i^H \quad \boldsymbol{w}_i^H]$,[4] and hence,

$$\rho_{2,1}(\phi_i) = \boldsymbol{r}_{2,1}^H(\phi_i)\boldsymbol{r}_{2,1}(\phi_i) = 2|\boldsymbol{w}_i|^2$$

[3] In CRB, the angle estimation error is proportional to M^{-3}

[4] Note that, in contrast to $\boldsymbol{J}_x(\theta)$ and $\boldsymbol{J}_y(\theta)$, the structures of the selection matrices $\boldsymbol{J}_x(\phi)$ and $\boldsymbol{J}_y(\phi)$ are independent of array geometry, and thus this result applies to an arbitrary array geometry.

and

$$\sigma_{2,1}(\phi_i) \approx \frac{2|\boldsymbol{w}_i|^2}{N-1}\frac{\sigma_n^2}{\eta_i}.$$

For a *single* source, $|\boldsymbol{w}_1|^2 = 1/M$, and $\hat{\text{SNR}}_1 = \eta_1/\sigma_n^2 \approx \text{SNR}_1$, thus

$$\sigma_{2,1}(\phi_1) \approx \frac{2}{M(N-1)}\frac{1}{\text{SNR}_1}.$$

It is seen that the estimation error, for $m = 2$, decays only in linear proportion with N. This is, of course, an extremely poor result. However, it is important to note that, for $m = 2$, the effective number of temporal samples used in actual phase computation is 2. This means that by choosing larger m values, the performance can be improved significantly. In fact, in the following section, it will be shown that, by choosing an appropriate value for m, the frequency estimation error can be made to decay in proportion to N^{-3}.

4.3 Effects of Data Extensions

The performance analysis outlined in the above section considers the data model of (3.21), with a spatial smoothing factor $L = 1$ and a temporal smoothing factor $m = 2$. In this section, I give analysis of how the data extension procedures affect the estimation performances. Moreover, I derive the optimum values of L and m, (L_o, m_o), that minimize the angle, frequency and joint estimation errors.

Consider an antenna array with an arbitrary geometry. Let for an (m, L) spatio-temporally smoothed data matrix, the pairs $(\boldsymbol{J}_x(\phi), \boldsymbol{J}_y(\phi))$ and $(\boldsymbol{J}_x(\theta), \boldsymbol{J}_y(\theta))$ be the selection matrices that produce the shift invariance pairs for the parameterized frequency and DOA estimations, respectively, then

$$\begin{cases} \boldsymbol{J}_x(\phi) := [\boldsymbol{I}_{m-1} \quad \boldsymbol{0}_1] \otimes \boldsymbol{I}_{M_L} \\ \boldsymbol{J}_y(\phi) := [\boldsymbol{0}_1 \quad \boldsymbol{I}_{m-1}] \otimes \boldsymbol{I}_{M_L}, \end{cases} \tag{4.27}$$

$$\begin{cases} \boldsymbol{J}_x(\theta) := \boldsymbol{I}_m \otimes \boldsymbol{J}_{1,x}(\theta) \\ \boldsymbol{J}_x(\theta) := \boldsymbol{I}_m \otimes \boldsymbol{J}_{1,y}(\theta). \end{cases} \tag{4.28}$$

where $\boldsymbol{J}_{1,x}(\theta)$ and $\boldsymbol{J}_{1,y}(\theta)$ are the selection matrices for the case $(m, L) = (1, 1)$ and are array geometry dependent. For a ULA, they are given by

$$\boldsymbol{J}_{1,x}(\theta) = [\boldsymbol{I}_{M-1} \quad \boldsymbol{0}_1]$$
$$\boldsymbol{J}_{1,y}(\theta) = [\boldsymbol{0}_1 \quad \boldsymbol{I}_{M-1}].$$

Now, let $\boldsymbol{U}_s$ be a unitary matrix that spans the column space of the extended data matrix, then for some invertible matrix $\boldsymbol{T}$, the matrices $\boldsymbol{A}_\phi := \boldsymbol{J}_x(\phi)\boldsymbol{U}_s\boldsymbol{T}^{-1}$

and $\boldsymbol{A}_\theta := \boldsymbol{J}_x(\theta)\boldsymbol{U}_s\boldsymbol{T}^{-1}$ are given by

$$\boldsymbol{A}_\phi = \begin{bmatrix} \boldsymbol{A}(L) \\ \boldsymbol{A}(L)\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{A}(L)\boldsymbol{\Phi}^{m-2} \end{bmatrix} \quad \text{and} \quad \boldsymbol{A}_\theta = \begin{bmatrix} \boldsymbol{A}'(L) \\ \boldsymbol{A}'(L)\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{A}'(L)\boldsymbol{\Phi}^{m-1} \end{bmatrix},$$

respectively, where $\boldsymbol{A}(L)$ is the array steering matrix corresponding to the L-factor spatially smoothed data matrix and $\boldsymbol{A}'(L)$ is a submatrix of $\boldsymbol{A}(L)$, whose row dimension depends on how the spatial smoothing is performed.

Let $\boldsymbol{w}_{\phi_i}^H$ and $\boldsymbol{w}_{\theta_i}^H$ be the i-th rows of $\boldsymbol{W}_\phi = \boldsymbol{A}_\phi^\dagger$ and $\boldsymbol{W}_\theta = \boldsymbol{A}_\theta^\dagger$, respectively, and let the bi-diagonal matrix $\boldsymbol{\Upsilon}(\phi_i)$ be defined as

$$\boldsymbol{\Upsilon}(\phi_i) = \begin{bmatrix} -\phi_i & 1 & & & \\ & -\phi_i & 1 & & \\ & & \ddots & \ddots & \\ & & & -\phi_i & 1 \end{bmatrix}. \tag{4.29}$$

Then, using the selection matrices defined in (4.27) and (4.28), it can be shown that

$$\begin{aligned} \boldsymbol{r}_{m,L}^H(\theta_i) &= \boldsymbol{w}_{\theta_i}^H [\boldsymbol{I}_m \otimes \boldsymbol{\Upsilon}(\theta_i)] \\ \boldsymbol{r}_{m,L}^H(\phi_i) &= \boldsymbol{w}_{\phi_i}^H [\boldsymbol{\Upsilon}(\phi_i) \otimes \boldsymbol{I}_{M_L}]. \end{aligned} \tag{4.30}$$

Note that, while $\boldsymbol{\Upsilon}(\theta_i)$ is array geometry dependent, $\boldsymbol{\Upsilon}(\phi_i)$ is independent of geometry. It is seen that, for ULA these matrices have the same structure (compare (4.29) with (4.22)). Let now $\boldsymbol{W}(L) = \boldsymbol{A}^\dagger(L)$ and $\boldsymbol{W}'(L) = \boldsymbol{A}'^\dagger(L)$, then it is seen that the two matrices

$$\boldsymbol{W}_m = \frac{1}{m-1} \begin{bmatrix} \boldsymbol{W}(L) & \boldsymbol{\Phi}^{-1}\boldsymbol{W}(L) & \cdots & \boldsymbol{\Phi}^{2-m}\boldsymbol{W}(L) \end{bmatrix}$$

and

$$\boldsymbol{W}'_m = \frac{1}{m} \begin{bmatrix} \boldsymbol{W}'(L) & \boldsymbol{\Phi}^{-1}\boldsymbol{W}'(L) & \cdots & \boldsymbol{\Phi}^{1-m}\boldsymbol{W}'(L) \end{bmatrix}$$

are left inverses of $\boldsymbol{A}_\phi$ and $\boldsymbol{A}_\theta$, respectively. I.e.,

$$\boldsymbol{W}_m\boldsymbol{A}_\phi = \boldsymbol{I} \quad \text{and} \quad \boldsymbol{W}'_m\boldsymbol{A}_\theta = \boldsymbol{I}.$$

It is well known, however, that the Moore-Penrose inverse (the pseudo-inverse) gives unique left inverses of these matrices with minimum (Frobenius) norms. Thus, with $\boldsymbol{W}_\phi = \boldsymbol{A}_\phi^\dagger$ and $\boldsymbol{W}_\theta = \boldsymbol{A}_\theta^\dagger$,

$$\|\boldsymbol{W}_\phi\|^2 \le \|\boldsymbol{W}_m\|^2 = \frac{\|\boldsymbol{W}(L)\|^2}{m-1} \quad \text{and}$$

$$\|\boldsymbol{W}_\theta\|^2 \le \|\boldsymbol{W}'_m\|^2 = \frac{\|\boldsymbol{W}'(L)\|^2}{m}.$$

If the underlying sources are *sufficiently separated*, $\boldsymbol{W}_\phi \approx \boldsymbol{W}_m$ and $\boldsymbol{W}_\theta \approx \boldsymbol{W}'_m$, and the above inequalities are tight. When the sources are close to each other, on the other hand, while the original array steering matrix (the steering matrix with $m = 1$) is near singular, the extended steering matrix may be well conditioned. This means that, for closely separated sources, the above inequalities may become loose and the given approximations might be too pessimistic. As a rule of thumb, in the case of a ULA for instance, we say two sources with DOAs α_1 and α_2 are *sufficiently separated* if

$$|\sin\alpha_1 - \sin\alpha_2| \geq \frac{1}{M\Delta},$$

where Δ is antenna spacing measured in fractions of signal wavelength. Note that, when this condition is satisfied, the peaks of the FFT of the columns of the array steering matrix $\boldsymbol{A}$ are at least $\frac{2\pi}{M}$ radians apart, which is equal to the BW of the M-point FFT bins. Thus, this states that, if the peaks of the FFT of $\boldsymbol{A}$ are separated at least by an amount equal to the BW of the bins of the M-point FFT, then the sources are said to be sufficiently separated. Considering the fact that the FFT matrix is a unitary matrix, this is a justifiable assumption. In the following, we shall assume that the sources under consideration are *sufficiently separated* and the above inequalities are tight. Under this condition, we may write

$$\begin{aligned}
\boldsymbol{w}_{\phi_i}^H &\approx \frac{1}{m-1}\left[\begin{array}{cccc} \boldsymbol{w}_i^H(L) & \phi_i^{-1}\boldsymbol{w}_i^H(L) & \cdots & \phi_i^{2-m}\boldsymbol{w}_i^H(L) \end{array}\right] \\
\boldsymbol{w}_{\theta_i}^H &\approx \frac{1}{m}\left[\begin{array}{cccc} \boldsymbol{w}_i'^H(L) & \phi_i^{-1}\boldsymbol{w}_i'^H(L) & \cdots & \phi_i^{1-m}\boldsymbol{w}_i'^H(L) \end{array}\right]
\end{aligned}$$

where $\boldsymbol{w}_i^H(L)$ and $\boldsymbol{w}_i'^H(L)$ are the i-th rows of $\boldsymbol{W}(L)$ and $\boldsymbol{W}'(L)$, and $\boldsymbol{w}_{\phi_i}^H$ and $\boldsymbol{w}_{\theta_i}^H$ are the i-th rows of $\boldsymbol{W}_\phi$ and $\boldsymbol{W}_\theta$, respectively. Putting these into (4.15) and using (4.18), it follows that

$$\begin{aligned}
\rho_{2,L}(\theta_i) &= \frac{1}{2}\|\boldsymbol{w}_i'^H(L)\boldsymbol{\Upsilon}(\theta_i)\|^2 \\
\rho_{2,L}(\phi_i) &= 2\|\boldsymbol{w}_i(L)\|^2
\end{aligned}$$

and

$$\begin{aligned}
\rho_{m,L}(\theta_i) &= \frac{2}{m}\rho_{2,L}(\theta_i) \\
\rho_{m,L}(\phi_i) &\approx \frac{1}{(m-1)^2}\rho_{2,L}(\phi_i).
\end{aligned} \tag{4.31}$$

The expressions for the parameterized angle and frequency estimation errors are obtained by replacing these into (4.20) and (4.26), respectively. Thus, with

$$\left\{\begin{array}{lcl} G_{\theta_i}(m) &=& \frac{2}{m(N-m+1)} \\ G_{\phi_i}(m) &=& \frac{1}{(m-1)^2(N-m+1)} \end{array}\right. \tag{4.32}$$

and

$$\begin{cases} G_{\theta_i}(L) &= \frac{1}{L}\rho_{2,L}(\theta_i) \\ G_{\phi_i}(L) &= \frac{1}{L}\rho_{2,L}(\phi_i) \end{cases}, \tag{4.33}$$

it follows that

$$\begin{aligned} \sigma_{m,L}(\theta) &\approx \frac{\sigma_n^2}{\eta} G_{\theta_i}(m) G_{\theta_i}(L), \\ \sigma_{m,L}(\phi) &\approx \frac{\sigma_n^2}{\eta} G_{\phi_i}(m) G_{\phi_i}(L). \end{aligned} \tag{4.34}$$

The above relations show that the estimation errors are separable functions in m and L. This means that the optimum values of the spatio-temporal smoothing factors (m_o, L_o) are independent of each other. In the following, we shall compute these values, considering each separately.

4.3.1 The Optimum Temporal Smoothing Factor

From (4.34), it is seen that the optimum values of m for the parameterized DOA and frequency estimations are obtained by minimizing $G_\theta(m)$ and $G_\phi(m)$, respectively. [5] Denoting by $m_o(\theta)$ and $m_o(\phi)$ the optimum temporal smoothing factors for the parameterized DOA and frequency estimations, respectively, we have

$$\begin{aligned} m_o(\theta) &= \frac{N+1}{2}, \quad \text{for } \theta, \\ m_o(\phi) &= \frac{2N+3}{3}, \quad \text{for } \phi. \end{aligned}$$

The above value of $m_o(\phi)$ agrees with a similar result reported in [31] for the harmonic retrieval problem. The corresponding variances of the parameterized DOA and frequency estimation are then

$$\begin{aligned} \sigma_{m_o,L}(\theta) &\approx \frac{\sigma_n^2}{\eta} \frac{4}{N^2} G_\theta(L), \\ \sigma_{m_o,L}(\phi) &\approx \frac{\sigma_n^2}{\eta} \frac{27}{4(N-1)^3} G_\phi(L), \end{aligned}$$

For JAFE, it makes more sense to look for an optimum m that minimizes the joint estimation error. To this end, we define a joint estimation error as the geometric mean of the variances of the angle and frequency estimation errors:

$$\sigma_{m,L}(\theta,\phi) \triangleq \sqrt{\sigma_{m,L}^2(\theta)\}\sigma_{m,L}^2(\phi)}, \tag{4.35}$$

[5] To simplify notations, in the remainder of this section, we omit the signal number index i, For instance, we write θ instead of θ_i.

which is equivalent to the arithmetic mean on the logarithmic scale. This definition has twofold advantages: first, it alleviate the scaling problems associated with the arithmetic mean and second, it preserves the separability of the error function in the variables m and L. Minimizing $\sigma_{m,L}(\theta,\phi)$ with respect to m and after some elaboration (approximation), we obtain

$$m_o \approx \frac{3N+2}{5}.$$

It is seen that m_o for the joint estimation is approximately equal to the average of its values obtained considering the angle and frequency estimations separately.

4.3.2 The Optimum Spatial Smoothing Factor

Referring to (4.34), it follows that minimizing the estimation errors with respect to L is equivalent to minimizing $G_\theta(L)$ and $G_\phi(L)$. To this end, first, we need to derive explicit expressions for these functions. Let M_L and M'_L be the number of rows in $\boldsymbol{A}(L)$ and $\boldsymbol{A}'(L)$, respectively, then for sufficiently distinct sources, the following approximations are valid:

$$G_\theta(L) \approx \frac{1}{LM'^2_L} \quad \text{and} \quad G_\phi(L) \approx \frac{1}{LM_L}.$$

For a single source, these relations are exact. For more sources, however, the above approximation is valid only if the steering vectors are sufficiently independent. This is always (asymptotically) satisfied for large M_L and M'_L values.

From their definitions, it is clear that the values of M_L and M'_L depend on the way the spatial smoothing is performed. For a ULA, for instance, if we assume a maximum overlap of subarrays as described in section 3.2.3, $M_L = M-L+1$ and $M'_L = M-L$. Putting these into the above expressions and minimizing $G_\theta(L)$ and $G_\phi(L)$ with respect to L, we obtain

$$L_o = \begin{cases} \frac{M}{3} & \text{for DOA, and} \\ \frac{M+1}{2} & \text{for frequency} \end{cases}$$

The corresponding variances of the parameterized DOA and frequency estimation are then

$$\begin{aligned} \sigma_{m,L_o}(\phi) &\approx \frac{\sigma_n^2}{\eta}\frac{27}{4M^3}G_\theta(m), \\ \sigma_{m,L_o}(\theta) &\approx \frac{\sigma_n^2}{\eta}\frac{4}{M^2-1}G_\phi(m). \end{aligned}$$

For JAFE, defining a joint estimation error as in (4.35), we obtain

$$L_o \approx \frac{2M}{5}.$$

As in the case of m_o, this value of L_o is approximately equal to the average of the optimum values obtained considering the angle and frequency estimations separately.

4.3.3 Forward-Backward Averaging

Forward-backward averaging [27, 90] is equivalent to doubling the number of temporal samples, with the rest of the data parameters remaining unchanged. Thus, for both θ and ϕ, this data extension provides a factor 2 improvement in the estimation accuracies. The important aspect of forward-backward averaging is that the resulting data matrix can be transformed into a real matrix of the same size (viz. 3.8, [43]). This provides a substantial reduction in complexity (viz. chapter 7, [27]), while improving the accuracy.

4.4 Errors in Computing the actual parameters

The final step in the JAFE algorithm is to transform the parameterized DOAs and the parameterized frequencies into their actual values in radians and Hz, respectively. In this section, I consider how these transformations affect the error behaviors.

4.4.1 Errors in computing the DOAs

The relation between the parameterized and the physical DOAs is defined by the array geometry. In the case of the ULA, for instance, this relation is given by

$$\alpha = \operatorname{asin}\left(\frac{\arg\theta}{2\pi\Delta}\right).$$

Assuming, the real and imaginary parts of θ are affected by statistically independent equal variance noises, the mean square errors on the angle estimates can be computed as (first-order approximations) (viz. [58, 65, 66, 91, 93])

$$\sigma_{m,L}(\alpha) := E\{\Delta\alpha^2\}_{m,L} = \frac{\sigma_{m,L}(\theta)}{2(2\pi\Delta\cos\alpha)^2}. \tag{4.36}$$

The final expressions for $\sigma_{m,L}(\alpha)$ is obtained by replacing $\sigma_{m,L}(\theta)$ with the expressions from (4.20).

4.4.2 Errors in computing the frequencies

As recalled from section 3.1, the actual signal frequency is computed from the parameterized frequency ϕ using the transformation

$$f = \frac{P}{2\pi} \arg \phi.$$

The first order approximation of the perturbation Δf on f is:

$$\Delta f \approx \mathrm{Re}\left(\frac{\partial f}{\partial \phi}\Delta\phi\right) = \frac{P}{2\pi}\mathrm{Re}(\Delta\phi),$$

and thus, the variance of the frequency estimation error is

$$\sigma_{m,L}(f) = \frac{1}{2}\frac{P^2}{4\pi^2}\sigma_{m,L}(\phi). \tag{4.37}$$

The final expression for $\sigma_{m,L}(f)$ is obtained by replacing $\sigma_{m,L}(\phi)$ with the expression from (4.26).

4.5 The Cramer-Rao Lower Bound

Putting a lower bound for any estimator proves to be extremely useful. It provides a benchmark against which we can compare the performance of any unbiased estimator. Moreover, it tells us the impossibility of finding an unbiased estimator whose variance is less than the bound. One such bound is the *Cramer-Rao Lower bond* (CRB) [39]. In this section, we derive the CRB for the JAFE algorithm.

Let us assume that an $M \times 1$ deterministic signal vector $\boldsymbol{s}(k;\boldsymbol{\mu})$ with unknown parameter vector

$$\boldsymbol{\mu} = \left[\begin{array}{cccc} \mu_1 & \mu_2 & \cdots & \mu_q \end{array}\right],$$

is observed in additive noise,

$$\boldsymbol{x}(k) = \boldsymbol{s}(k;\boldsymbol{\mu}) + \boldsymbol{\eta}(k) = \left[\begin{array}{c} s_1(k;\boldsymbol{\mu}) \\ s_2(k;\boldsymbol{\mu}) \\ \cdots \\ s_M(k;\boldsymbol{\mu}) \end{array}\right] + \boldsymbol{\eta}(k) \in \mathbb{C}^M \tag{4.38}$$

where $\boldsymbol{\eta}(k)$ is an $M \times 1$ noise vector. Here, the dependence of the signal $\boldsymbol{s}(k)$ on $\boldsymbol{\mu}$ is made explicit by writing $\boldsymbol{s}(k;\boldsymbol{\mu})$. Assume that $\boldsymbol{\eta}(k)$ is a white Gaussian noise (WGN) with variance σ^2, and that we have collected N time samples of

the signal $\boldsymbol{x}(k)$. Then, the log likelihood function of the signal (defined as the logarithm of the probability density function) is given by

$$\mathcal{L}(\boldsymbol{x};\boldsymbol{\mu}) = -\frac{MN}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{k=1}^{N}\left[\boldsymbol{x}(k) - \boldsymbol{s}(k;\boldsymbol{\mu})\right]^H\left[\boldsymbol{x}(k) - \boldsymbol{s}(k;\boldsymbol{\mu})\right]. \tag{4.39}$$

Let the gradient of the signal vector $\boldsymbol{s}(k;\boldsymbol{\mu})$ with respect to $\boldsymbol{\mu}$ be denoted by $\boldsymbol{D}_k(\boldsymbol{\mu})$, i.e.,

$$\boldsymbol{D}_k(\boldsymbol{\mu}) = \left[\begin{array}{cccc} \frac{\partial \boldsymbol{s}(k;\boldsymbol{\mu})}{\partial \mu_1} & \frac{\partial \boldsymbol{s}(k;\boldsymbol{\mu})}{\partial \mu_2} & \cdots & \frac{\partial \boldsymbol{s}(k;\boldsymbol{\mu})}{\partial \mu_q} \end{array}\right],$$

then the so-called Fisher information matrix is given by

$$\boldsymbol{I}(\boldsymbol{\mu}) = \frac{1}{\sigma^2}\mathrm{Re}\left(\sum_{k=1}^{N}\boldsymbol{D}_k^H(\boldsymbol{\mu})\boldsymbol{D}_k(\boldsymbol{\mu})\right). \tag{4.40}$$

The CRB for estimating the i-th parameter μ_i is obtained from the inverse of the Fisher information matrix (viz. [39]) as

$$\mathrm{CRB}(\mu_i) = [\boldsymbol{I}^{-1}(\boldsymbol{\mu})]_{ii}$$

4.5.1 CRB for the JAFE Data Model

Consider a simplified version of the JAFE data model, in which the modulating signals are set to have bandwidths $BW = 0$,

$$\boldsymbol{x}(k) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{B}\boldsymbol{\phi}^k + \boldsymbol{\eta}(k) =: \boldsymbol{s}(k;\boldsymbol{\mu}) + \boldsymbol{\eta}(k), \tag{4.41}$$

where

- $\boldsymbol{A}(\boldsymbol{\theta}) = \left[\begin{array}{ccc} \boldsymbol{a}(\theta_1) & \boldsymbol{a}(\theta_2) & \boldsymbol{a}(\theta_d) \end{array}\right] \in \mathbb{C}^{M,d}$, and $\boldsymbol{a}(\theta)$ is the array response vector for a signal with a parameterized DOA of θ.
- $\boldsymbol{\theta} = [\theta_1 \ \cdots \ \theta_d]^T$ is a vector containing the parameterized DOA of the d signals.
- $\boldsymbol{\phi}^k = \left[\begin{array}{cccc} \phi_1^k & \phi_2^k & \cdots & \phi_d^k \end{array}\right]^T$, where ϕ_i is the parameterized frequency of the i-th signal.
- $\boldsymbol{B} = \mathrm{diag}\{\beta_i\}_{i=1}^d$ is a diagonal gain matrix, where $\beta_i \in \boldsymbol{R}^+$ is the amplitude of the i-th signal as received by the antenna array.
- $\boldsymbol{\eta}(k)$ is an $M \times 1$ white Gaussian noise vector.

Let $\boldsymbol{\beta} = [\beta_1 \ \cdots \ \beta_d]^T$ be a vector continning the channel gains, then the conditioning parameters that affect the signal likelihood function are collected into the $3d \times 1$ vector $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \left[\begin{array}{ccc} \boldsymbol{\theta}^T & \boldsymbol{\phi}^T & \boldsymbol{\beta}^T \end{array} \right].$$

Define

$$\begin{aligned} \boldsymbol{D}_\theta &= \left[\begin{array}{ccc} \frac{\partial \boldsymbol{a}}{\partial \theta_1}(\theta_1) & \cdots & \frac{\partial \boldsymbol{a}}{\partial \theta_d}(\theta_d) \end{array} \right] \\ \boldsymbol{\Phi}^k &= \text{diag}\{\boldsymbol{\phi}^k\}. \end{aligned}$$

Evaluating the derivative of $\boldsymbol{s}(k;\boldsymbol{\mu})$ with respect to each parameter we get the following

$$\begin{aligned} \frac{\partial \boldsymbol{s}(k;\boldsymbol{\mu})}{\partial \boldsymbol{\theta}} &= \boldsymbol{D}_\theta \boldsymbol{B}\boldsymbol{\Phi}^k =: \boldsymbol{D}_k(\boldsymbol{\theta}), \\ \frac{\partial \boldsymbol{s}(k;\boldsymbol{\mu})}{\partial \boldsymbol{\phi}} &= k\boldsymbol{A}\boldsymbol{B}\boldsymbol{\Phi}^{k-1} =: \boldsymbol{D}_k(\boldsymbol{\phi}), \\ \frac{\partial \boldsymbol{s}(k;\boldsymbol{\mu})}{\partial \boldsymbol{\beta}} &= \boldsymbol{A}\boldsymbol{\Phi}^k =: \boldsymbol{D}_k(\boldsymbol{\beta}). \end{aligned}$$

Let $\boldsymbol{I}_k(\boldsymbol{\mu})$ be defined as

$$\boldsymbol{I}_k(\boldsymbol{\mu}) = \frac{1}{\sigma^2}\text{Re}\left[\begin{array}{c} \boldsymbol{D}_k(\boldsymbol{\theta})^H \\ \boldsymbol{D}_k(\boldsymbol{\phi})^H \\ \boldsymbol{D}_k(\boldsymbol{\beta})^H \end{array} \right] \left[\begin{array}{c} \boldsymbol{D}_k(\boldsymbol{\theta})^H \\ \boldsymbol{D}_k(\boldsymbol{\phi})^H \\ \boldsymbol{D}_k(\boldsymbol{\beta})^H \end{array} \right]^H. \tag{4.42}$$

Then, the Fisher information matrix is

$$\boldsymbol{I}(\boldsymbol{\mu}) = \sum_{k=1}^{N} \boldsymbol{I}_k(\boldsymbol{\mu}) = \frac{1}{\sigma^2}\text{Re}\left[\begin{array}{ccc} \boldsymbol{\Delta} & \boldsymbol{P}^H & \boldsymbol{Q}^H \\ \boldsymbol{P} & \boldsymbol{\Gamma} & \boldsymbol{R}^H \\ \boldsymbol{Q} & \boldsymbol{R} & \boldsymbol{\Lambda} \end{array} \right], \tag{4.43}$$

where N is the number of time samples and

$$\begin{aligned} \boldsymbol{\Delta} &= \sum_{k=1}^{N} \boldsymbol{\Phi}^{-k}\boldsymbol{B}\boldsymbol{D}_\alpha^H \boldsymbol{D}_\alpha \boldsymbol{B}\boldsymbol{\Phi}^k, & \boldsymbol{P} &= \sum_{k=1}^{N} k\boldsymbol{\Phi}^{-k}\boldsymbol{B}\boldsymbol{D}_\theta^H \boldsymbol{A}\boldsymbol{B}\boldsymbol{\Phi}^{k-1}, \\ \boldsymbol{\Sigma} &= \sum_{k=1}^{N} k^2\boldsymbol{\Phi}^{1-k}\boldsymbol{B}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{B}\boldsymbol{\Phi}^{k-1}, & \boldsymbol{Q} &= \sum_{k=1}^{N} \boldsymbol{\Phi}^{-k}\boldsymbol{B}\boldsymbol{D}_\theta^H \boldsymbol{A}\boldsymbol{\Phi}^k, \\ \boldsymbol{\Lambda} &= \sum_{k=1}^{N} \boldsymbol{\Phi}^{-k}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{\Phi}^k, & \boldsymbol{R} &= \sum_{k=1}^{N} k\boldsymbol{\Phi}^{1-k}\boldsymbol{B}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{\Phi}^k. \end{aligned}$$

In the simulation results, the CRB is used as a bench mark to evaluate the performance of the JAFE algorithm. It is seen from the results that the JAFE algorithm approaches the CRB only when the spatio-temporal smoothing factors (viz. section 3.2.2) are chosen to be close to their optimum values. This could be computationally too expensive. In section 5.4, it will be shown that, with a multi-resolution spatio-temporal sampling, the CRB is attained at a much lower complexity.

4.6 Simulation Examples

In this simulation example, we consider a 4 element ULA with baseline separation of $\Delta = 1/2$. We assume that two far field, equal power signals s_1 and s_2 are impinging on the antenna array. The DOA and center frequency of s_1 are $\alpha_1 = 10$ Degrees and $f_1 = 2$ MHz, and those of s_2 are $\alpha_2 = 55$ Degrees and $f_2 = 5$ MHz respectively. The source signals are narrow-band (25 kHz) amplitude modulated signals. The data is sampled at a rate of 20 MHz, and the processing is done over $N = 32$ time samples. All simulation results are based on 100 Monte-Carlo runs. The behaviors are summarized in Fig. 4.1 through Fig. 4.4.

Fig. 4.1 shows how temporal smoothing affects the parameter estimation errors. From the plots, it is seen that the DOA estimation error is minimum for $m = N/2$ and the frequency estimation error is minimum for $m = 2N/3$ as predicted in section 4.3. It is seen that, when $m \to 2$, the frequency/DOA estimation errors increase sharply and that, for the algorithm to attain the CRB, m should be close to its optimum value. Choosing large m, however, increases the computational complexity. This means that one has to find a compromise between complexity and accuracy.

In Fig. 4.2, it is shown that, apart from improving the estimation accuracy, temporal smoothing also provides robustness against rank loss when there exist multiple signals with the same DOA. This is in agreement with the identifiability conditions discussed in section 3.3.

The effect of spatial smoothing on the estimation errors is summarized in Fig. 4.3. The simulation was run using a ULA with $M = 16$ elements, $N = 16$, $m = 10$ and SNR $= 20$dB. The DOAs and the center frequencies of the two sources under consideration are the same as before. As predicted in section 4.3, the parameterized DOA and frequency estimation errors are minimum for $L = \frac{M}{3}$ and $L = \frac{M+1}{2}$, respectively. Moreover, in Fig. 4.4 it is seen that, apart from performance improvement, spatial smoothing achieves rank restoration when several signals have the same center frequencies. In

Fig. 4.4a, a seemingly unexpected behavior is seen. That is, for large frequency separation, the DOA estimation error increased when L was changed from 1 to 2. This should not be surprising because for $M = 4$ and $L = 2$, the ratio $(M-1)^2 > L(M-L)^2$ (cf. section 4.3.2) and thus $E_2 > E_1$, where E_1 and E_2 are estimation errors corresponding to $L = 1$ and $L = 2$, respectively.

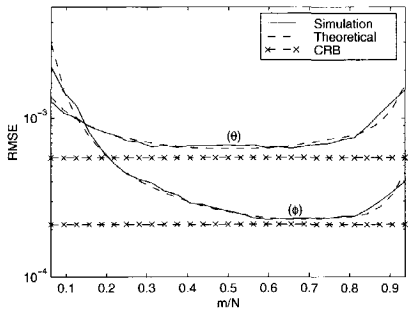


Fig. 4.1: The parameterized DOA and frequency estimation errors as functions of temporal smoothing factor m at SNR = 30dB. It is seen that the theoretical behaviors perfectly agree with the simulation results.

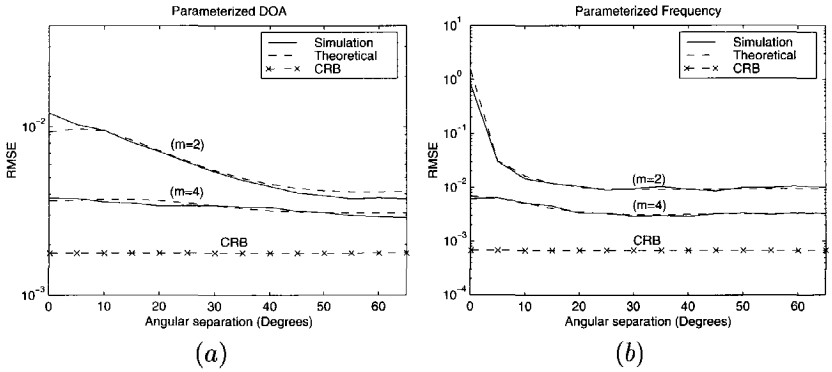


Fig. 4.2: Behavior of (a) the parameterized DOA and (b) the parameterized frequency estimation errors as functions of angular separation. Note the improvement obtained via temporal smoothing, particularly at small angular separations. (SNR = 20dB).

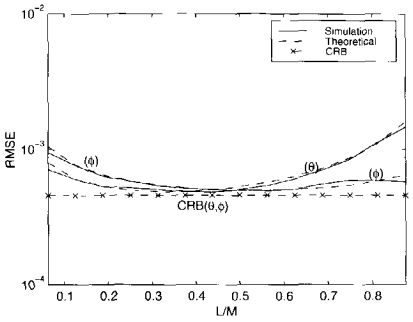


Fig. 4.3: The parameterized DOA and frequency estimation errors as functions of spatial smoothing factor L at SNR = 20dB. It is seen that the theoretical behaviors perfectly agree with the simulation results.

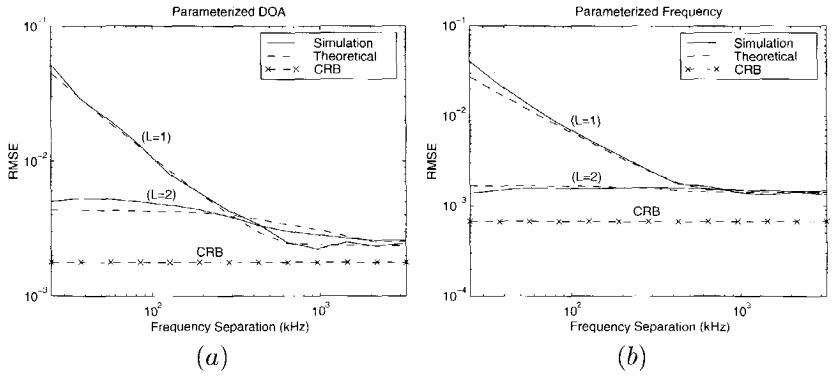


Fig. 4.4: Behavior of (a) the parameterized DOA and (b) the parameterized frequency estimation errors as functions of frequency separation. Note the superior performance of the spatially smoothed data approach at small frequency separations. (SNR = 20dB.)

4.7 Discussion

In this chapter, I have presented an analysis of the ESPRIT based JAFE Algorithm. Using a simple perturbation model, I was able to derive analytical expressions for the estimation errors, and for the optimum values of the spatial and temporal smoothing factors. The optimum spatio-temporal smoothing

factors so obtained are seen to be different from those derived in chapter 3 for maximizing the number of identifiable sources. Thus, it means that one cannot satisfy simultaneous optimality in both identifiability and accuracy. Moreover, I have shown that the JAFE algorithm achieves the CRB only when the spatio-temporal smoothing factors are close to their optimum values. However, choosing the smoothing factors close to their optimum values may be computationally too expensive. In this case, a technique called the multi-resolution JAFE that alleviates this problem without causing too much degradation in accuracy, can be implemented. This technique is discussed in the following chapter (chapter 5).

Finally, it should be noted that the presented performance analysis is independent of the joint diagonalization technique that may have been used in solving the joint eigenvalue decomposition (EVD) problems (viz. section 3.2.2). In the analysis I have tacitly assumed that an optimum joint diagonalization of the matrix pencil problem has been achieved. The actual performance, therefore, depends on the quality of the joint diagonalization procedure employed. Though there are several accounts in the literature on this problem, it is an open research topic to find a reliable optimal solution. One way to evaluate the quality of a joint diagonalization method is to compare it against the expected performance derived in this chapter.

Chapter 5

A MULTI-RESOLUTION ESPRIT ALGORITHM

Contents

As described in chapter 2, the ESPRIT algorithm makes use of a single shift invariance structure present in the array response vector $\boldsymbol{a}(\theta)$ to estimate the DOA of the incoming signals. Here $\theta = e^{j\mu}$, and μ is a phase shift to be estimated. In narrowband direction-of-arrival estimation, the phase shift is due to the difference in arrival times of the wavefront at the elements of an antenna

array. For a uniform linear array (ULA), $\boldsymbol{a}(\theta) = [1 \quad \theta \quad \theta^2 \quad \cdots]^T$ and $\mu = 2\pi\Delta\sin(\alpha)$, where Δ is the distance between the elements (in wavelengths), and α is the angle of arrival measured with respect to the normal of the array axis as shown in Fig. 2.4.

It has been shown in section 4.4.1 that the accuracy of the estimation of $\sin(\alpha)$ is directly proportional to $\frac{1}{\Delta}$ (crf. [58, 65, 66, 91, 93]). Thus, it is preferable to have a large baseline separation Δ, so that we collect a large phase shift μ. Unfortunately, however, we cannot collect more than a single cycle, $-\pi \leq \mu < \pi$, because the inverse of the mapping $\mu \rightarrow \theta = e^{j\mu}$ is ambiguous outside this range. To prevent aliasing, we thus have to ensure that $\Delta \leq \frac{1}{2}$, which is essentially Shannon's sampling theorem in the spatial domain.

The idea behind multi-resolution parameter estimation is to obtain *two* or more estimates of μ: the first based on a small baseline or short sampling period, yielding a coarse estimate μ_1 of μ without aliasing, and the second based on a large baseline or (much) larger sampling period, providing an aliased estimate μ_2 of μ at a finer scale. These two estimates are combined to obtain a final estimate $\hat{\mu} = 2\pi n + \mu_2$, where the integer number of cycles n is estimated from μ_1.

Similar works have been reported in the literature. In particular, Zoltowski et.al. [104] discuss a similar problem of angle-frequency estimation using multiple scales in time and space. However, their solutions are very much directed by engineering considerations, which incurs a certain sacrifice in elegance and clarity. In particular, the coarse frequency estimation is done by applying ESPRIT to a small set of DFT values around spectral peaks which are determined via peak searching algorithms. The fine frequency estimates and the angle estimates are obtained sequentially and for each estimated coarse frequency independently, which assumes that they are sufficiently unique. Here, we derive a one-shot joint estimation procedure referred to as the MR-ESPRIT [47, 48, 50].

There is a connection of MR-ESPRIT to MI-ESPRIT [84, 97] as well. MI-ESPRIT, like the MR-ESPRIT, exploits the multiple shift-invariance structure present in multi-baseline arrays. A distinction is that MI-ESPRIT is formulated in terms of (iterative) subspace fitting, and basically attempts to find more accurate beamforming vectors by considering multiple shift invariances. The original paper [84] did not specifically recognize the fact that also more accurate direction estimates can be found. In [97], a non-iterative MI-ESPRIT is given. There, the aliasing is resolved by searching for an optimum solution among formerly computed candidates. The approach presented here, on the other hand, resolves aliasing by merely solving a set of analytic equations. Moreover, the corresponding parameters are grouped automatically without the need for any extra processing, which is not the case in MI-ESPRIT.

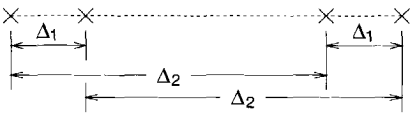


Fig. 5.1: Multi-resolution spatial sampling

In this chapter, after introducing the MR-ESPRIT, I apply it to DOA estimation, frequency estimation, and a combination of the two. To make things simple, I will concentrate on a uniform linear array (ULA). The concept can easily be extended to multi-dimensional arrays and for multi-dimensional ESPRIT.

5.1 The MR-ESPRIT

The original ESPRIT algorithm is based on arrays with a doublet structure, i.e., consisting of several antenna pairs with the same baseline vectors. However, the chosen array geometries often admit other pairings with different baselines. For instance, the array structure shown in Fig. 5.1 combines two spatial sampling rates. The minimal number of antennas to have two baseline vector pairs is four. With more antennas, several interesting configurations are possible.

As discussed in chapter 2, the M-dimensional array response vector $\boldsymbol{a}(\alpha)$ represents the response of the M-element antenna array to a narrowband signal from a direction α. It can be parameterized in several ways. The usual parameterization is in terms of $\theta = e^{j2\pi\Delta\sin(\alpha)}$, where Δ is a reference distance measured in fractions of signal wavelength. In our case of an array with two baselines, we can (redundantly) parameterize the array by two parameters, $\theta_1 = e^{j2\pi\Delta_1\sin(\alpha)}$ and $\theta_2 = e^{j2\pi\Delta_2\sin(\alpha)}$. In the case of the array of Fig. 5.1, we have

$$\boldsymbol{a}(\theta_1,\theta_2) = \begin{bmatrix} 1 \\ \theta_1 \\ \theta_2 \\ \theta_1\theta_2 \end{bmatrix}. \tag{5.1}$$

The idea is to treat the two parameters θ_1 and θ_2 as independent and estimate both of them from the measurement data, and only then combine them into a single estimate of $\sin(\alpha)$. Estimation is done by exploiting the *dual shift-*

invariance structure of $\boldsymbol{a}(\theta_1, \theta_2)$, i.e., in the above example

$$\boldsymbol{a}_{x1} = \begin{bmatrix} 1 \\ \theta_2 \end{bmatrix}, \boldsymbol{a}_{y1} = \begin{bmatrix} \theta_1 \\ \theta_1\theta_2 \end{bmatrix} \quad \Rightarrow \quad \boldsymbol{a}_{y1} = \boldsymbol{a}_{x1}\theta_1,$$

$$\boldsymbol{a}_{x2} = \begin{bmatrix} 1 \\ \theta_1 \end{bmatrix}, \boldsymbol{a}_{y2} = \begin{bmatrix} \theta_2 \\ \theta_1\theta_2 \end{bmatrix} \quad \Rightarrow \quad \boldsymbol{a}_{y2} = \boldsymbol{a}_{x2}\theta_2,$$

For more general arrays with a dual shift-invariance structure, we can define selection matrices $\boldsymbol{J}_{xi}$ and $\boldsymbol{J}_{yi}$ $(i = 1, 2)$ such that the above relations hold for $\boldsymbol{J}_{xi}\boldsymbol{a}$ and $\boldsymbol{J}_{yi}\boldsymbol{a}$.

Let μ_i $(i = 1, 2)$ be the argument of θ_i. Then, if the distance $\Delta_i < \frac{1}{2}$, the angle of arrival α of the wavefront can be uniquely determined from μ_i using the transformation

$$\alpha = \arcsin(\frac{\mu_i}{2\pi\Delta_i}).$$

However, when $\Delta_i > \frac{1}{2}$, because of aliasing we get a set of cyclically related candidates for α:

$$\alpha(n) = \arcsin(\frac{\mu_i + 2\pi n}{2\pi\Delta_i}).$$

In MR-ESPRIT we combine the non-aliased and the aliased estimates of the parameters to obtain a better estimation accuracy. The resulting algorithm is very similar to the case of joint azimuth-elevation estimation [86].

Thus, to be specific, consider d narrowband sources $s_i(t)$ impinging on the antenna array. Collecting N samples of the M antenna outputs at a rate P into an $M \times N$ data matrix $\boldsymbol{X}$ as in section 2.2.1, we obtain the data model

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{N} = \boldsymbol{a}_1 s_1 + \cdots + \boldsymbol{a}_d s_d + \boldsymbol{N}$$

where the columns of $\boldsymbol{A} \in \mathbb{C}^{M,d}$ are the array response vectors $\{\boldsymbol{a}_i\}$, the rows of $\boldsymbol{S}\mathbb{C}^{d,N}$ are the sampled source signals and $\boldsymbol{N} \in \mathbb{C}^{M,N}$ is the noise matrix. Assuming $d < M$, the first step of the algorithm is to estimate a basis $\boldsymbol{U}_s$ of the column span of $\boldsymbol{X}$, typically using an SVD. $\boldsymbol{U}_s$ and $\boldsymbol{A}$ are related by a $d \times d$ nonsingular matrix $\boldsymbol{T}$ as (viz. section 2.3)

$$\boldsymbol{U}_s = \boldsymbol{A}\boldsymbol{T}$$

The second step in the algorithm is to form submatrices of $\boldsymbol{U}_s$ using the proper selection matrices,

$$\boldsymbol{U}_{xi} = \boldsymbol{J}_{xi}\boldsymbol{U}_s, \qquad \boldsymbol{U}_{yi} = \boldsymbol{J}_{yi}\boldsymbol{U}_s. \quad (i = 1, 2)$$

The shift-invariance structure of the array implies that

$$\boldsymbol{U}_{xi} = \boldsymbol{A}_i\boldsymbol{T}, \qquad \boldsymbol{U}_{yi} = \boldsymbol{A}_i\boldsymbol{\Theta}_i\boldsymbol{T},$$

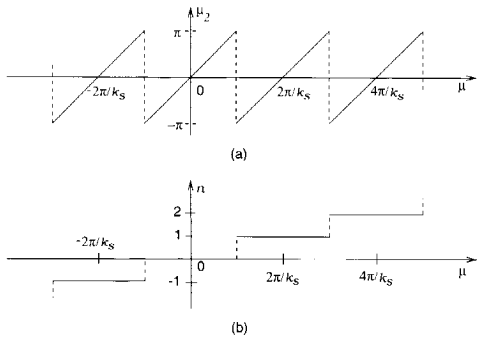


Fig. 5.2: (a) The aliased spatial frequency μ_2 as a function of the alias-free spatial frequency μ. (b) The corresponding winding number n.

where $\boldsymbol{A}_i$ is a submatrix of $\boldsymbol{A}$ and the diagonal matrix $\boldsymbol{\Theta}_i = \text{diag}\{\theta_{ij}\}_{j=1}^{d}$ contains the d shift parameters of the d sources with reference to the i-th baseline. The final step is to estimate the parameters by considering

$$\begin{aligned} \boldsymbol{E}_1 &= \boldsymbol{U}_{x1}^{\dagger}\boldsymbol{U}_{y1} &= \boldsymbol{T}^{-1}\boldsymbol{\Theta}_1\boldsymbol{T}, \\ \boldsymbol{E}_2 &= \boldsymbol{U}_{x2}^{\dagger}\boldsymbol{U}_{y2} &= \boldsymbol{T}^{-1}\boldsymbol{\Theta}_2\boldsymbol{T}. \end{aligned}$$

It is seen that the data matrices $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are jointly diagonalizable by the same matrix $\boldsymbol{T}$, and can be solved using any of the joint diagonalization techniques discussed in section 3.6. After $\boldsymbol{T}$ has been found, we also have estimates of $\{(\theta_{1k}, \theta_{2k})\}$ for each of the d sources.

It remains, for each source, to combine[1] θ_1 and θ_2 into an estimate of the argument μ of θ. Let us assume that $\Delta_1 \leq \frac{1}{2}$, so that μ_1 (argument of θ_1) is not aliased and is a coarse estimate of μ. Also assume that $\Delta_2 \gg \frac{1}{2}$, so that in μ_2 aliasing occurs: the estimate μ is proportional to μ_2 plus an appropriate integer multiple of 2π (see Fig. 5.2). It follows that we have two estimates of $2\pi \sin(\alpha)$,

$$2\pi \sin(\alpha) = \frac{1}{\Delta_1}\mu_1 = \frac{1}{\Delta_2}(2\pi n + \mu_2)\,. \tag{5.2}$$

The winding number n is determined as the best fitting integer to match the two right hand side expressions,

$$n = \text{round}\left(\frac{1}{2\pi}\left(\frac{\Delta_2}{\Delta_1}\mu_1 - \mu_2\right)\right) =: \text{round}(\hat{n}). \tag{5.3}$$

[1]Here we drop the subscript k in θ_{ik} for readability purpose.

The ratio $k_s := \frac{\Delta_2}{\Delta_1}$ can be interpreted as the (spatial) gain in accuracy. In particular, the estimate of $2\pi \sin(\alpha)$ based on μ_2 is a factor k_s more accurate than that based on μ_1. Thus a more accurate estimate of the spatial frequency μ can be obtained as

$$\mu = \frac{1}{k_s}(2\pi n + \mu_2). \tag{5.4}$$

or by optimally combining the two estimates as

$$\mu = \frac{1}{k_s^2}\mu_1 + (1 - \frac{1}{k_s^2})\frac{1}{k_s}(2\pi n + \mu_2) \tag{5.5}$$

5.2 Analysis

As described above, the steps involved in MR-ESPRIT are similar to that of the JAFE algorithm. Hence, the performance analysis follows essentially the same procedure. However, in the MR-ESPRIT case, there are more parameters involved in the estimation process, such as the winding number and the gain in resolution. In this section, we take a look how these parameters are affected by the presence of noise (measurement as well as system noise).

5.2.1 The winding number

Consider the relations given in (5.3) and (5.4), where we have tacitly assumed that $\Delta_2 = k_s\Delta_1$ holds perfectly. In practice however, due to measurement errors, this holds only approximately. Let Δk_s represent the error on k_s such that $\Delta_2 = (k_s + \Delta k_s)\Delta_1$. Also assume that μ_1 and μ_2 are determined with estimation errors $\Delta\mu_1$ and $\Delta\mu_2$, respectively, where $\Delta\mu_1$ and $\Delta\mu_2$ are independent random variables, with $\mathrm{E}\{\Delta\mu_1^2\} = \mathrm{E}\{\Delta\mu_2^2\} = \sigma_\mu^2$. [2] With these assumptions, the error Δn on $\hat{n}$ in (5.3) can be approximated as

$$\Delta n \approx \frac{\partial \hat{n}}{\partial k_s}\Delta k_s + \frac{\partial \hat{n}}{\partial \mu_1}\Delta\mu_1 + \frac{\partial \hat{n}}{\partial \mu_2}\Delta\mu_2.$$

Replacing the value of $\hat{n}$ from (5.3) into the above equation we obtain

$$\Delta n = \frac{1}{2\pi}\mu_1\Delta k_s + \frac{1}{2\pi}(k_s\Delta\mu_1 - \Delta\mu_2). \tag{5.6}$$

For a given array configuration, the first term in (5.6) is a constant. It represents the offset on $\hat{n}$ due to the array imperfection. On the other hand, both parameters $\Delta\mu_1$ and $\Delta\mu_2$ in the second term are zero mean Gaussian processes (viz

[2] This is a valid assumption because phase estimation error is not a function of the antenna separation distance.

chapter 4 [58, 91, 93]). [3] Consequently, Δn is also a Gaussian process with a mean $\frac{1}{2\pi}\mu_1 \Delta k_s$ and a variance

$$\sigma_n^2 = \mathrm{E}\{(\Delta n - \frac{1}{2\pi}\mu_1 \Delta k_s)^2\} = \frac{1}{4\pi^2}(k_s^2 + 1)\sigma_\mu^2. \tag{5.7}$$

A typical distribution function of Δn is shown in Fig. 5.3. It is seen from (5.3)

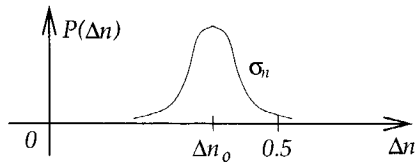


Fig. 5.3: A typical probability distribution function of Δn, $(\Delta n_o = \frac{1}{2\pi}\mu_1 \Delta k_s)$

that n is determined correctly if $|\Delta n| < 0.5$. However, since Δn is a random process, we can satisfy this only with some uncertainty (confidence level). In particular, given a required confidence level $\mathcal{L}$, we find the conditions under which the probability

$$P(|\Delta n| < 0.5) > \mathcal{L} \tag{5.8}$$

Assuming that $P(\cdot)$ has a Gaussian distribution, it can be shown that, [29]

$$P(|\Delta n| < 0.5) = \frac{1}{2}\mathrm{erf}\left(\frac{\pi + \mu_1 \Delta k_s}{\sigma_\mu \sqrt{2(k_s^2+1)}}\right) + \frac{1}{2}\mathrm{erf}\left(\frac{\pi - \mu_1 \Delta k_s}{\sigma_\mu \sqrt{2(k_s^2+1)}}\right), \tag{5.9}$$

where σ_μ represents the root mean square measurement error in μ. A family of curves $P(|\Delta n| < 0.5)$, for $\mu_1 = \pi$ (representing the worst case scenario) and an arbitrarily chosen value of $\Delta k_s = 0.75$, as functions of σ_μ (for different values of k_s) are shown in Fig. 5.4. To obtain more explicit expressions, let the function $f_\kappa(x)$ be defined as

$$f_\kappa(x) = \frac{1}{2}\mathrm{erf}\left((\pi + \mu_1 \kappa)x\right) + \frac{1}{2}\mathrm{erf}\left((\pi - \mu_1 \kappa)x\right), \tag{5.10}$$

then $P(|\Delta n| < 0.5)$ may be expressed in terms of $f_\kappa(x)$ as

$$P(|\Delta n| < 0.5) = f_{\Delta k_s}\left(\frac{1}{\sigma_\mu \sqrt{2(k_s^2+1)}}\right).$$

[3] More precisely, these are Gaussian processes if the input noise is Gaussian.

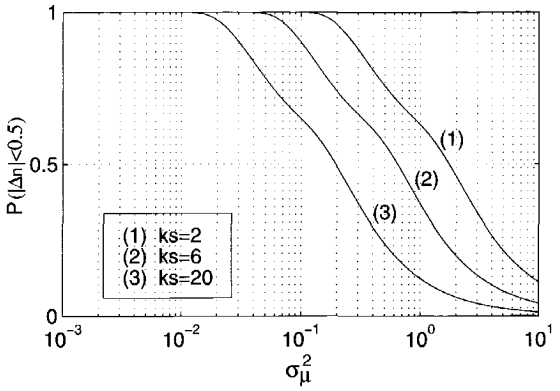


Fig. 5.4: A family of curves $P(|\Delta n| < 0.5)$ as functions of σ_μ for $\mu_1 = \pi$ and $\Delta k_s = 0.75$

Now, putting this into (5.8) and solving for k_s, we get

$$k_s < \sqrt{\frac{1}{2\sigma_\mu^2}\left(\frac{1}{f_{\Delta k_s}^{-1}(\mathcal{L})}\right)^2 - 1} \quad =: k_{\max}, \tag{5.11}$$

where $f_{\Delta k_s}^{-1}(\cdot)$ is the inverse function of $f_{\Delta k_s}(\cdot)$. From this relation, it is clear that the resolution gain factor cannot be made arbitrarily large. It is bounded from above by a number which is a function of the estimation error and the array imperfection factor Δk_s. Particularly, one can clearly see that, as the estimation error increases, the maximum value of k_s decreases. This is in perfect agreement with intuitive perception. For instance, for the case $\mu_1 = \pi$, $\Delta k_s = 0$ and $\mathcal{L} = 0.998$, the bounds on k_s at $\sigma_\mu = 0.1$ and 0.05 are 10.1 and 20.3, respectively. Note that, when $\Delta k_s = 0$, (5.9) reduces to

$$P(|\Delta n| < 0.5) = \text{erf}\left(\frac{\pi}{\sigma_\mu\sqrt{2(k_s^2+1)}}\right),$$

and the expression for $k_{\max}$ becomes

$$k_{\max} = \sqrt{\frac{\pi^2}{2\sigma_\mu^2}\left(\frac{1}{\text{erf}^{-1}(\mathcal{L})}\right)^2 - 1} \tag{5.12}$$

5.2.2 Dependence of $k_{\max}$ on SNR

To establish the relation between $k_{\max}$ and SNR, we first need to determine the dependence of σ_μ (the phase estimation error) on the SNR. To this end, let us assume that we have collected N time samples of the output of an M element antenna array that has a dual shift invariance structure. Moreover, for $M_i < M$ $(i = 1, 2)$, let $\boldsymbol{J}_{xi} \in \mathbb{R}^{M_i, M}$ and $\boldsymbol{J}_{yi} \in \mathbb{R}^{M_i, M}$, $i = 1, 2$ be the selection matrices used to construct the shift invariance matrix pairs with reference to the i-th base line. M_i is the number of elements in the selected subarrays, and for convenience, we assume that $M_1 = M_2 = M_o$. Using the results of chapter 4, we can then modify (4.25) to obtain

$$\sigma^2_{\alpha,i} = \frac{1}{SNR}\left(\frac{1}{M_o^2 N}\left(\frac{1}{2\pi\Delta_i \cos(\alpha)}\right)^2\right), \tag{5.13}$$

where $\sigma_{\alpha,i}$ is the root mean square error (RMSE) obtained with reference to the i-th base line separation Δ_i. Recall that $\mu_i = 2\pi\Delta_i \sin(\alpha)$ and, hence

$$\begin{aligned} \mu_i + \Delta\mu_i &= 2\pi\Delta_i \sin(\alpha + \Delta\alpha) \\ &\approx 2\pi\Delta_i \left(\sin(\alpha) + \Delta\alpha \cos(\alpha)\right). \end{aligned}$$

This implies that $\Delta\mu_i = (2\pi\Delta_i \cos(\alpha))\,\Delta\alpha$ and

$$\sigma^2_\mu = (2\pi\Delta_i \cos(\alpha))^2\, \sigma^2_{\alpha,i}, \tag{5.14}$$

Here, the index reference to the baseline in $\sigma_\mu = E\{(\Delta\mu_i)^2\}$ is dropped because $\Delta\mu_i$ is independent of Δ_i. Now, using (5.13), σ^2_μ is expressed in terms of array parameters as

$$\sigma^2_\mu = \frac{1}{SNR}\left(\frac{1}{M_o^2 N}\right), \tag{5.15}$$

Finally, putting (5.15) into (5.11), we find the following expression for $k_{\max}$:

$$k_{\max} = \sqrt{\frac{\mathrm{SNR}}{2}\left(M_o^2 N\left(\frac{1}{f^{-1}_{\Delta k_s}(\mathcal{L})}\right)^2\right) - 1} \tag{5.16}$$

Note that (5.13) and, therefore, (5.16) are derived assuming that there is only one source in the channel. For more than one source (d sources, say), let $\sigma^2_{\mu_j}$ represent the variance of the phase estimation error for the j-th source,[4], then

[4]For more than one sources σ_{μ_j} depends on the SNR in a more complicated way (see chapter 4).

the bound on k_s is generalized as

$$k_{\max} = \min_j \sqrt{\frac{1}{2\sigma_{\mu_j}^2}\left(\frac{1}{f_{\Delta k_s,j}^{-1}(\mathcal{L})}\right)^2 - 1},$$

where $f_{\Delta k_s,j}$ is defined the same way as $f_{\Delta k_s}$ (5.10), but with μ_1 replaced by $\mu_{1,j}$, where $\mu_{1,j}$ is the j-th spatial frequency measured with reference to Δ_1, and $\sigma_{\mu_j}^2 = E\{(\Delta\mu_{1,j})^2\} = E\{(\Delta\mu_{2,j})^2\}$ is assumed to be independent of the base line separation.

5.2.3 Bias on μ due to array imperfections and a self calibrating MR-ESPRIT

Once the winding number n is determined correctly, the next step is to use (5.4) to estimate the spatial frequency μ. If the array is imperfect, the estimate of μ will be biased. The bias (offset) $\Delta\mu$ on μ due to Δk_s can be approximated by (viz. (5.4))

$$\Delta\mu \approx \frac{\partial\mu}{\partial k_s}\Delta k_s = -\frac{1}{k_s^2}(2\pi n + \mu_2)\Delta k_s, \tag{5.17}$$

which indicates that, for a given value of k_s, angles associated with large winding numbers are more affected by Δk_s than those associated with smaller winding numbers.

To minimize the bias, a *self calibrating* MR-ESPRIT, which first estimates the resolution gain factor and then uses this estimated value of k_s, to resolve the ambiguity. To be specific, suppose that we have been processing a sequence of data blocks indexed with some time reference t. Let T be a finite positive integer, and for $j = 1, \ldots, d$, let $\mu_{1,j}(t)$ and $\mu_{2,j}(t)$ represent the coarse and fine spatial frequency estimates of the j-th wave front at a time reference $n \in [t_o, t_o+T-1]$, respectively. Let also $\hat{n}_j(t)$ be the estimate of the corresponding winding number obtained assuming $\Delta k_s = 0$. The idea is to first estimate the resolution gain factor as

$$\hat{k}_s = \frac{1}{T \cdot d}\sum_{t=t_o}^{T-1}\sum_{j=1}^{d}\frac{2\pi\hat{n}_j(t) + \mu_{2j}(t)}{\mu_{1j}(t)}, \tag{5.18}$$

and then insert this estimate into (5.4) for the computation of the spatial frequency μ. Assuming that the mean of the estimation errors is zero, $\hat{k}_s$ asymptotically converges to its true value. An improvement is attained as long as Δk_s is such that $\hat{n}_j(t)$ is correctly estimated. The performance of a self calibrating MR-ESPRIT is compared against a non calibrating MR-ESPRIT in the simulation results.

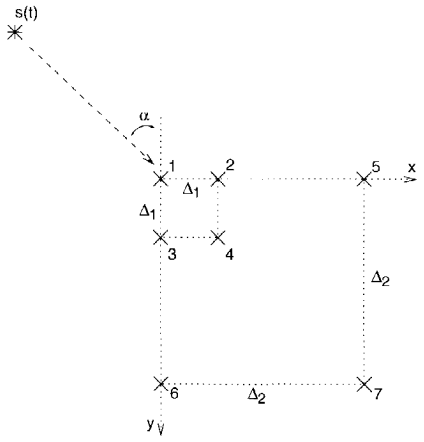


Fig. 5.5: 2-D multi-baseline antenna array

5.3 2-D MR-ESPRIT

The Multi-resolution concept described in section 5.1 can easily be extended into multi-dimensional antenna array. In this section, I show how this can be achieved for a rectangular array structure shown in Fig. 5.5. Let α be the direction of arrival of a signal measured as shown in the figure and let λ be its wavelength. Let for $i = 1, 2$, θ_i and γ_i be defined as

$$\begin{aligned} \theta_i &= e^{j\pi\Delta_i \sin(\alpha)} \\ \gamma_i &= e^{j\pi\Delta_i \cos(\alpha)}. \end{aligned}$$

They represent the parameterized DOAs relative to the baseline separations in the x- and y-directions, respectively. These can be estimated by appropriately defining selection matrices. For instance for the simple example of Fig. 5.5, let a_m represent the response of the m-th antenna element for a signal with direction components $(\theta_i, \gamma_i), i = 1, 2$, then from the shift invariance structure, it follows that

$$\begin{bmatrix} a_2 \\ a_4 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_3 \end{bmatrix} \theta_1 \qquad \begin{bmatrix} a_5 \\ a_7 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_6 \end{bmatrix} \theta_2$$
$$\begin{bmatrix} a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \gamma_1 \qquad \begin{bmatrix} a_6 \\ a_7 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_5 \end{bmatrix} \gamma_2,$$

After θ_i and γ_i are estimated, it remains to combine θ_1 and θ_2 into a single estimate of the argument μ_x of θ_x, and γ_1 and γ_2 into a single estimate of the argument μ_y of θ_y. As in section 5.1, assume that $\Delta_1 < \lambda$, so that $\mu_{x,1} = \arg(\theta_1)$

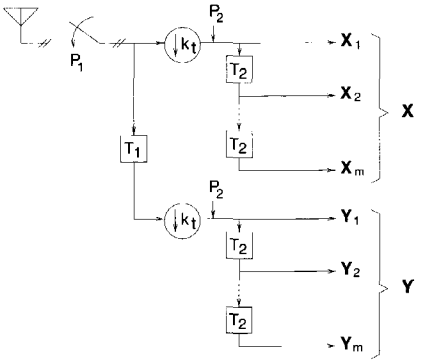


Fig. 5.6: Multi-resolution temporal sampling

and $\mu_{y,1} = \arg(\gamma_1)$ are non aliased and are coarse estimates of the phase shifts μ_x and μ_y. Also assume that $\Delta_2 \gg \lambda$, so that in $\mu_{x,2} = \arg(\theta_2)$ and $\mu_{y,2} = \arg(\gamma_2$ aliasing occur, i.e., the estimates of μ_x and μ_y are proportional to $\mu_{x,2}$ and $\mu_{y,2}$ plus some integer multiple of 2π,

$$\sin(\alpha) = \frac{1}{2\pi\Delta_1}\mu_{x,1} = \frac{1}{2\pi\Delta_2}(2\pi n_x + \mu_{x,2}) \tag{5.19}$$

$$\cos(\alpha) = \frac{1}{2\pi\Delta_1}\mu_{y,1} = \frac{1}{2\pi\Delta_2}(2\pi n_y + \mu_{y,2}). \tag{5.20}$$

The winding numbers n_x and n_y are determined as the best fitting integers to match the two right hand expressions. This results in a similar expression as in (5.3). Once the winding numbers are determined, the DOA is computed as

$$\alpha = \text{atan}\left(\frac{2\pi n_x + \mu_{x,2}}{2\pi n_y + \mu_{y,2}}\right) \tag{5.21}$$

5.4 MR-Joint angle-frequency estimation

Up to now in this chapter, we have been looking at the MR-ESPRIT applied for DOA estimation only. In this section, I discuss how the multi-resolution idea can be extended for JAFE. The resulting algorithm is referred to as the MR-JAFE algorithm.

5.4.1 Model

If the frequency band to be monitored is much wider than the bandwidth of the signals, the data sampling rate P in section 5.1 has to be chosen very large. This implies that the frequency estimates will not be very accurate. To overcome this problem, we can employ the multi-resolution ideas of section 5.1, but now in the temporal domain.

Thus suppose we have two sampling rates, P_1 (fast) and P_2 (intermediate, compared to the source bandwidths), related via a temporal gain factor $k_t = \frac{P_1}{P_2} \gg 1$. It usually suffices to collect only two subsequent samples at the fastest rate P_1, for every m samples at the intermediate rate P_2 as shown in Fig. 5.6. For $i = 1, \ldots, m$, the outputs $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ are all $M \times N$ matrices, where M is the number of antenna elements and N is the number of temporal samples. This means that the data matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are each of size $mM \times N$ and have the same structure as $\boldsymbol{X}_m$ in (3.2). Let $\boldsymbol{Z}$ be the overall data matrix obtained by stacking $\boldsymbol{X}$ and $\boldsymbol{Y}$,

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{bmatrix} \in \mathbb{C}^{2mM,N},$$

then from the structures of $\boldsymbol{Y}$ and $\boldsymbol{X}$, it follows that, with $\phi_{j,i} = e^{j2\pi \frac{f_j}{P_i}}$ and $\boldsymbol{\Phi}_j = \text{diag}\{\phi_{j,i}\}_{i=1}^{d}$, $\boldsymbol{Z}$ has the factorization

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{A}_m \\ \boldsymbol{A}_m \boldsymbol{\Phi}_1 \end{bmatrix} \boldsymbol{S}, \tag{5.22}$$

where $\boldsymbol{A}_m$ is constructed from the array steering matrix $\boldsymbol{A}$ as (see (3.3))

$$\boldsymbol{A}_m = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{A}\boldsymbol{\Phi}_2 \\ \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}_2^{m-1} \end{bmatrix} \in \mathbb{C}^{2mM,d},$$

and $\boldsymbol{S} \in \mathbb{C}^{d,N}$ is a matrix collecting N temporal samples of the d impinging signals. The estimation of the parameters follows by defining three types of selection matrices operating on the data matrix $\boldsymbol{Z}$:

$$\begin{aligned}
\boldsymbol{J}_{x\phi 1} &= [1 \quad 0] \otimes \boldsymbol{I}_m \otimes \boldsymbol{I}_M \\
\boldsymbol{J}_{y\phi 1} &= [0 \quad 1] \otimes \boldsymbol{I}_m \otimes \boldsymbol{I}_M \\
\boldsymbol{J}_{x\phi 2} &= \boldsymbol{I}_2 \otimes [\boldsymbol{I}_{m-1} \quad \boldsymbol{0}] \otimes \boldsymbol{I}_M \\
\boldsymbol{J}_{y\phi 2} &= \boldsymbol{I}_2 \otimes [\boldsymbol{0} \quad \boldsymbol{I}_{m-1}] \otimes \boldsymbol{I}_M \\
\boldsymbol{J}_{x\theta} &= \boldsymbol{I}_2 \otimes I_m \otimes [\boldsymbol{I}_{M-1} \quad \boldsymbol{0}] \\
\boldsymbol{J}_{y\theta} &= \boldsymbol{I}_2 \otimes I_m \otimes [\boldsymbol{0} \quad \boldsymbol{I}_{M-1}]
\end{aligned}$$

Let $\boldsymbol{U}_s$ be a unitary matrix that spans the column space of $\boldsymbol{Z}$, obtained using similar processing steps as before (SVD of $\boldsymbol{Z}$). Let, $\boldsymbol{U}_{x\theta} = \boldsymbol{J}_{x\theta}\boldsymbol{U}_s$, $\boldsymbol{U}_{y\theta} = \boldsymbol{J}_{y\theta}\boldsymbol{U}_s$, and for $i = 1, 2$, let $\boldsymbol{U}_{x\phi i} = \boldsymbol{J}_{x\phi i}\boldsymbol{U}_s$, $\boldsymbol{U}_{y\phi i} = \boldsymbol{J}_{y\phi i}\boldsymbol{U}_s$. From the shift invariance property, we obtain the model

$$\begin{aligned} \boldsymbol{U}_{x\phi 1}^{\dagger}\boldsymbol{U}_{y\phi 1} &= \boldsymbol{T}^{-1}\boldsymbol{\Phi}_1\boldsymbol{T} \\ \boldsymbol{U}_{x\phi 2}^{\dagger}\boldsymbol{U}_{y\phi 2} &= \boldsymbol{T}^{-1}\boldsymbol{\Phi}_2\boldsymbol{T} \\ \boldsymbol{U}_{x\theta}^{\dagger}\boldsymbol{U}_{y\theta} &= \boldsymbol{T}^{-1}\boldsymbol{\Theta}\boldsymbol{T} \end{aligned}$$

A joint diagonalization of the three matrices now provides estimates of (ϕ_1, ϕ_2, θ) of each source. If the array has an additional spatial multi-resolution structure, then $\boldsymbol{\Theta}$ splits into $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$ and can be estimated with multi-resolution ESPRIT as discussed in section 5.1.

To complete our estimation procedure, we need to estimate for each source the argument μ of ϕ by combining ϕ_1 and ϕ_2 in the same way as we did for the DOA estimation in section 5.1. Let $f_{\max}$, be the upper bound of the frequency band of interest, we assume that the two sampling frequencies P_1 and P_2 are such that $P_1 > f_{\max}$ and $P_2 \ll f_{\max}$. Based on these two sampling rates we obtain two estimates of the frequency f: one (based on P_1) at a coarser scale with no aliasing and one (based on P_2) at a finer scale but with aliasing, i.e.,

$$f = \frac{P_1}{2\pi}\mu_1 = \frac{P_2}{2\pi}(2\pi n + \mu_2). \tag{5.23}$$

The winding number n is determined as before using

$$n = \text{round}\frac{1}{2\pi}(\frac{P_1}{P_2}\mu_1 - \mu_2)$$

Note that if μ_1 and μ_2 have estimation errors of equal size, then the noise on the second estimate is a factor $k_t = \frac{P_1}{P_2}$ smaller than that on the first estimate. Thus, we would either use the second equation to estimate f_i, or optimally combine the two estimates using an equation similar to (5.5).

5.4.2 Complexity reduction achieved by MR-JAFE

Recall from chapter 3 that the data sampling rate used in constructing the data matrix of (3.2) is P times the Nyquist rate of the baseband signals. Since P can be quite large, it would be very expensive to construct a full dat matrix of all samples. In section 5.4, I have shown that with the MR-JAFE, it is sufficient to subsample: collect 2 subsequent samples at a rate $P_1 = P$ followed by m samples at a rate $P_2 = P/k_t$. In the following, I give analysis comparing the computational complexities of the MR-JAFE and the JAFE algorithms.

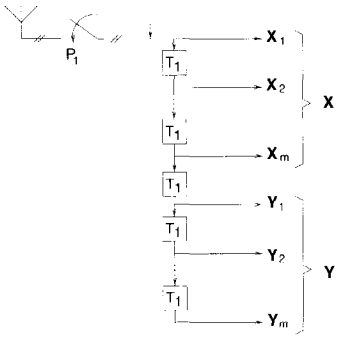


Fig. 5.7: Collecting data for JAFE using a single sampling rate

Though only the frequency estimation process is considered, a similar result is obtained for the DOA estimation also.

To be specific, suppose that an M element ULA with baseline separation Δ wavelengths is impinged by a single far field signal $s(t)$ with DOA α and center frequency f. Let the band of interest be such that we sample with the rate P to satisfy the Nyquist sampling rate, and let $T := 1/P$. To make the comparison fair, in the first instance, assume that we have collected the $2mM \times N - m + 1$ partitioned data matrix $\boldsymbol{Z}$ using the measurement setup of Fig. 5.7. This data matrix has the same number of rows as that obtained using Fig. 5.6, but its structure is similar to that of $\boldsymbol{X}_m$ in (3.2) with m replaced by $2m$. In chapter 4, for this data model, it has been shown that the variances of the frequency estimation error, ε, corresponding to the optimum value of the temporal smoothing factor $m_{opt} = \frac{2N+3}{3}$ is such that

$$\varepsilon \propto \frac{1}{T^2} \frac{1}{(N-1)^3}. \tag{5.24}$$

In the second instance, assume that we have collected a $2mM \times N_t - m + 1$ MR data matrix $\boldsymbol{Z}$ using the measurement setup of Fig. 5.6, where $N_t = \lfloor N/k_t^{2/3} \rfloor$. The variances of the frequency estimation error ($\hat{\varepsilon}$), corresponding to the optimum value of the temporal smoothing factor $\hat{m}_{opt} = \frac{2N_t+3}{3}$ is then obtained by replacing N with N_t, T with $k_t T$ in (5.24),

$$\hat{\varepsilon} \propto \frac{1}{k_t^2 T^2} \frac{1}{(N_t-1)^3}. \tag{5.25}$$

It is seen that, for $N_t \gg 1$,

$$\varepsilon \approx \hat{\varepsilon}.$$

However, since $k_t > 1$ and

$$\begin{aligned} \hat{m}_{opt} &= \frac{m_{opt}}{k_t} < m_{opt}, \\ N_t &= \frac{N}{k_t^{2/3}} < N, \end{aligned}$$

the complexity of the MR-JAFE is much smaller than that of the direct approach. Generally, if we let $\varepsilon(m)$ and $\varepsilon_{mr}(m)$ represent the variances of the direct and the multi-resolution approaches at a temporal smoothing factor of m, respectively, then

$$\varepsilon_{mr}(m) \approx \varepsilon(k_t m)$$

This essentially states that, with $N_t \approx N/k_t^{2/3}$, the MR-JAFE achieves the same performance as that of the direct JAFE with a temporal smoothing factor of only m/k_t. Since the computational complexity is proportional to $(mM)^2 N$, this provides with almost a factor $k_t^{2.7}$ reduction in the complexity.

5.5 Simulations

In this section, I give simulation results that demonstrate the claims of this chapter. The simulation example considers a processing band of 10 MHz and a linear antenna array with $M = 4$ antenna elements arranged as in Fig. 5.1 with $\Delta_1 = \frac{1}{2}$ and varying Δ_2. The data is collected into a 4×64 matrix at a sampling rate of $P_1 = 20$MHz. Two sources emitting narrowband signals (25 kHz) at center frequencies $\boldsymbol{f} = [6, 6.5]$ MHz, and propagating in distinct directions with DOAs $\boldsymbol{\alpha} = [40, 45]$ degrees are considered.

The simulation results are shown in Fig. 5.8 through Fig. 5.11. From the first plot, it is seen that the accuracy of MR-ESPRIT is indeed proportional to the gain factor k_s. An upper limit for this gain is reached when the winding numbers n can no longer be estimated accurately. This is shown in Fig. 5.9, where the RMSE of the parameter estimator as a function of varying k_s is plotted. To make the figure less crowded, only the behavior corresponding to DOA $= 45$ degrees is shown. It is seen that, for a given SNR, there exists a limit on k_s beyond which the performance of the estimator degrades sharply. Moreover, this bound is seen to be proportional to the SNR as expected. A simple intuitive analysis can give us a good incite. In Fig. 5.2, it is seen that as k_s increases, the number of alias regions increases proportionally and the extent of the alias regions $2\pi/k_s$ gets increasingly narrow. In the limit, $2\pi/k_s$ becomes comparable to the noise variance and, as the result, the alias zone can no longer be correctly estimated.

Simulations showing the improvements in the biases due to self calibration are shown in Fig. 5.10, where again for the sake of readability, only the behavior corresponding to DOA $= 45$ degrees is shown. . The results correspond to $k_s = 20$, and the averaging for the self calibration is taken with $T = 20$. It is seen that when the bias $\Delta k_s > 0.05$, the self calibrating MR-ESPRIT gives lower bias. This means that, if array placement errors are expected to be large, it is advantageous to implement a self calibrating MR-ESPRIT. Moreover, one can see that there exists a limit on Δk_s beyond which the self calibration process fails. This is because, when Δk_s exceeds this point the winding number n is not correctly estimated, and consequently the estimated k_s is also far removed from the actual value. Moreover, it is seen that as the SNR is decreased the breakaway point gets smaller because, as seen in (5.9), the effect of Δk_s on n is sever at low SNR.

The effect of MR temporal sampling on the frequency estimation is summarized in Fig. 5.11,[5] where results obtained via the MR-JAFE, with an accuracy gain factor of $k_t = 2^{3/2}$ are compared against those of the direct estimation method. The superiority of the MR-based approach is obvious, particularly at small m values. In fact, as discussed in section 5.4, the MR-JAFE achieves the same accuracy as the direct method at a temporal smoothing factor of only $m_r = m/2$.

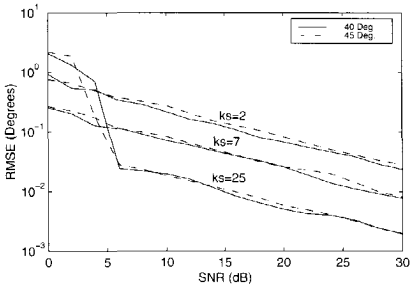


Fig. 5.8: The root mean square error of the frequency estimates as functions of SNR. ($k_s = 2$ corresponds to ULA)

5.6 Conclusion

In this chapter, I have shown that the MR-ESPRIT uses a multiple shift invariance structure induced via multiple spatio-temporal sampling rates to estimate signal parameters. A short sampling period is required to avoid aliasing, a

[5] With an MR spatial sampling, a similar behavior is exhibited in DOA estimation.

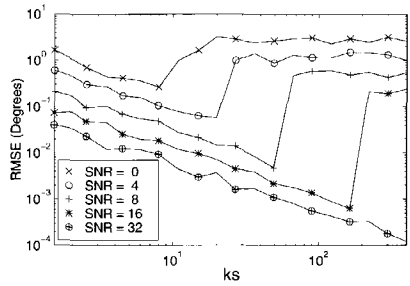


Fig. 5.9: Root mean square errors of the DOA estimates, corresponding to the wavefront with DOA $= 45^o$, as functions of k_s.

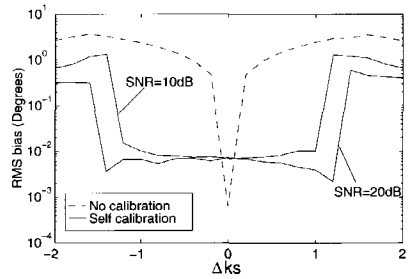


Fig. 5.10: The bias on DOA estimation due to imperfect array ($k_s = 20$, $T = 20$). The non-calibrated behavior is shown for SNR=20dB only, its behavior at 10dB is approximately the same.

long sampling period is preferred for accuracy. It has been shown that the MR-ESPRIT algorithm allows to combine the two estimates to obtain a better estimation accuracy. The ratio of the longest sampling period to the shortest one is a measure of the gain in accuracy. Because of various factors, such as noise and measurement error, the achievable gain in accuracy is shown to be bounded.

The MR-ESPRIT is also successfully applied to joint angle-frequency estimation, resulting in the MR-JAFE algorithm. I have shown that the MR-JAFE achieves the same performance level as the JAFE algorithm with a factor k^2 less complexity, where k is the gain in accuracy.

Finally, it should be noted that all the data stacking techniques, namely, temporal smoothing, spatial smoothing and forward backward averaging can be

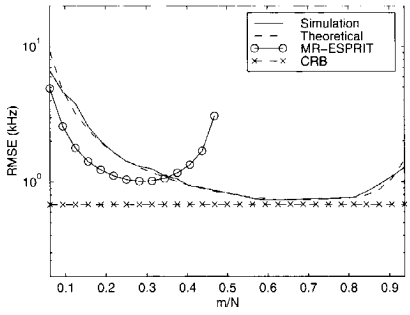


Fig. 5.11: MR-ESPRIT based JAFE improving the performance of the frequency estimation in the small m region. Only the plot corresponding to DOA $= 45^o$ is shown.

incorporated into the above algorithm. Moreover, the forward backward averaged data can be transformed into the real domain as discussed in section 3.8.

Chapter 6

PRACTICAL CONSIDERATIONS

Contents

One way of analyzing the robustness of a given algorithm is to see how it behaviors under non-ideal situations. To this end, in this chapter, I consider different practical cases where the data models do not perfectly fit into the JAFE model discussed in chapter 3. These include, effects of frequency hopping sources, effects of moving sources and moving receivers, effects of frequency and angle spreading, and effects of antenna coupling.

The chapter is organized as follows. First, in section 6.1, I show that, in spite of model mismatch caused by frequency hopping (FH) signals, the JAFE algorithm can estimate the angle and the pre- and post-hopping frequencies of a FH signal. Following this, in section 6.2, I discuss effects of frequency and/or angel spreading on the JAFE algorithm. Here, I show that, in addition to angel and frequency, the algorithm can also be used to estimate the bandwidths of signals. When the signal sources and the receiver are moving, the DOA can no longer be assumed to be fixed, and the movement might introduce some smearing of the signals' frequency spectra. In section 6.3, these effects are considered. In most cases, the array elements are assumed to behave independent of each other. In practice, however, there exists mutual coupling between them. In section 6.4, it is shown that unless the actual array behavior is appropriately modeled, the so-called high resolution parameter estimation algorithms will not achieve their theoretical performances. Finally, the chapter is concluded with some useful remarks and recommendations. At each stage, I give simulation results demonstrating the results claimed in the chapter.

6.1 Slow Frequency Hopping Signals

Frequency hopping (FH) signals are a class of signals whose carrier frequencies change periodically. The FH approach allows the system to incorporate frequency diversity, which minimizes the performance degradation due to slow fading channels, as well as interference diversity, which ensures that the system is not subject to a worst case interference scenario. FH signals can be classified as fast FH (FFH) and slow FH (SFH) depending on the ratio of the frequency hopping rate to the symbol rate. When the hoping rate is smaller than the symbol rate, the hopping is said to be slow, otherwise it is said to be fast. SFH signals appear in military as well as communication scenarios [18]. From some perspectives, it is required to localize these SFH signal sources. For instance, in emergency and police applications, it is required to blindly localize mobile telephones. If the signals occupy a fixed frequency band, the results of chapter 2 can directly be used to localize the sources. On the other hand, if the sources hop in a wide frequency band, we need to extend these results. In this

section, we discuss how this can be achieved.

The proposed method makes use of the joint angle frequency estimation (JAFE) algorithm of chapter 3 to separate and track SFH signals. As mentioned in that chapter, one advantage of joint estimation is that the signal parameters are paired for free. Here, this property is exploited to track signal sources whose carrier frequencies change randomly within a certain frequency band. The idea is to use the JAFE algorithm to simultaneously estimate the angle and frequency of each source, and then to use the angle information to correctly identify the source and its hopping frequency.

6.1.1 The data model

Consider an M element uniform linear antenna array (ULA), with baseline separation of Δ, where Δ is measured in terms of fractions of signal wavelength. Let $\boldsymbol{a}(\theta)$ be the response of the array to a far field narrowband signal with the parameterized direction of arrival (DOA) $\theta := e^{j2\pi\Delta\sin(\alpha)}$, where α is the signal DOA measured with respect to the normal of the array axis as described in Fig. 2.4,

$$\boldsymbol{a}(\theta) = \begin{bmatrix} 1 \\ \theta \\ \vdots \\ \theta^{M-1} \end{bmatrix}.$$

Let f_i be the intermediate frequency (IF) of the i-th source and suppose that d narrowband sources $s_i(t)e^{jc\pi f_i t}$, $i = 1, \ldots, d$ are impinging on a ULA. Let the width of the band of interest be such that, after demodulation to IF, we can sample at a rate P to satisfy the Nyquist rate. Without multi-path, the noise free data model of the modulated sources at the receiver is (viz. section 3.1)

$$\boldsymbol{x}(\tfrac{n}{P}) = \sum_{1}^{d} b_i \boldsymbol{a}(\theta_i) e^{j\frac{2\pi}{P} f_i n} s_i(\tfrac{n}{P}),$$

where b_i is the is amplitude of the i-th source. Let $\phi_i = e^{j\frac{2\pi}{P} f_i}$, $\boldsymbol{B} = \text{diag}\{b_i\}_{i=1}^{d}$, $\boldsymbol{\Phi} = \text{diag}\{\phi_i\}_{i=1}^{d}$, $\boldsymbol{A}$ be an $M \times d$ matrix that contains the d steering vectors corresponding to the d sources and $\boldsymbol{s}$ be a $d \times 1$ signal vector, then the above can be written in matrix form as

$$\boldsymbol{x}(\tfrac{n}{P}) = \boldsymbol{A}\boldsymbol{\Phi}^n \boldsymbol{B}\boldsymbol{s}(\tfrac{n}{P}), \tag{6.1}$$

Assume, for the moment, that a single narrowband signal $s(t)$, propagating in a distinct path and frequency hopping at a rate smaller than the signal symbol

rate is impinging on an M element ULA. Although the results can easily be generalized to more than one sources, we restrict ourselves to a single source scenario for simplicity reasons.

Suppose that we have collected N samples of the array output (3.1) at a rate P into an $M \times N$ data matrix $\boldsymbol{X}$. Also suppose that, at time instant $n_o \in [0, N-1]$, the signal $s(t)$ hops from f_a to f_b. We assume that in one data matrix there is at most one frequency hop per signal. This puts bounds on the time extent of the data matrix, which should be less than the shortest hopping period. Short data extent means limited estimation accuracy. However, since we are considering SFH signals, the extent of the data matrix can be made short enough to allow at most a single frequency hop per source without significantly degrading the estimation accuracy. Let ϕ_a and ϕ_b, and $\theta(f_a)$ and $\theta(f_b)$ be the parameterization of the center frequencies and DOAs of the signal at frequencies f_a and f_b, respectively. We model the FH signal with two intermittent sources, one operating during the interval $[0, n_o]$ and the other one operating in the interval $[n_o + 1, N-1]$. To get compact expressions, let for some phase offset x,

$$\boldsymbol{F}_{n_o} = \begin{bmatrix} s(0) & \phi_a s(\frac{1}{P}) & \cdots & \phi_a^{n_o} s(\frac{n_o}{P}) \\ 0 & 0 & \cdots & 0 \\ & 0 & \cdots & 0 \\ & \phi_b^{n_o+1+x} s(\frac{n_o+1}{P}) & \cdots & \phi_b^{N-1+x} s(\frac{N-1}{P}) \end{bmatrix} \tag{6.2}$$

and

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ \theta(f_a) & \theta(f_b) \\ \vdots & \vdots \\ \theta(f_a)^{M-1} & \theta(f_b)^{M-1} \end{bmatrix} \tag{6.3}$$

then the noise free data at the output of a ULA can be expressed as

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{F}_{n_o} \tag{6.4}$$

where $\boldsymbol{B} = b\boldsymbol{I}$ is the amplitude of the signal. For the sake of simplicity, in the remainder of the section, we shall assume that the phase offset $x = 0$. Let

$$\boldsymbol{F}(t) = \begin{bmatrix} s(t) & s(t+\frac{1}{P})\phi_a & \cdots & s(t+\frac{N-1}{P})\phi_a^{N-1} \\ s(t) & s(t+\frac{1}{P})\phi_b & \cdots & s(t+\frac{N-1}{P})\phi_b^{N-1} \end{bmatrix}$$

and let the finite support function $\boldsymbol{y}(t, t_o)$ be defined as

$$\boldsymbol{y}(t, t_o) = \begin{cases} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, & \text{if } 0 \le t \le t_o \\ \\ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, & \text{if } t_o < t \le N-1 \end{cases}$$

and

$$\boldsymbol{Y}(t,t_o) := \left[\ \boldsymbol{y}(t,t_o) \quad \boldsymbol{y}(t+\tfrac{1}{P},t_o) \quad \cdots \quad \boldsymbol{y}(t+\tfrac{N-1}{P},t_o)\ \right].$$

Then, we can express $\boldsymbol{X}$ more compactly as

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{B}(\boldsymbol{F}(0) \odot \boldsymbol{Y}(t,\frac{n_o}{P})), \tag{6.5}$$

where $\odot$ stands for a point-wise matrix product. If there are d signals in the channel and say $r \leq d$ of the signals hop to a new set of carrier frequencies, the above model can be generalized to the case where $\boldsymbol{A}$ is an $M \times (d+r)$ matrix, $\boldsymbol{B}$ is a $(d+r) \times (d+r)$ diagonal matrix and $\boldsymbol{F}(0)$ is a $(d+r) \times N$ matrix.

6.1.2 Applying the JAFE Algorithm

To apply the JAFE algorithm we need to construct a data matrix with a dual shift invariance property. To this end, we stack time shifts of the array output into $mM \times N - m + 1$ block Hankel data matrix $\boldsymbol{X}_m$ as discussed in section 3.2.2. With the model of $\boldsymbol{X}$ in (6.5), $\boldsymbol{X}_m$ can be expressed as,

$$\boldsymbol{X}_m = \begin{bmatrix} \boldsymbol{A}\boldsymbol{B}(\boldsymbol{F}(0) \odot \boldsymbol{Y}(t,\frac{n_o}{P})) \\ \boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{B}(\boldsymbol{F}(\frac{1}{P}) \odot \boldsymbol{Y}(t,\frac{n_o-1}{P})) \\ \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}^{m-1}\boldsymbol{B}(\boldsymbol{F}(\frac{m-1}{P}) \odot \boldsymbol{Y}(t,\frac{n_o-m+1}{P})) \end{bmatrix}. \tag{6.6}$$

In section 3.2.2, we have made the assumption that $\boldsymbol{F}(0) \approx \cdots \approx \boldsymbol{F}(\frac{m}{P}) =: \boldsymbol{F}_s$, from which we get the approximation

$$\boldsymbol{X}_m \approx \begin{bmatrix} \boldsymbol{A}\boldsymbol{B}(\boldsymbol{F}_s \odot \boldsymbol{Y}(t,\frac{n_o}{P})) \\ \boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{B}(\boldsymbol{F}_s \odot \boldsymbol{Y}(t,\frac{n_o-1}{P})) \\ \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}^{m-1}\boldsymbol{B}(\boldsymbol{F}_s \odot \boldsymbol{Y}(t,\frac{n_o-m+1}{P})) \end{bmatrix}. \tag{6.7}$$

For the algorithm to work, the above data model has to be reduced to the form given in (3.3). That is, we want to find a k such that $\boldsymbol{X}_m$ can be approximated with

$$\boldsymbol{X}_k = \boldsymbol{A}_m\boldsymbol{B}(\boldsymbol{F}_s \odot \boldsymbol{Y}(t,n_o-k)), \tag{6.8}$$

where $\boldsymbol{A}_m$ is as defined in (3.4). As a simple rule, we say k is optimum if it minimizes the number of non zero terms in the difference

$$\boldsymbol{X}_m - \boldsymbol{X}_k.$$

Lemma 6.1 *Consider a single far field, narrowband, frequency hopping signal impinging on an M element ULA. Let $\boldsymbol{X}_m$ be an extended data matrix obtained by stacking m time shifts of the array output as in (6.6). Suppose that, for some $k \in [0, m-1]$, we approximate $\boldsymbol{X}_m \in \mathbb{C}^{mM, N-m+1}$ with*

$$\boldsymbol{X}_k = \boldsymbol{A}_m \boldsymbol{B} (\boldsymbol{F}_s \odot \boldsymbol{Y}(t, n_o - k),$$

then the optimum value of k that minimizes the number of non zero terms in $\boldsymbol{X}_m - \boldsymbol{X}_k$ is given by $k_{opt} = \frac{m}{2}$

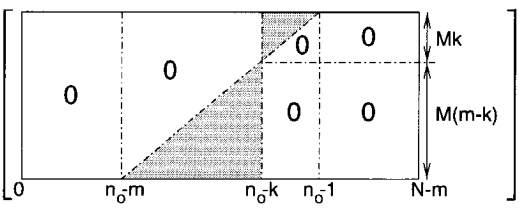
Proof:- Let $\boldsymbol{F}_k := \boldsymbol{F}_s \odot \boldsymbol{Y}(t, n_o - k)$, and let $\boldsymbol{a}_j$ be the j-th column of $\boldsymbol{A}_m$ and b_j be the jj-th entry of the diagonal matrix B, then the difference

$$\varepsilon = \boldsymbol{X}_k - \boldsymbol{X}_m \tag{6.9}$$

may be written as

$$\varepsilon = \sum_{p=1}^{m} \sum_{j=1}^{2} b_j \boldsymbol{a}_j (\boldsymbol{F}_k - \boldsymbol{F}_p) \tag{6.10}$$

From the definition of $\boldsymbol{Y}(t, n_o - k)$, it is seen that for $p = 0, \cdots, m-1$, $\boldsymbol{F}_k$ and $\boldsymbol{F}_p$ are identical except for $|p-k|$ non zero entries in the transition region. Thus, the above sum has non-zero entries only in the shaded triangular region shown in (6.11).

$\varepsilon =$ 

. (6.11)

Hence, the number of non-zero terms is given by

$$Q(k) = \frac{M}{2}((m-k)^2 + k^2)$$

Setting the derivative of this to zero the optimum value of k is obtained,

$$k_{opt} = \frac{m}{2}$$

□

Let ε_{opt}^2 be the modeling error defined as

$$\varepsilon_{opt}^2 = \|\boldsymbol{X}_{k_{opt}} - \boldsymbol{X}_m\|_F^2.$$

Then, noting that the maximum value of the magnitude of each of the non-zero entries is 2,

$$\varepsilon_{opt}^2 \leq 4b^2 Q(k_{opt}) = Mb^2 m^2$$

Let σ^2 be the noise power. If the variance of the modeling error defined by

$$\sigma_\epsilon^2 := \frac{\varepsilon_{opt}^2}{mMN}$$

is such that $\sigma_\epsilon^2 < \sigma^2$, the modeling error is dominated by the noise and is, therefore, said to be negligible. Defining the signal to noise ratio as SNR $= b^2/\sigma^2$, this is satisfied if

$$\frac{m}{N} < \frac{1}{\text{SNR}}. \tag{6.12}$$

The above states that, when the temporal stacking parameter m is smaller than $\frac{N}{\text{SNR}}$, the error introduced by model mismatch may be neglected,. In this case, the factorization in (6.5) is a useful approximation of the original data model. Nevertheless, due to the approximation, some biases are introduced into the parameter estimates. It is shown in the simulation results of section 6.1.7 that these biases are minimal when the condition in (6.12) is satisfied.

Example 6.1 *Consider two sources with carrier frequencies* $\boldsymbol{f}_a = [2, 3]$ *MHz and angles of arrivals DOA* $= [30, 45]$ *degrees, respectively impinging on a ULA with* $M = 4$ *elements. The outputs of the antenna array are arranged into a block Hankel data matrix with* $N = 64$ *and varying* m. *Let us assume that at time* $n_o = 32$, *the center frequency of the first source changes from* 2 *MHz to* 4 *MHz. The behaviors of the singular values of the data matrix as functions of* m, *for SNR* $= 10$ *dB, are shown in Fig. 6.1. From the plots we see that, for*

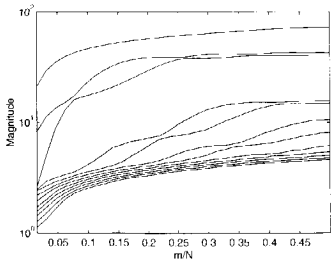


Fig. 6.1: The singular values as functions of $m(n_o = 32)$

$m/N < 0.1 = 10^{-SNR/10}$, *the rank of the data matrix is essentially equal to* 3 *(the number of angle frequency pairs including the pre- and post-hopping frequencies). However, as* m *is increased beyond this point, additional singular*

values start to pop up from the noise floor, and the model becomes increasingly erroneous.

6.1.3 Undetectable Frequency hops

From the model of (6.2), we see that the number of temporal samples associated to the pre- and post-hopping frequencies are n_o and $N - n_o$, respectively. Moreover, I have shown that, under ideal conditions, the extended data matrix will have a rank that shows the presence of frequency hopping signals. The question is now whether this is also the case when there is some noise. The answer to this question is non-affirmative. To show this, consider the model of (6.5) and let $\boldsymbol{Y}(f, n_o)$ be the Fourier transform of $\boldsymbol{Y}(t, n_o)$. $\boldsymbol{Y}(f, n_o)$ consists of two sinc functions with the amplitudes n_o and $N - n_o$, respectively. The Fourier transform of $\boldsymbol{X}$ can then be expressed as

$$\boldsymbol{X}(f) = \boldsymbol{A}\boldsymbol{B}(\boldsymbol{F}_s(f) * \boldsymbol{Y}(f, n_o)) \tag{6.13}$$

where $*$ stands for a row-wise convolution. Since the two rows of $|\boldsymbol{Y}(f, n_o)|^2$ have peak amplitudes of n_o and $N - n_o$, respectively, the corresponding rows of $|\boldsymbol{F}_s(f) * \boldsymbol{Y}(f, n_o)|^2$ will have peaks that are proportional to n_o and $N - n_o$, respectively. More precisely, if we let b represent the amplitude of the incoming signal, then the peaks $\boldsymbol{p}^2$ of the two rows of $|\boldsymbol{B}(\boldsymbol{F}_s(f) * \boldsymbol{Y}(f, n_o))|^2$ are given by

$$\boldsymbol{p}^2 = b^2 \begin{bmatrix} n_o \\ N - n_o \end{bmatrix}.$$

Let σ^2 be the variance (power) of the noise, then the frequency hopping is detectable if $\boldsymbol{p}^2 > \sigma^2$. Defining the SNR as b^2/σ^2, the detectablity condition becomes

$$n_o > \frac{1}{\text{SNR}} \quad \text{and} \quad N - n_o > \frac{1}{\text{SNR}}.$$

This states that, when the hopping instant is close to the boundaries of the data matrix, it is not detected, and the data matrix given in (6.4) will essentially have a rank that does not show the presence of a hopping source. This property is illustrated in the following example.

Example 6.2 *Consider the two sources of example 6.1. The plots of the* 8 *largest singular values for* $m = 4$ *and as* n_o *is varied from* 0 *to* N *are shown in Fig. 6.2. As clearly seen from the plots, the third largest singular value gradually detaches itself from the noise level. For* n_o *sufficiently away from* 0 *and* N*, the data covariance matrix is essentially of rank 3, and the JAFE algorithm produces three pairs of angle-frequency estimates. On the other*

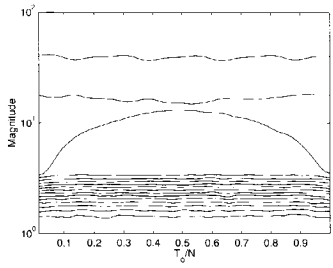


Fig. 6.2: The singular values as functions of n_o

hand, as n_o approaches to 0 or N (i.e., as n_o approaches to the edges of the data block), the data matrix becomes of rank 2. Consequently, the JAFE algorithm produces only two pairs of angle-frequency estimates. No frequency hopping is detected in this case. When $n_o \to N$ only the pre-hopping frequency is detected and when $n_o \to 0$ only the post-hopping frequency is detected.

6.1.4 Estimation and Tracking

The first step in the JAFE algorithm is to estimate the signal subspace from the collected data. For the narrowband signal model, and a temporally smoothing factor m satisfying the condition in (6.12), the effective rank of the temporally smoothed data matrix of (6.6) is equal to the number of signal sources. According to our discussion in section 6.1.1, some of the detected signals might not physically exist, they may be results of FH. This is determined only after all the frequency-DOA pairs have been determined. Once the number of sources (physically existing plus those generated by FH signals) is estimated, the second step is to use the JAFE algorithm to estimate all the angle frequency pairs (viz section 3.2.2). Assuming that the signals have enough spatial separations and that they have distinct DOAs, the the third step is to identify the frequency hopping signals by matching the frequencies associated to the same DOA. In example 6.2, we have seen that, under noisy conditions, when the hopping instant n_o is close to the edges of the data block, only one of the frequencies is detected. In this case, however, two consecutive data blocks produce two different frequencies associated to the same DOA. Assuming the DOAs are unique, we can easily determine the hopping source and its new carrier frequency.

6.1.5 Signal Copy

From some perspectives, it is also required to listen to the signals. For this, we need to locate the exact switching instant. This is achieved by analyzing the phase of the reconstructed IF signal associated with the frequency hop. Let $\boldsymbol{w}(\alpha)$[1] be the beamforming vector that correspond to the direction α (angle of arrival of the hopping signal s), then the signal $\boldsymbol{s} = \boldsymbol{w}(\alpha)\boldsymbol{X}$ is approximately given by

$$\boldsymbol{s} = \left[\begin{array}{ccccccc} s(0) & \phi_a s(\frac{1}{P}) & \ldots & \phi_a^{n_o} s(\frac{n_o}{P}) & \phi_b^{n_o+1} s(\frac{n_o+1}{P}) & \ldots \end{array} \right],$$

Let $\phi(n) = \phi_a^n$, for $n \leq n_o$ and $\phi(n) = \phi_b^n$, otherwise. For $n = 1, 2, 3, \ldots$, define an instantaneous signal frequency $f(n)$ as

$$f(n) = \frac{P}{2\pi} \text{imag} \left(\log(\phi(n) s(\frac{n}{P})) - \log(\phi(n-1) s(\frac{n-1}{P})) \right) \tag{6.14}$$

Assuming that $s(\frac{n}{P}) \approx s(\frac{n-1}{P})$, we have

$$f(n) = \begin{cases} f_a, & n \leq n_o \\ f_b - f_a & n = n_o + 1 \\ f_b, & n > n_o + 1 \end{cases}, \tag{6.15}$$

$f(n)$ is plotted in Fig. 6.3, for the simulation setup described in *Example 6.1*. It

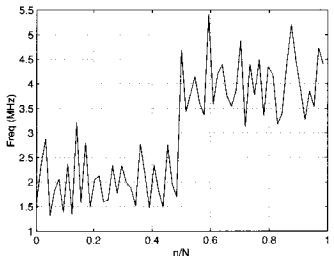


Fig. 6.3: Behavior of the instantaneous frequency $f(n)$ (SNR = 10 dB)

is seen that by observing the instantaneous frequency of the IF signal, one can locate the switch instant. If the source changes frequency at a known period, we need to determine the switch moment only once per connection.

[1]Note that $\boldsymbol{w}(\alpha)$ is readily determined from the estimated array steering matrix.

6.1.6 Analysis

In this section, we give analysis of the two steps involved in frequency estimation and tracking of FH signals. These are 1) estimating the angle frequency pairs (JAFE) and 2) the estimation of the instantaneous frequency, which is used to determine the hopping instant.

Analysis of JAFE

A detailed analysis of the JAFE algorithm is given in chapter 4. In the following, I present a summary of these results. It should be noted that, from our perspective, the results in chapter 4 are a bit optimistic as they don't consider the modeling error introduced by the frequency hopping signals. As it will soon be clear, this does not introduce any discrepancy in the analysis presented. Let $\boldsymbol{X}_m \in \mathbb{C}^{mM,N-m+1}$ be a data matrix obtained by performing an m-factor temporal smoothing (viz. sections 3.1). It has been shown in chapter 4 that the variances of the angle and frequency estimation errors, respectively, are such that

$$\begin{aligned} \sigma_\theta^2 &\propto \frac{2}{(M-1)^2 m(N-m+1)} \\ \sigma_\phi^2 &\propto \frac{1}{M(m-1)^2(N-m+1)} \end{aligned}$$

Let $\mu = m/N, \epsilon = 1/N$, and let K_ϕ and K_θ be the proportionality constants for the above equations, then it follows that

$$\sigma_\theta^2 = K_\theta \frac{2}{(M-1)^2 N^2 \mu(1-\mu+\epsilon)} \tag{6.16}$$

$$\sigma_\phi^2 = K_\phi \frac{1}{MN^3(\mu-\epsilon)^2(1-\mu+\epsilon)} \tag{6.17}$$

Now, let $\eta_o = n_o/N$, then the number of temporal samples associated with pre-hopping parameters is $\eta_o N$, and that corresponding to the post-hopping parameters is $(1-\eta_o)N$. Thus, the variances of the estimation errors of the pre- and post-hopping parameterized DOAs, respectively, are given by

$$\sigma_{\theta,o}^2 = \frac{\sigma_\theta^2}{\eta_o^2} \quad \text{and} \quad \sigma_{\theta,f}^2 = \frac{\sigma_\theta^2}{(1-\eta_o)^2}. \tag{6.18}$$

Similarly, the variances of the errors on the pre- and post-hopping parameterized frequencies, respectively, are

$$\sigma_{\phi,o}^2 = \frac{\sigma_\phi^2}{\eta_o^3} \quad \text{and} \quad \sigma_{\phi,f}^2 = \frac{\sigma_\phi^2}{(1-\eta_o)^3}. \tag{6.19}$$

These results are verified using simulations in section 6.1.7.

Estimating the Hopping instant

Consider the situation where a continuous signal hops from frequency f_a to f_b. Let $\Delta f_{ab} = f_b - f_a$ represent the size of the frequency hop. We are interested to know the probability of correctly detecting the frequency hopping instant. As described above, the hopping instant is determined by inspecting the instantaneous frequency $f(n)$. In this section, we take a closer look at the computation of $f(n)$ given in (6.14). Let $\nu(n)$ [2] be a noise term at time n, and let M be the number of antenna elements. Then, the measurement sample at the output of the beamformer is approximately given by

$$\hat{\phi}(n)\hat{s}(n) \approx \phi(n)s(n) + \frac{1}{M}\nu(n),$$

from which it follows that $\hat{\phi}(n)\hat{s}(n)$ is Gaussian with variance

$$\sigma^2_{\hat{\phi}} = \frac{1}{M^2}\frac{1}{\text{SNR}}, \tag{6.20}$$

where SNR is the signal to noise ratio defined as $E\{\frac{s^2(n)}{\nu^2(n)}\}$. Here, the argument n is dropped because the variance is independent of n. The next step is to compute the statistical behavior of $f(n)$ defined in (6.14). To this end, assuming $s(n) \approx s(n-1)$, we may approximate (6.14) as

$$f(n) \approx \frac{P}{2\pi}\text{imag}(\log\hat{\phi}(n) - \log\hat{\phi}(n-1)). \tag{6.21}$$

Noting that $f(n)$ is a function of the two parameters $\hat{\phi}(n)$ and $\hat{\phi}(n-1)$, its variance can easily be computed using the transformation rules derived in [39]:

$$\sigma^2_f = \begin{bmatrix} \frac{\partial f(n)}{\partial\hat{\phi}(n)} & \frac{\partial f(n-1)}{\partial\hat{\phi}(n-1)} \end{bmatrix} \begin{bmatrix} \sigma^2_{\hat{\phi}} & 0 \\ 0 & \sigma^2_{\hat{\phi}} \end{bmatrix} \begin{bmatrix} \left(\frac{\partial f(n)}{\partial\hat{\phi}(n)}\right)^H \\ \left(\frac{\partial f(n-1)}{\partial\hat{\phi}(n-1)}\right)^H \end{bmatrix}$$

Substituting for $f(n)$ from (6.21) and after some simplification this reduces to

$$\sigma^2_f = \frac{P^2}{4\pi^2}\sigma^2_{\hat{\phi}} = \frac{P^2}{4\pi^2 M^2}\frac{1}{\text{SNR}}, \tag{6.22}$$

For a Gaussian input noise the distribution of the instantaneous frequencies at the hopping instant typically looks like the one shown in Fig. 6.4. Let

[2] We assume that $\nu(n)$ is Gaussian with a zero mean.

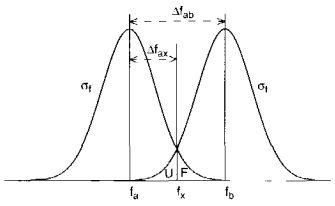


Fig. 6.4: Typical probability distribution function of the frequency estimation error for a frequency hopping signal

$f_x = f_a + \Delta f_{ax}$ be the frequency threshold for determining whether the instantaneous frequency $f(n)$ represents a hop or not. Which means, according to the particular situation of Fig. 6.4, if $f(n)$ exceeds f_x, there is considered to exist a frequency hop. This in turn implies that, even if there is a frequency hop, it cannot be detected when $f(n)$ is in the region marked by U, and we receive a false alarm when $f(n)$ is in the region marked with F. Let the probability of $f(n)$ being in the U and the F regions be $p(U)$ and $p(F)$, respectively, then [39]

$$\begin{aligned} p(U) &= \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\Delta f_{ax} - \Delta f_{ab}}{\sqrt{2}\sigma_f}\right) \\ p(F) &= \frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\Delta f_{ax}}{\sqrt{2}\sigma_f}\right). \end{aligned}$$

For a given hop size Δf_{ab}, we want to find the optimum threshold level (the optimum value of Δf_{ax}) that minimizes the sum of the above two probabilities. Let $p = p(U) + p(F)$, then this means that we want to minimize

$$p = 1 + \frac{1}{2}\text{erf}\left(\frac{\Delta f_{ax} - \Delta f_{ab}}{\sqrt{2}\sigma_f}\right) - \frac{1}{2}\text{erf}\left(\frac{\Delta f_{ax}}{\sqrt{2}\sigma_f}\right). \tag{6.23}$$

By setting the derivative of this sum with respect to Δf_{ax} to zero and after some elaboration, we get

$$\Delta f_{\text{opt}} = \frac{\Delta f_{ab}}{2}.$$

Replacing this into (6.23), the optimum value of p is obtained, denoting this value with p_o, we have

$$p_o = 1 - \text{erf}\left(\frac{\Delta f_{ab}}{2\sqrt{2}\sigma_f}\right) = 1 - \text{erf}\left(\Delta f_{ab}\frac{\pi M\sqrt{\text{SNR}}}{\sqrt{2}P}\right) \tag{6.24}$$

Let p_m represent the maximum allowable error rate, and let's say that we are interested in finding the lower limit on Δf_{ab} that guarantees an error rate of $p_o \leq p_m$. Solving for Δf_{ab} from (6.24), we obtain

$$\Delta f_{ab} \geq \frac{\sqrt{2}P}{\pi M \sqrt{\text{SNR}}} \text{erf}^{-1}(1 - p_m),$$

where $\text{erf}^{-1}(\cdot)$ is the inverse error function. This shows that, for a given M, SNR and p_m, the right hand expression represents the minimum allowable frequency hop size.

6.1.7 Simulation results

The simulation results are based on the setup conditions specified in *Example 6.1*. We take a close look at various sets of behaviors. In the first case, the behavior of the estimation error as a function of SNR is studied. The results are shown in Fig. 6.5 for $N_o = 0.55N$. As seen, the variance of the angle and frequency estimation errors decrease in proportion to the inverse of the SNR. This is in perfect agreement to what one would expect.

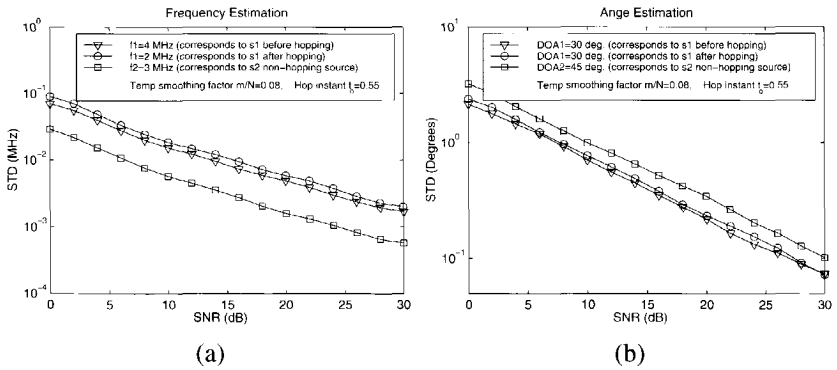
In the second set of plots (Fig. 6.6), the standard deviations of the angle and frequency estimation errors are plotted as functions of η_o. From the plots, it is seen that, while the estimation errors associated with the pre-hopping parameters improved with increasing η_o, those of the post-hopping parameters have degraded. This is in perfect agreement to what was discussed in section 6.1.6, where it has been shown that the variances of the angle and frequency estimation errors are proportional to the square inverse and the cube inverse of the number of temporal samples, respectively. Noting that the number of temporal samples associated with the pre- and post-hopping parameters are η_o and $1 - \eta_o$, respectively, the behaviors shown in Fig. 6.6 are expected.

In the third case (Fig. 6.7), the behaviors of the estimation errors as functions of the temporal smoothing factor m are plotted. It is seen that, as m is increased, the estimation errors initially drop quickly, and then saturate. Since computational complexity of the algorithm increases in proportion with m^2, it makes sense to choose m taking into consideration the cost-accuracy trade-off, preferably near the saddle point.

In the fourth set of plots (Fig. 6.8), the biases introduced by the shift invariance violating columns as functions of m are shown. From Fig. 6.8*a*, it is seen that, with increasing m, while the biases in the frequency estimates of the hopping source gradually decrease, the bias in the frequency estimate of the non-hopping source degrades. This indicates that as m is increased, the

non-frequency hopping source gets more and more contaminated by the shift invariance violation, whereas the hopping source is helped by the averaging. This is partly because the averaging process mixes the rows of the data matrix creating a more even distribution of the shift invariance violation. From Fig. 6.8*b*, one can see that a similar behavior is exhibited by the biases in the DOA estimates. Particularly, it is seen that, while the biases in the DOA of the FH source steadily go down, the bias of that of the non hopping source starts to rise after initial drop. The saddle point corresponds to the point where the shift invariance violating columns counter balance the improvement gained via temporal smoothing. It should, however, be noted that the biases in the DOA estimation are much smaller (one order of magnitude smaller) than the standard deviation, and thus are negligible. Note that the frequency estimates are based on shift structure induced via temporal smoothing, whereas the DOA estimates are obtained by considering shift invariance structure induced by the antenna array. This makes the DOA estimates less affected by the FH, and justifies the behaviors observed in Fig. 6.8.

As a final plot, in Fig. 6.9, the theoretically expected probability of committing error in determining the hoping instant (viz. (6.24)) and the corresponding result obtained via simulation are summarized. It is seen that the two behaviors coincide satisfactorily, verifying the validity of the analysis given in section 6.1.6.



(a) (b)

Fig. 6.5: Frequency and DOA Estimation error as functions of SNR. (a) behavior for frequency, and (b) behavior for DOA

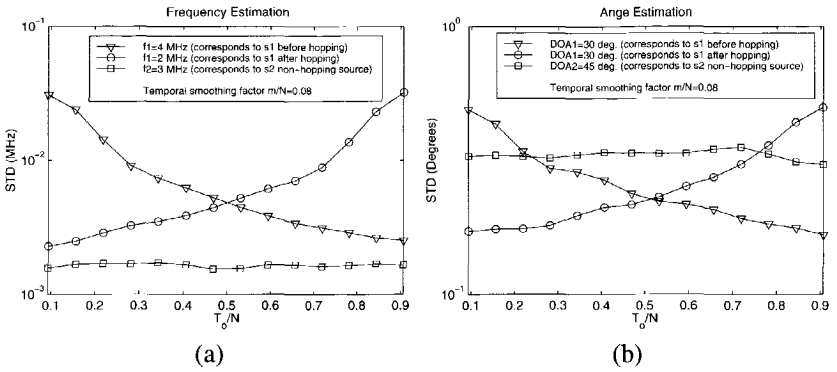


Fig. 6.6: Frequency and DOA Estimation error as functions of $\eta_o = n_o/N$. (a) Behavior for frequency, and (b) behavior for DOA

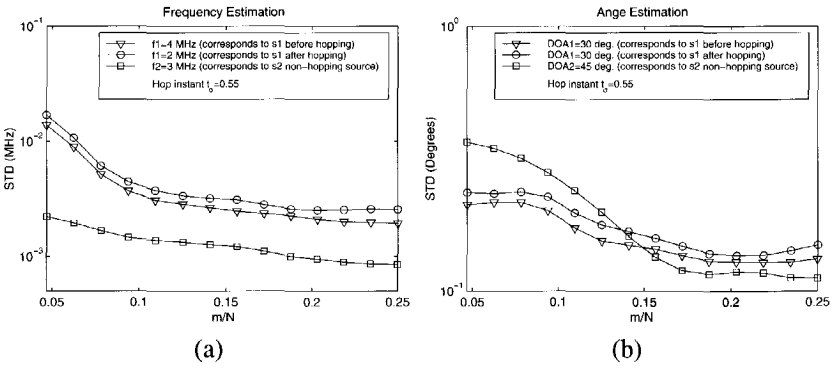
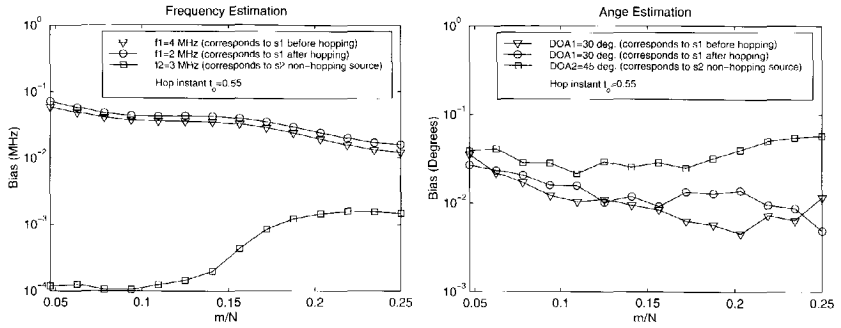
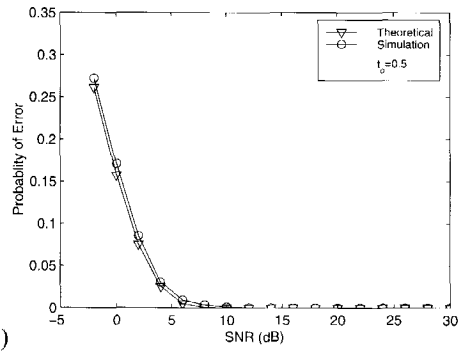


Fig. 6.7: Dependence of the estimation errors on the temporal smoothing factor m. (a) behavior of frequency estimation, and (b) behavior of DOA estimation (SNR=20dB)



(a) (b)

Fig. 6.8: Dependence of the biases on temporal smoothing factor m. (a) behavior of frequency estimation, and (b) behavior of DOA estimation (SNR=20dB)



(b)

Fig. 6.9: Probability of committing error in determining the hoping instant

6.2 Effects of Frequency and Angle Spreads on the JAFE Algorithm

In generating the data model in section 3.1, I have assumed that the signals under consideration satisfy the narrowband assumption described in section 2.1.1, and that they propagate from a point in the far field. In this section, I consider cases where these assumptions are violated. I.e., I consider cases

where frequencies and angles are spread over certain frequency and angle regions, respectively.

6.2.1 Frequency Spreading (Modulation)

Recall, from section 2.1.1, that the narrowband assumption states that

1. the signal bandwidth W is much smaller than the sampling rate P, and

2. the sampling rate P is much smaller than the nominal carrier frequency f_c of the band of interest.

As the signal bandwidth increases the first of the above two narrowband assumptions becomes less and less valid. At a certain point, the modeling error becomes so large that the algorithm breaks down. In this section, after discussing a simple modulation model, I will derive condition under which the narrowband assumption remains valid.

Works related to the analysis presented here are found in the literature. Particularly, Zatman [101] gives a similar analysis assuming that both of the two narrowband assumptions mentioned above are violated.[3] The consequence of this assumption is that the DOAs of the signals now become functions of instantaneous frequency. Zatman was able to come up with an analytic expression for the DOA spread by making the assumption that the signal has a rectangular spectrum in the band of interest. His main goal was not bandwidth estimation. In fact, he considered a single source example to show that modulation introduces biases in the DOA estimates. His derivation fails to give meaningful results when more than a single source is involved. In my case, however, I show that the JAFE algorithm can be used to estimate the bandwidths of multiple signals.

Moreover, unlike Zatman, I assume that only the first of the two narrowband assumptions is violated. I.e., I consider the situation where P is much smaller than the nominal carrier frequency f_c, and the signal bandwidth W is comparable to P. In this case, the array steering matrices are essentially frequency independent. Note that this is a justifiable assumption in radio communication systems, where we have a number of narrowband signals spread over a certain relatively wide frequency range.

[3] For acoustic and sonar communication, this is a valid assumption.

The model

Consider, the data model given in (3.2). It has been shown that, when the assumption $P \ll f_c$ is valid, this data matrix has the factorization given in (3.3). For modulated signals the terms in this expression must be given as functions of frequency, in which case, under noise free conditions we write [4]

$$\boldsymbol{X}_m(f) = \boldsymbol{A}_m(f)\boldsymbol{F}_s(f) \in \mathbb{C}^{mM,N-m+1}, \tag{6.25}$$

and the data covariance matrix as a function of the instantaneous frequency is then given by

$$\boldsymbol{R}(f) = \boldsymbol{A}_m(f)\boldsymbol{P}_s(f)\boldsymbol{A}_m^H(f) \in \mathbb{C}^{mM,mM},$$

where $\boldsymbol{P}_s(f) = \boldsymbol{F}_s(f)\boldsymbol{F}_s(f)^H$ is the instantaneous signal covariance matrix. If the incoming signals are assumed to be independent, $\boldsymbol{P}_s(f)$ is a diagonal matrix, where the diagonal entries $p_i(f)$ represent the received signal powers. Let $\boldsymbol{a}_{mi}(f)$ be the i-th column of $\boldsymbol{A}_m(f)$, then $\boldsymbol{R}(f)$ can be decomposed as

$$\boldsymbol{R}(f) = \sum_{i=1}^{d} p_i(f)\boldsymbol{a}_{mi}(f)\boldsymbol{a}_{mi}^H(f) =: \sum_{i=1}^{d} \boldsymbol{R}_i(f) \in \mathbb{C}^{mM,mM},$$

where $\boldsymbol{R}_i(f)$ represents the covariance matrix for the i-th signal at the instantaneous frequency f. For non zero bandwidth, the i-th signal covariance matrix $\boldsymbol{R}_i$ may be constructed by integrating $\boldsymbol{R}_i(f)$ over the desired frequency band, i.e.,

$$\boldsymbol{R}_i = \int_{f_o-\frac{W}{2}}^{f_o+\frac{W}{2}} p_i(f)\boldsymbol{a}_{mi}(f)\boldsymbol{a}_{mi}(f)^H df \in \mathbb{C}^{mM.mM},$$

where W is the bandwidth of the signal and f_o is its center frequency. Let T be the sampling period and $\phi(f) := e^{j2\pi fT}$, then from (3.2), we note that the i-th array steering vector $\boldsymbol{a}_{mi}(f)$ has the structure,

$$\boldsymbol{a}_{mi}(f) = \begin{bmatrix} \boldsymbol{a}_i \\ \boldsymbol{a}_i\phi(f) \\ \vdots \\ \boldsymbol{a}_i\phi(f)^{m-1} \end{bmatrix} \tag{6.26}$$

where the M-vector $\boldsymbol{a}_i$ is the i-th column of $\boldsymbol{A}$ (the array steering matrix prior to temporal smoothing). Since we have assumed that the band of interest is

[4] This notation should not be confused with the Fourier transform of $\boldsymbol{X}_m$. Here, $\boldsymbol{X}_m(f)$ represents the data that would have been collected if the signal had been filtered with an ideal bandpass filter with infinitively small bandwidth centered at the instantaneous frequency f.

much smaller than the nominal carrier frequency, $\boldsymbol{a}_i$ can be assumed to be frequency independent. Let $\boldsymbol{Q}_i = \boldsymbol{a}_i \boldsymbol{a}_i^H$, then it follows that the jk-th $M \times M$ block entry of $\boldsymbol{R}_i \in \mathbb{C}^{mM,mM}$ is

$$\boldsymbol{Q}_i^{(jk)} = \boldsymbol{Q}_i \int_{f_o - \frac{W}{2}}^{f_o + \frac{W}{2}} p_i(f) \phi_i(f)^{j-k} df.$$

Replacing $\phi(f) = e^{j2\pi fT}$ and letting $\delta_{jk} = j - k$, the above can be expressed as

$$\boldsymbol{Q}_i^{(jk)} = \boldsymbol{Q}_i \int_{f_o - \frac{W}{2}}^{f_o + \frac{W}{2}} p_i(f) e^{j2\pi \delta_{jk} fT} df. \tag{6.27}$$

If we now assume that $p_i(f)$ has a rectangular spectrum between $f_o - \frac{W}{2}$ and $f + \frac{W}{2}$, the solution to the above integral is

$$\boldsymbol{Q}_i^{(jk)} = \boldsymbol{Q}_i \text{sinc}(\delta_{jk} WT) e^{j2\pi \delta_{jk} f_o T}.$$

Note that, except for the term $\text{sinc}(\delta_{jk} WT)$, the above expression exactly looks like the model for a single frequency signal. The worst deviation from the ideal case occurs when $\delta_{jk} = \pm(m-1)$, where m is the temporal smoothing factor. Thus, adopting the same condition used by Compton [16] from spatial to temporal context, we regard the incoming signal narrowband when

$$\text{sinc}((m-1)WT) \approx 1. \tag{6.28}$$

In a noisy situation, the condition in (6.28) could be too pessimist. In [101], a tighter condition, which takes into account the effect of noise, has been arrived at by tackling the problem from a different angle. There, it is argued that a signal can no longer be assumed to be narrowband if the bandwidth of the signal is such that its rank one representation is not possible. For a single source case, for instance, the signal is assumed to be narrowband as long as the second largest eigenvalue of the covariance matrix remains at the noise floor. Let λ_2 be the second largest eigenvalue of the covariance matrix and σ^2 be the noise power, then the condition for narrwband is given by

$$\lambda_2 \leq \sigma^2$$

In [34, 101] an explicit expression for λ_2 as a function of signal bandwidth has been derived. The results are obtained by modeling the single wideband signal with two independent zero bandwidth signals. Let f_1 and f_2 be the frequencies of these signals and ρ_1 and ρ_2 be their powers, respectively. The covariance matrix $\boldsymbol{R}$ is thus modeled by

$$\boldsymbol{R} = \rho_1 \boldsymbol{a}_{m1}(f_1) \boldsymbol{a}_{m1}^H(f_1) + \rho_2 \boldsymbol{a}_{m2}(f_2) \boldsymbol{a}_{m2}^H(f_2). \tag{6.29}$$

Note that $\rho_1 + \rho_2 = p$, where p is the power of the original wideband signal. The eigenproblem is to find values of λ_i and $\boldsymbol{u}_i$, $i = 1, 2$ such that

$$\boldsymbol{R}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i, \tag{6.30}$$

where $\boldsymbol{u}_i$ is given by a linear combination of $\boldsymbol{a}_{m1}(f_1)$ and $\boldsymbol{a}_{m2}(f_2)$, i.e., for two complex constants $\mu_{i,1}$ and $\mu_{i,2}$,

$$\boldsymbol{u}_i = \mu_{i,1}\boldsymbol{a}_{m1}(f_1) + \mu_{i,2}\boldsymbol{a}_{m2}(f_2)$$

Let us assume that the two signals in (6.29) have equal power, i.e., $\rho_1 = \rho_2 = \frac{p}{2}$. Then substituting the above into (6.30) and using (6.29), we obtain the following solution for $\lambda_{1,2}$ (viz [34, 101])

$$\lambda_{1,2} = mM\frac{p}{2}(1 \pm |\psi|), \tag{6.31}$$

where, $|\psi|$ is the cosine of the angle between $\boldsymbol{a}_{m1}(f_1)$ and $\boldsymbol{a}_{m2}(f_2)$ in an mM-dimensional vector space, and is given by

$$|\psi| = \frac{|\boldsymbol{a}_{m1}^H(f_1)\boldsymbol{a}_{m2}(f_2)|}{\sqrt{\boldsymbol{a}_{m1}^H(f_1)\boldsymbol{a}_{m1}(f_1)\boldsymbol{a}_{m2}^H(f_2)\boldsymbol{a}_{m2}(f_2)}}$$

With the model of $\boldsymbol{a}_{mi}(f)$ in (6.26), and w_f defined as $w_f = f_2 - f_1$, we get

$$|\psi| = \frac{1}{m}\left|\sum_{i=0}^{m-1} e^{j2\pi w_f T}\right|$$

For $w_f T \ll 1$ the above is approximately given by

$$|\psi| \approx \text{sinc}(mw_f T) \tag{6.32}$$

For the narrowband assumption to be valid, we require that the smallest of the two eigenvalues given in (6.31) be under the noise level,

$$mM\frac{p}{2}(1 - |\psi|) \leq \sigma^2 \Rightarrow \text{sinc}(mw_f T) \geq 1 - \frac{2\sigma^2}{mMp}. \tag{6.33}$$

The similarity between (6.33) and (6.28) is obvious, i.e., when the noise approaches zero, (6.33) reduces to (6.28). Zatman [101] did not, in particular, recognize this similarity. Solving for σ^2 from (6.33) we get an equivalent noise level that has the same effect as w_f,

$$\sigma_{w_f}^2 = \frac{mMp}{2}(1 - \text{sinc}(mw_f T)). \tag{6.34}$$

It is seen that, when $w_f = 0$, $\sigma_{w_f}^2 = 0$ and we get a rank one representation of the incoming signal as expected. These results are demonstrated using simulations in section 6.2.3.

Estimating Signal Bandwidth

For narrowband signals, the rank of the data covariance matrix is equal to the number of signals present. I.e., there is a rank one representation of each signal. For sufficiently wideband signals, however, the effective rank of the covariance matrix is larger than the number of signals, and the zero bandwidth signal model does no longer apply. In fact, as the bandwidth is increased, eigenvalues pop up from the noise floor one by one. In this case we get a model where the incoming signal is represented with multiple, independent single-frequency-signals. As the signals originate from the same source, they must have the same angular bearing. Let q be the number of these single-frequency-signals, and for $j = 1, \cdots, q$, let f_j be the corresponding frequency estimates. The center frequency of the incoming wide band signal is computed as the mean of the f_j,

$$\hat{f}_o = \frac{\sum_{j=1}^{q} f_j}{q} \tag{6.35}$$

and the band widths of the signals are estimated by requiring that the variances of the model and reality be the same, i.e.,

$$\frac{1}{W}\int_{f_o-\frac{W}{2}}^{f_o+\frac{W}{2}} f^2 p_i(f) df = \frac{\sum_{j=1}^{q}(\hat{f}_o - f_j)^2}{q}. \tag{6.36}$$

If we assume that the signal has a rectangular spectrum, (i.e., $p_i(f) = 1$) in the interval $[f_o - W/2, f_o + W/2]$, then solving for W, we get

$$W = \sqrt{12(\frac{\sum_{j=1}^{q}(\hat{f}_o - f_j)^2}{q})}, \tag{6.37}$$

The validity of (6.36) is verified using simulation results in section 6.2.3.

6.2.2 Angle Spreading

Angle spreading occurs when there are local scatterers near the signal source or when the propagating wavefront bounces off an extended reflector as shown in Fig. 6.10. Let ϖ be a positive real constant such that the DOA α of the incoming signal is spread over the interval $[\alpha_o - \frac{\varpi}{2}, \alpha_o + \frac{\varpi}{2}]$. Then the m-th entry of the array steering vector is obtained by taking the integral over this interval

$$a_m = \frac{1}{\varpi}\int_{\alpha_o-\frac{\varpi}{2}}^{\alpha_o+\frac{\varpi}{2}} b(\alpha) e^{j2\pi m\Delta \sin\alpha} d\alpha, \tag{6.38}$$

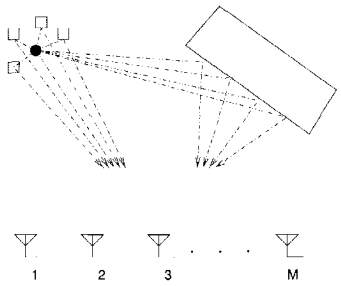


Fig. 6.10: Angle spreading due to local scatterers and an extended reflector

where $b(\alpha)$ represents the amplitude of the ray from the instantaneous angle α. Let $\mu = \alpha - \alpha_o$, $\theta = e^{j2\pi\Delta \sin\alpha_o}$. Then, taking the approximation

$$\sin\alpha = \sin(\alpha_o + \mu) \approx \sin\alpha_o + \mu\cos\alpha_o$$

and making change of variables and letting $b'(\mu) = b(\alpha_o + \mu)$, the above integral can be re-written as

$$a_m = \frac{1}{\varpi}\int_{-\frac{\varpi}{2}}^{\frac{\varpi}{2}} b'(\mu)e^{j2\pi m\Delta(\sin\alpha_o + \mu\cos\alpha_o)}d\mu,$$

If we assume that the DOA is uniformly distributed over the angle range (i.e., $b'(\mu) = 1$), the solution to the above integral becomes

$$a_m = \text{sinc}(m\varpi\Delta\cos\alpha_o)\theta^m \tag{6.39}$$

It is seen that except for the $\text{sinc}(m\varpi\Delta\cos\alpha_o)$ term the above equation looks like the model for the case of a point source model. Thus, adopting the results from the previous section, we consider the angle spread to be narrow if

$$\text{sinc}((M-1)\varpi\Delta\cos\alpha_o) \approx 1$$

When this condition is satisfied, there is essentially no decorrelation between the signals received at the opposite ends of the antenna array. For $\Delta = 1/2$, if we assume $\text{sinc}(x) \approx 1$, for $|x| < 0.01$, then the above condition is satisfied for

$$\varpi \leq \frac{0.02}{(M-1)\cos\alpha_o}$$

The dependence on $\cos\alpha_o$ is justified since the time delay between the opposite ends of the antenna array is proportional to the cosine of the angle of arrival. For instance, for $\alpha_o = 45$ degrees, the maximum allowable angular spread is $\varpi_{max} = 0.4$ degrees.

6.2.3 Simulations

In this simulation example, I consider two equal energy amplitude modulated signals with angle-frequency pairs $(\alpha_1, f_1) = (15, 10.2)$ and $(\alpha_2, f_2) = (30, 10.1)$, where the angle is in degrees and frequency in MHz. The band of interest is in the range 9.5 MHz to 10.5 MHz. After IF demodulation (with a local oscillator operating at 10 MHz), the IF signal is sampled at a sampling frequency of 1 MHz. The data is collected by an M element ULA into $mM \times N$ data matrix, where $M = 4$, $N = 128$ and the temporal smoothing factor m is set to 4. No spatial smoothing is performed.

In Fig. 6.11, the data covariance matrix eigenvalues are plotted as functions of signal bandwidth. For this simulation, the amplitude of the second source is set to zero, effectively resulting in a single source scenario. As seen from the graph, the theoretically predicted behavior of the two largest eigenvalues perfectly agrees with the simulation result. The second largest eigenvalue emerges out of the noise floor at the point where (6.34) is satisfied. This point is marked by a dotted vertical line in Fig. 6.11.

In the second instance, two plots showing the behaviors of estimation error as functions of bandwidth and SNR, respectively are shown. In Fig. 6.12a, the signal bandwidth is varied from 50 kHz to 100 kHz, and the frequency estimation error for three SNR values are plotted. In Fig. 6.12b, the behavior of the estimation error is shown from a different angle. Here, taking the bandwidth as a parameter, the estimation error is plotted as a function of SNR. From the graphs, one can clearly see that, for a given SNR, there exists a point beyond which the error due to modulation dominates. At this transition point, the error caused by the signal modulation and the error generated by noise are equal.

In Fig. 6.13a, the biases in the frequency estimates caused by model mismatch are shown. It is seen that for $W = 100$ kHz the biases in the frequency estimates come close to 4% of the bandwidth. For some applications, this bias might be too much. In such cases, a more accurate signal model, where a signal is represented with multiple independent sources as discussed in section 6.2.1, can be employed. Let q be the number of these independent signals and, for $j = 1, \cdots q$, let f_j be their estimated center frequencies, then the frequency estimate is obtained using (6.35). In Fig. 6.13b, the biases of the frequency estimation based on this relation are shown. It is seen that, as compared to the former approach, the biases are now reduced by an order of magnitude.

In Fig. 6.14, results of signal bandwidth estimation based on the model discussed in section 6.2.1 are shown. It is seen that the results obtained via simulation agree with the theoretically predicted behavior. The offset of about 5 kHz, is caused by the fact that in the derivation of the theoretical result we

have assumed brick type passband behavior, which cannot be satisfied with the simulation data.

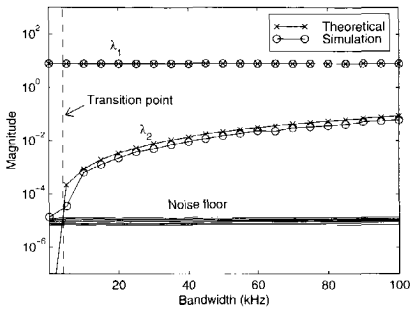
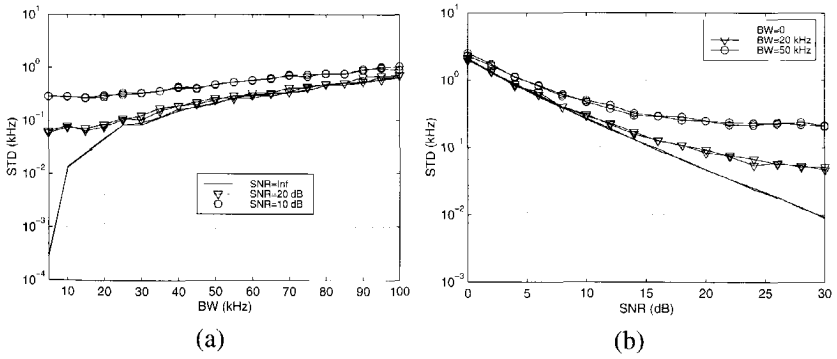


Fig. 6.11: Behavior of the two largest eigenvalues as a function of signal bandwidth (SNR=50 dB)



(a) (b)

Fig. 6.12: Behavior of the estimation errors of the center frequencies of two signals with IF of $f_1 = 200$ kHz and $f_2 = 200$, a) as a function of bandwidth for SNR=10 dB, 20 dB and ∞, and b) as a function of SNR for bandwidth=0, 20 and 50 kHz

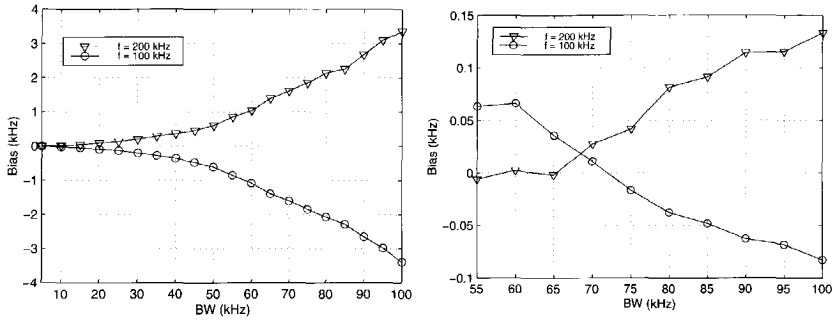


Fig. 6.13: Dependence of the biases on the signal bandwidth, (a) behavior assuming rank one representation of signal, and (b) behavior assuming rank two representation of signal, (SNR=20dB)

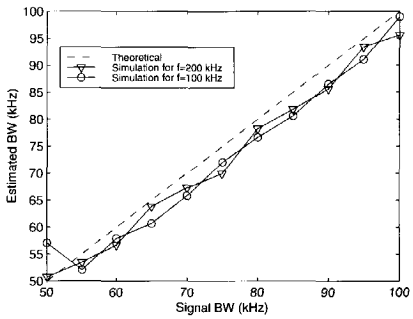


Fig. 6.14: Results of bandwidth estimation (SNR=20 dB)

6.3 Moving Sources and Moving Receivers

In this section I consider the cases where either the source or the receiver or both are moving. I derive a mathematical model for each of these cases, and see how it affects the performance of the estimation algorithm.

Let a signal source $s(t) = e^{j2\pi ft}$ be moving with a velocity v_s in the angular

direction of θ_s to the source-receiver-line[5] as shown in Fig. 6.15. Let τ be

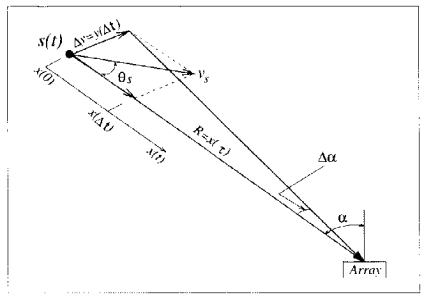


Fig. 6.15: Modeling a moving source

the time it takes the signal to travel from position $x(0)$ to the receiver such that $R = x(\tau)$. At time t the transmitter-receiver separation is given by $x(\tau) - x(t)$. Thus, the signal received at the antenna array can be expressed as

$$s_r(t) = s(t - \{\frac{x(\tau) - x(t)}{c}\}) \tag{6.40}$$

Putting $x(t) = v_s t \cos(\theta_s)$ and $c = f\lambda$, we obtain,

$$s_r(t) = s(t\{1 + \frac{v_s}{f\lambda}\cos(\theta_s)\} - \tau) = e^{j2\pi((f + \frac{v_s}{\lambda}\cos(\theta_s))t - f\tau)}. \tag{6.41}$$

Let $f_m s = \frac{v_s}{\lambda}$, then at the receiver, the source frequency appears to be

$$f' = f + f_{ms}\cos(\theta_s). \tag{6.42}$$

$f_D = f_{ms}\cos(\theta_s)$ is called the *Doppler shift* and f_{ms} is its maximum value. If the speed of the signal source is known, the Doppler shift can be compensated for. In many practical applications, however, the speed is not only unknown but also variable. In such cases, it is important to either estimate the speed of the signal source or mathematically model its effect. In general, the Doppler shift becomes apparent only if the shift is greater than the resolution of the estimator. For instance, if the mean standard deviation of the estimator is-say-100 Hz, then Doppler shifts with $f_D \leq 100$ Hz would not be detected by the estimator[6].

To see how the DOA is affected by the movement of the source, we consider the motion in the y-direction as shown in Fig. 6.15. Let Δt be the time window over which the data is collected. Then between the initial and the final samples

[5]Here we say the pair (v_s, θ_s) describe the movement of the source.

[6]At a carrier frequency of 1 GHz this corresponds to a speed of $v \approx 100$ Km/h).

the source has moved by $\Delta y = v_s \Delta t \sin(\theta_s)$ perpendicular to the transmitter receiver line. Assuming $\Delta\alpha$ is small, we have $\Delta y = R\Delta\alpha$. Thus,

$$\Delta\alpha = \frac{v_s \Delta t}{R} \sin(\theta_s). \tag{6.43}$$

For instance, $v = 100$ km/h, $R = 50$m and $\Delta t = 1$msec., give $\Delta\alpha = 0.035$ degrees. This is a very small value and can be neglected.

Next, consider the case where the receiver is moving, and let its movement be represented by the pair (v_r, θ_r) as shown in Fig. 6.16. Following a simi-

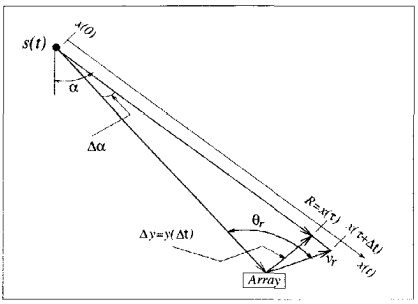


Fig. 6.16: Modeling a moving receiver

lar procedure as above, it is easy to show that a moving receiver affects the frequency and the DOA in the same manner as a moving source. I.e., if the receiver moves with a velocity v_r in the direction θ_r as shown in Fig. 6.16, the source frequency appears to be

$$f' = f + f_{mr} \cos(\theta_r)$$

where $f_{mr} = \frac{v_r}{\lambda}$. If Δt represents the time window over which the data is collected, the change in angle over this interval is given as before by

$$\Delta\alpha = \frac{v_r \Delta t}{R} \sin(\theta_r).$$

When both the source and the receiver are moving, one can easily show (with superposition) that the frequency and DOA are affected according to the relations

$$f' = f \pm f_{ms} \cos(\theta_s) \pm f_{mr} \cos(\theta_r) \tag{6.44}$$

and

$$\Delta\alpha = \left| \frac{v_s \Delta t}{R} \sin(\theta_s) \pm \frac{v_r \Delta t}{R} \sin(\theta_r) \right|, \tag{6.45}$$

respectively. The frequency components are positive when the x-components of v_s and v_r point towards the receiver and the source, respectively, and the angle components are added when the velocities have opposite y-components (see Fig. 6.15 and Fig. 6.16).

6.4 Effect of Non-ideal Array Behavior

Many computer simulation results confirm the superior performance of high resolution DOA estimation algorithms such as ESPRIT and MUSIC. In actual arrays, distortions caused by non ideal behaviors (antenna gain and phase, and mutual coupling errors) degrade the expected performances of these methods drastically. Usually, these problems are deliberately or unknowingly ignored by many theoreticians. In this section, I present experimental results showing that, if these methods have to achieve their theoretical performance, one has to take into consideration the non-ideal array behaviors. Consider an arbitrary geometry antenna array with M elements being impinged by a single far field signal with the DOA of α. Under non-ideal conditions, the signal at the output of the antenna can be modelled as [75]

$$\boldsymbol{x}(t) = \boldsymbol{C}\boldsymbol{a}(\alpha)s(t) + \boldsymbol{\nu}(t),$$

where $\boldsymbol{x}(t)$ is the measured data vector, $\boldsymbol{a}(\alpha)$ is the ideal array steering vector for a narrowband signal from direction α, $s(t)$ is the instantaneous signal value, $\boldsymbol{\nu}(t)$ is a noise vector, and the complex matrix $\boldsymbol{C}$ is a *distortion matrix* that accounts for the combined effects of the antenna gain and phase, and coupling errors. In an actual system, we measure $\boldsymbol{a}_c(\alpha) = \boldsymbol{C}\boldsymbol{a}(\alpha)$, and not $\boldsymbol{a}(\alpha)$. This means that for correct parameter estimation, we need to estimate the distortion matrix $\boldsymbol{C}$.

In the literature, there are a number of works that address this issue. Generally, these methods may be categorized into two classes: blind methods and deterministic methods. The blind approaches, [19, 59, 75, 95], are on-line array calibration techniques. They try to estimate both the distortion matrix and the unknown DOAs by making use of some presumed structures. The methods are very appealing, however, they suffer from inconsistency. In fact, it is shown in [37] that these methods give non-unique solutions, when the distortion matrix is allowed to have arbitrary structure. The deterministic methods are off-line array calibration techniques, [37, 76]. They compare the estimated array steering vector $\boldsymbol{a}_c(\alpha)$ against the ideal steering vector $\boldsymbol{a}(\alpha)$ to estimate the distortion matrix. The problem of these approaches is that they fail to model the dependence of $\boldsymbol{C}$ on the environment. [7]

[7] Scaterrers local to the receiver alter the coupling pattern of an array.

A probably better result may be obtained by combining the above two approaches. That is, to use a deterministic method to get a global approximation of the distortion matrix, and a blind approach to refine this estimate. The goal in this section is not to develop methods for obtaining better distortion estimates, it is rather to give experimental results that demonstrate distortion calibration is an essential element in high-resolution DOA estimation methods.

6.4.1 Model

In this section, I formally develop the data model for non-ideal array. To be specific, consider a single narrowband signal with DOA of α impinging on a ULA, with M elements. Let the N-vector

$$\boldsymbol{x}_m^H(t) = [x_m(t) \;\; x_m(t+1) \;\; \cdots \;\; x_m(t+N-1)]$$

be the t-th sampled signal vector measured at the output of the m-th antenna element. If we assume that there is no mutual coupling between the antennas, $\boldsymbol{x}_m(t)$ is given as

$$\boldsymbol{x}_m^H(t) = b_m \boldsymbol{s}^H(t - \tau_m - \mu_m), \tag{6.46}$$

where, $\boldsymbol{s}^H(t) = [s(t) \;\; s(t+1) \;\; \cdots \;\; s(t+N-1)]$ is the sampled input signal, b_m is antenna radiation gain, μ_m is the phase distortion in the m-th channel and τ_m is the propagation delay measured from a reference antenna position. For ULA, we have

$$\tau_m = (m-1)\tau_o \sin(\alpha),$$

where τ_o is the time it takes for the signal to propagate between two adjacent antenna elements and α is the DOA measured with reference to the normal of the array axis. Assuming that $s(t)$ is a narrowband signal with a center frequency f_o, (6.46) can be approximated as

$$\begin{aligned} \boldsymbol{x}_m^H(t) &= b_m e^{-j2\pi f_o \mu_m} e^{-j2\pi(m-1) f_o \tau_o} \boldsymbol{s}^H(t) \\ &=: \gamma_m \theta^{m-1} \boldsymbol{s}^H(t), \end{aligned}$$

where $\gamma_m = b_m e^{-j2\pi f_o \mu_m}$ and $\theta = e^{-j2\pi f_o \tau_o}$. Now collecting the M antenna outputs into the $M \times N$ matrix $\boldsymbol{X}(t)$, we get the following model

$$\boldsymbol{X}(t) = \boldsymbol{\Gamma} \boldsymbol{a}(\theta) \boldsymbol{s}^H(t), \tag{6.47}$$

where the complex diagonal matrix $\boldsymbol{\Gamma} = \text{diag}\{\gamma_i\}_{i=1}^M$ represents the channel phase distortion and $\boldsymbol{a}(\theta) = [1 \;\; \theta \;\; \cdots \;\; \theta^{M-1}]^T$ is the ideal parameterized array steering vector. In the above model, we have assumed that each antenna element acts independently. In actual case, however, the reflected radiation

from one element couples to its neighbors, as do currents that propagate along the surface of the array. Under this condition, the output of each antenna is the sum of the primary incident signal and the secondary reflected signals from the neighboring elements. Thus, for some $M \times M$ matrix $\boldsymbol{D}$, we have

$$\boldsymbol{X}(t) = \boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{a}(\theta)s^H(t) =: \boldsymbol{C}\boldsymbol{a}(\theta)s^H(t) \tag{6.48}$$

6.4.2 Estimating The Distortion Matrix

In this section, I give a simple off-line method of estimating the distortion matrix $\boldsymbol{C}$. The method is similar to those discussed in [37] and [76]. The distinction is that, in my approach after estimating the array steering vectors, I normalize each steering vector with respect to its first entry. This gives us a unique solution for the distortion matrix. Note that the solutions described in [37] and [76] are unique up to some complex multiplicative constant which poses problems when working with several independent snapshots. In the following, we assume that the antennas have flat frequency response at the frequency band of interest. Further, we assume that there is only a single source in the channel, though generalization for more sources is possible. We collect the data for Q distinct source positions. Let the data associated with the q-th source position, $1 \leq q \leq Q$, be denoted by $\boldsymbol{X}_q$. Then the distortion matrix $\boldsymbol{C}$ is estimated as follows:

For $q = 1, \cdots, Q$ **Do**

- Collect an $M \times N$ data matrix $\boldsymbol{X}_q$
- Compute the SVD of $\boldsymbol{X}_q$: $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H = \boldsymbol{X}_q$
- Estimate the signal subspace $\boldsymbol{u}_q$ as the column of $\boldsymbol{U}$ that corresponds to the largest singular value.
- Normalize $\boldsymbol{u}_q$ with respect to its first entry. If $\boldsymbol{a}_q$ denotes the q-th array steering vector to be estimated and $\boldsymbol{a}_q(m)$ its m-th entry, then set

$$\boldsymbol{a}_q(m) = \boldsymbol{u}_q(m)/\boldsymbol{u}_q(1), m = 1, \cdots, M.$$

- Construct augmented array steering matrix $\boldsymbol{A}_c$ as

$$\boldsymbol{A}_c = [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_Q]$$

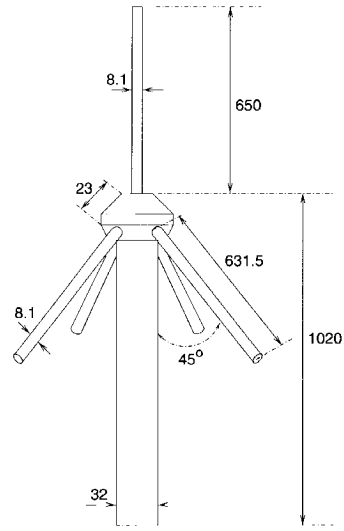


Fig. 6.17: A monopole antenna element. The dimensions are in mm.

End Do
Finally, Compute the distortion matrix as

$$\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{A}_c^H(\boldsymbol{A}_c\boldsymbol{A}_c^H)^{-1}, \tag{6.49}$$

where $\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1) \ \cdots \ \boldsymbol{a}(\theta_Q)]$ is a matrix containing the Q ideal array steering vectors corresponding to the Q source positions. Note that, for this method to work Q must be greater than or equal to the number of antenna elements M.

6.4.3 Simulation Results

In this section, the antenna coupling behavior is analyzed using NEC antenna simulation setup. With this simulation, I demonstrate that the antenna coupling model derived in this chapter is useful and appropriate. A narrowband far field signal at center frequency of 109 MHz is impinging on a ULA with $M = 4$ elements. Each antenna element is a mono-pole with the geometry shown in Fig. 6.17 and the baseline separation is 109 cm. A simulation data was collected for DOAs varying from 0 to 90 degrees with a step of 5 degrees. The distortion matrix was estimated with the procedure described in section 6.4.2 using the measurements corresponding to the DOA $= [0 \ \ 15 \ \ 55 \ \ 90]$ degrees. The ESPRIT and MUSIC algorithms were applied to estimate the

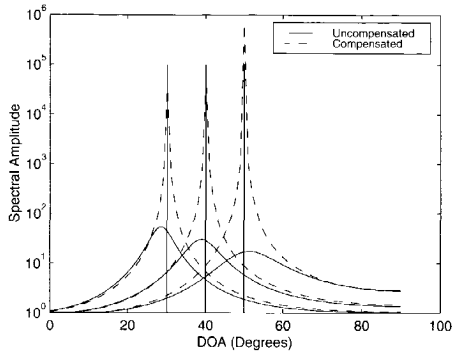
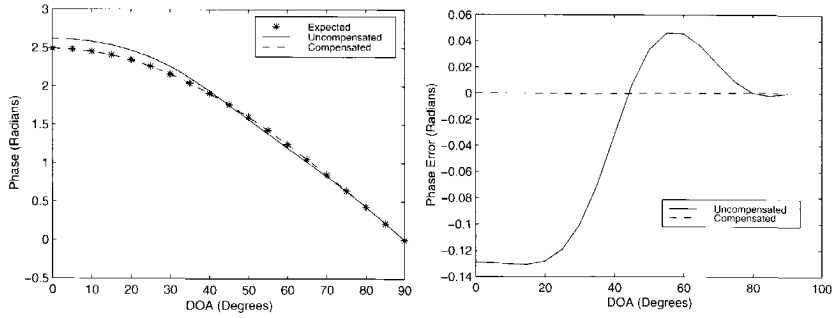


Fig. 6.18: The music Spectra, for a single source scenario. The true DOAs are indicated by solid vertical lines, and the plots are shown for three different DOA values.

DOAs for each simulation data. The performance improvements attained with the distortion compensation are summarized in figures 6.18 and 6.19 for MUSIC and ESPRIT approaches, respectively. Note that in the MUSIC approach, the DOAs are obtained by searching spectral peaks in the so-called "MUSIC Spectrum", whereas in the ESPRIT algorithm, the DOAs are computed from phase estimates obtained via shift invariance considerations. As a final result,



(a) (b)

Fig. 6.19: a) Phase behaviors as functions of DOA, and b) Errors in the phase estimates before compensation (dashed line) and after compensation (solid line)

in table 6.1, the expected DOAs and the estimated DOAs (before and after

distortion compensation) are summarized.

From the results, it is seen that the calibration procedure almost completely eliminates the biases. The slight error is partly because of noise and partly because of the fact that in estimating the coupling matrix, the global optimum was not attained.

Table 6.1: Effect of distortion compensation on DOA estimates (single source)

Expected	Uncompensated		Compensated	
	ESPRIT	MUSIC	ESPRIT	MUSIC
0	0	0	0	0
5.00	5.05	4.95	4.99	4.97
10.00	10.00	9.94	9.99	9.98
15.00	14.81	14.94	15.00	15.00
20.00	19.44	19.19	20.00	20.02
25.00	23.99	23.94	24.99	25.02
30.00	28.64	28.94	29.98	30.02
35.00	33.62	33.94	34.97	35.01
40.00	39.18	38.94	39.97	40.00
45.00	45.44	44.94	44.98	45.00
50.00	52.28	51.40	49.99	50.00
55.00	59.52	58.35	55.00	55.00
60.00	67.10	65.94	60.01	60.00
65.00	75.58	73.39	65.01	65.00
70.00	–	84.26	70.01	70.00
75.00	–	90.00	75.01	75.00
80.00	–	90.00	80.01	80.00
85.00	–	90.00	85.00	85.00
90.00	–	90.00	90.00	90.00

To see whether the compensation helps when more than one sources are involved, the estimation process was repeated for a two source case. Two narrowband signals with IF center frequencies 4 and 5 MHz and DOAs 75 and 60 degrees, respectively, were considered. The MUSIC and ESPRIT results are summarized in table 6.2. It is seen that the compensation almost completely resolves the coupling problem. Particularly, it is interesting to see that, before the compensation, MUSIC gave only a single DOA estimate, that was completely wild. This is clearly seen in the MUSIC spectra shown in Fig. 6.20.

Generally, the radiation patterns of the antenna elements and, consequently,

Table 6.2: Effect of distortion compensation on DOA estimates (two sources, $f_1 = 109$ MHz, $f_2 = 110$ MHz. The coupling matrix is computed at 109 MHz.)

Expected	Uncompensated		Compensated	
	ESPRIT	MUSIC	ESPRIT	MUSIC
(15,30)	(14.9,28.7)	(27.2,–)	(15.0,30.2)	(15.0,30.4)
(35,50)	(33.9,52.4)	(39.1,–)	(35.0,50.1)	(34.9,49.3)
(55,70)	(59.6,–)	(54.4,–)	(55.0,70.7)	(55.0,66.0)
(75,90)	(–,–)	(67.5,–)	(75.0, –)	(74.9, –)

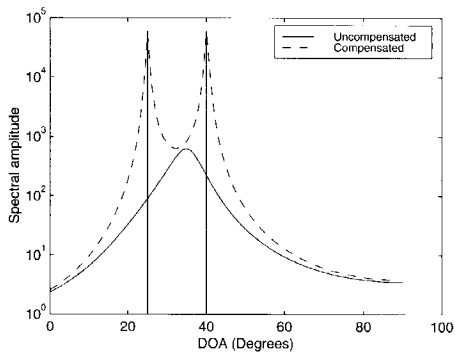


Fig. 6.20: The music Spectra, for two source case. The true DOAs are indicated by solid vertical lines.

the coupling pattern of the antenna array are functions of frequency. Since the coupling matrix is computed at 109 MHz, this is why it is seen that the biases at 110 MHz are much higher than those at 109 MHz. For a good result it is useful to select antenna elements with invariant radiation pattern in the band of interest. If this cannot be achieved, the compensation process should be repeated for each signal frequency. As the frequency estimates are byproducts of the JAFE algorithm, this can easily be achieved in the context of this thesis. However, the computational complexity will now be a factor d higher, where d is the number of carrier frequencies. When compensation is conducted at both 109 and 110 MHz frequencies, the estimated DOAs of the two sources are improved as shown in table 6.3.

Table 6.3: Effect of distortion compensation on DOA estimates (two sources, $f_1 = 109$ MHz, $f_2 = 110$ MHz. The Coupling matrix is computed for each frequency separately)

Expected	Uncompensated		Compensated	
	ESPRIT	MUSIC	ESPRIT	MUSIC
(15,30)	(14.9,28.7)	(27.2,–)	(15.0,30.0)	(15.0,30.0)
(35,50)	(33.9,52.4)	(39.1,–)	(35.0,50.0)	(34.9,50.0)
(55,70)	(59.6,–)	(54.4,–)	(55.0,70.0)	(55.0,69.9)
(75,90)	(–,–)	(67.5,–)	(75.0,90.0)	(74.9,90.0)

6.5 Conclusion

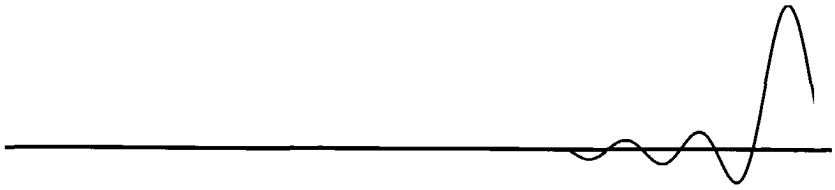
In this chapter, I have presented analysis of the effects of model mismatches on the Joint Angle-Frequency Estimation (JAFE) algorithm. Particularly, the effects of frequency hopping and modulation, moving receivers and sources, and distortions caused by antenna coupling and phase errors are considered.

I have shown that the JAFE algorithm can be elegantly used to detect, estimate and track frequency hopping signals. With the analysis, I have shown that, for a given SNR and number of antenna elements M, there exists a lower limit on the size of the frequency hop which guarantees a specified performance level. I have also applied the algorithm to estimate the bandwidths of modulated signals, and have shown that the method gives a good estimates of the bandwidths when the signal energies are symmetrically distributed in the frequency bands.

When the signal source and the receiving system are moving, the signal parameters become functions of time. I have shown that, for surface transportation systems (top speed of about 100 km/h), the model can be considered to be piece-wise stationary, and the algorithm continues to work without significant degradation.

Finally, I have shown that, if high-resolution DOA estimation algorithms are required to achieve their theoretically expected performances, it is mandatory to appropriately compensate the distortions caused by non-ideal array behaviors. Here, a simple off-line compensation method was employed. For better results, it may be useful to fine-tune the distortion matrix using on-line blind compensation algorithms such as those described in [19,59,75,95]. Moreover, the solution should consider the effect of frequency on the coupling matrix.

Chapter 7



IMPLEMENTATIONAL ISSUES

Contents

In this chapter, I give some indications of how much computation is involved in the implementation of the JAFE algorithm. I base my discussion on an arbitrary antenna array with M elements, and data collected using the measurement set up of Fig. 5.6. This allows a multi-resolution (MR) temporal sampling. I omit a MR spatial sampling in order not to put any restriction on the array geometry. This means that the MR parameter estimation technique of chapter 5 is applied

to frequency estimation only. To be specific, assume that the antenna array, with array steering vector $\boldsymbol{A}$ (viz. chapter 2) is impinged by d narrowband far field signals. The data is collected into a $2mM \times N - m + 1$ augmented data matrix $\boldsymbol{Z}$ as discussed in section 5.4, where N is the number of temporal samples and m is the temporal smoothing factor. In reference to Fig. 5.6, $\boldsymbol{Z}$ is given as

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{bmatrix}. \tag{7.1}$$

I assume that

- Signals propagate in distinct paths
- a two dimensional antenna array, where $\boldsymbol{A}$ has two shift parameters (parameterized DOAs), one relative to the x axis denoted by θ and one relative to the y axis denoted by γ (viz. section 2.4.1). The argument μ_x of θ is proportional to the sine of the angle of arrival α and the argument μ_y of γ is proportional to its cosine.
- $\boldsymbol{Z} \in \mathbb{C}^{2mM, N-m+1}$ contains two temporal sampling periods T_1 and T_2, which may be used for multi-resolution frequency estimation. We normalize to $T_2 = 1$.

For $i = 1, \cdots, d$ and $k = 1, 2$, let f_i be the center frequency of the i-th signal and $\phi_{i,k}$ be defined as $\phi_{i,k} = e^{j2\pi T_k f_i}$, respectively. Moreover, let the $d \times N - m + 1$ matrix $\boldsymbol{F}_s$ be the collection of $N - m + 1$ time samples of the d sources defined in (3.5). Put $\boldsymbol{\Phi}_k := \text{diag}\{\phi_{i,k}\}_{i=1}^{2}$, and for some integer $m \geq 1$, let $\boldsymbol{A}_m$ be defined as

$$\boldsymbol{A}_m = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{A}\boldsymbol{\Phi}_2 \\ \vdots \\ \boldsymbol{A}\boldsymbol{\Phi}_2^{m-1} \end{bmatrix}$$

Then, $\boldsymbol{Z}$ can be expressed as

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{A}_m \\ \boldsymbol{A}_m\boldsymbol{\Phi}_1 \end{bmatrix} \boldsymbol{F}_s + \boldsymbol{N}, \tag{7.2}$$

where $\boldsymbol{N}$ is a noise term. In the following, after summarizing the different operations involved in the JAFE algorithm, I present estimation of the computational complexity of each stage.

7.1 Basic Operations Involved in JAFE Algorithm

1. *Singular value decomposition (SVD) of the data matrix*: An efficient implementation of the SVD (which can partly be implemented in adaptive form) is obtained using a two step strategy. First, compute the QR decomposition of $\boldsymbol{Z}^H$:

$$\boldsymbol{Z}^H = \boldsymbol{Q}\boldsymbol{R},$$

and then the SVD of $\boldsymbol{R}^H$:

$$\boldsymbol{R}^H = \boldsymbol{U}\Sigma\boldsymbol{V}^H.$$

The unitary matrix $\boldsymbol{U}$ spans the column space of $\boldsymbol{Z}$. After rank reduction (removal of the noise subspace), we get a $2mM \times d$ matrix $\boldsymbol{U}_s$ that spans the signal subspace. $\boldsymbol{U}_s$ and $\boldsymbol{A}_m$ are related via an invertible $d \times d$ matrix $\boldsymbol{T}$:

$$\boldsymbol{U}_s = \boldsymbol{A}_m\boldsymbol{T}$$

2. *Applying selection matrices to $\boldsymbol{U}_s$*: This requires no computation, a set of memory accesses only. However, these memory accesses could be time consuming. Thus, for our purpose, we assume that a memory access is equal to to a half multiplication. To estimate the two temporal parameters $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$, and the two spatial parameters $\boldsymbol{\Theta} := \text{diag}\{\theta_i\}_{i=1}^d$ and $\boldsymbol{\Gamma} = \text{diag}\{\gamma_i\}_{i=1}^d$, we need to define four pairs of selection matrices. For $M_x, M_y < M$ let the pairs $(\boldsymbol{I}_{x,\theta}, \boldsymbol{I}_{y,\theta}) \in \mathbb{R}^{M_x \times M}$ and $(\boldsymbol{I}_{x,\gamma}, \boldsymbol{I}_{y,\gamma}) \in \mathbb{R}^{M_y \times M}$ be the selection matrices operating on $\boldsymbol{A}$ to generate shift invariance matrix pairs to estimate the parameterized DOAs θ and γ, respectively. For a 4-UCA array, for instance, these matrices are as given in (2.26). The four pairs of selection matrices that operate on $\boldsymbol{U}_s$ to generate the above four parameters, respectively, are then given by

$$\begin{cases} \boldsymbol{J}_{x\phi_1} &= [1 \quad 0] \otimes \boldsymbol{I}_m \otimes \boldsymbol{I}_M \\ \boldsymbol{J}_{y\phi_1} &= [0 \quad 1] \otimes \boldsymbol{I}_m \otimes \boldsymbol{I}_M \end{cases} \in \mathbb{R}^{mM,2mM}$$

$$\begin{cases} \boldsymbol{J}_{x\phi_2} &= \boldsymbol{I}_2 \otimes [\boldsymbol{I}_{m-1} \quad \boldsymbol{0}] \otimes \boldsymbol{I}_M \\ \boldsymbol{J}_{y\phi_2} &= \boldsymbol{I}_2 \otimes [\boldsymbol{0} \quad \boldsymbol{I}_{m-1}] \otimes \boldsymbol{I}_M \end{cases} \in \mathbb{R}^{2(m-1)M,2mM}$$

$$\begin{cases} \boldsymbol{J}_{x,\theta} &= \boldsymbol{I}_2 \otimes I_m \otimes \boldsymbol{I}_{x,\theta} \\ \boldsymbol{J}_{y,\theta} &= \boldsymbol{I}_2 \otimes I_m \otimes \boldsymbol{I}_{y,\theta} \end{cases} \in \mathbb{R}^{2mM_x,2mM}$$

$$\begin{cases} \boldsymbol{J}_{x,\gamma} &= \boldsymbol{I}_2 \otimes I_m \otimes \boldsymbol{I}_{x,\gamma} \\ \boldsymbol{J}_{y,\gamma} &= \boldsymbol{I}_2 \otimes I_m \otimes \boldsymbol{I}_{y,\gamma} \end{cases} \in \mathbb{R}^{2mM_y,2mM}$$

3. *Generating a joint eigenvalue problem*: This is accomplished by computing the following matrix ratios.

$$\begin{aligned} \boldsymbol{E}_{\phi_1} &= (\boldsymbol{J}_{x\phi_1}\boldsymbol{U}_s)^\dagger(\boldsymbol{J}_{y\phi_1}\boldsymbol{U}_s) \\ \boldsymbol{E}_{\phi_2} &= (\boldsymbol{J}_{x\phi_2}\boldsymbol{U}_s)^\dagger(\boldsymbol{J}_{y\phi_2}\boldsymbol{U}_s) \\ \boldsymbol{E}_{\theta} &= (\boldsymbol{J}_{x,\theta}\boldsymbol{U}_s)^\dagger(\boldsymbol{J}_{y,\theta}\boldsymbol{U}_s) \\ \boldsymbol{E}_{\gamma} &= (\boldsymbol{J}_{x,\gamma}\boldsymbol{U}_s)^\dagger(\boldsymbol{J}_{y,\gamma}\boldsymbol{U}_s) \end{aligned} \tag{7.3}$$

4. *Solving the joint eigenvalue problem*: The joint eigenvalue decomposition problem is solved by jointly diagonalizing

$$\begin{aligned} \boldsymbol{E}_{\phi_1} &= \boldsymbol{T}^{-1}\boldsymbol{\Phi}_1\boldsymbol{T} \\ \boldsymbol{E}_{\phi_2} &= \boldsymbol{T}^{-1}\boldsymbol{\Phi}_2\boldsymbol{T} \\ \boldsymbol{E}_{\theta} &= \boldsymbol{T}^{-1}\boldsymbol{\Theta}_x\boldsymbol{T} \\ \boldsymbol{E}_{\gamma} &= \boldsymbol{T}^{-1}\boldsymbol{\Theta}_y\boldsymbol{T}. \end{aligned}$$

7.2 The Computational Loads

In the following, I give the computational load involved in each of the processing stages discussed above. To give some feeling of some of the symbols, I give the number of floating point operations required by each stage for a 4-UCA antenna array, where $M = 4$, $N = 128$, $d = 5$ and $m = 2$. For these choices, the dimensions of the various matrices defined in the previous sections are as shown in table 7.1.

Table 7.1: Dimensions of Matrices

Matrix	Dimen.	Matrix	Dimen.
$\boldsymbol{Z}$	32×126	$\boldsymbol{R}$	32×32
$\boldsymbol{J}_{x\phi_1}, \boldsymbol{J}_{y\phi_1}$	16×32	$\boldsymbol{J}_{x\phi_2}, \boldsymbol{J}_{y\phi_2}$	24×32
$\boldsymbol{J}_{x\theta}, \boldsymbol{J}_{y\theta}$	16×32	$\boldsymbol{J}_{x\gamma}, \boldsymbol{J}_{y\gamma}$	16×32
$\boldsymbol{U}_s$	$32 \times d$	$^{**}\boldsymbol{E}_i$	$d \times d$

$^{**}i = \{\phi_1, \phi_2, \theta, \gamma\}$

7.2.1 Singular value decomposition

Computationally, the SVD stage is the most demanding one. Depending on ones interest, the SVD can be implemented in adaptive or non adaptive manner.

In either case, an efficient SVD implementation is obtained using a two stage processing strategy as discussed in the foregoing section.

- Use adaptive (or non adaptive) QR decomposition to compute

$$\boldsymbol{Z}^H = \boldsymbol{Q}\boldsymbol{R}$$

Recalling that $\boldsymbol{Z}$ is an $2mM \times N - m + 1$ matrix, this computation requires [22] [1].

$$\mathcal{N}_{svd,1} = 32m^2M^2(N - m + 1 - \tfrac{2}{3}mM).$$

- The second step is to compute the SVD of the $2mM \times 2mM$ matrix $\boldsymbol{R}$. Since we are interested with the left singular vectors only, we don't need to compute the complete decomposition. This reduces the computational complexity considerably. The required number of flops amounts to [22]

$$\mathcal{N}_{svd,2} = 192m^3M^3.$$

Summing the above two components $\mathcal{N}_{svd,1}$ and $\mathcal{N}_{svd,2}$, we obtain the total number of floating point operations required for the computation of the SVD of the $2mM \times N - m + 1$ data matrix $\boldsymbol{Z}$,

$$\mathcal{N}_{svd} = 32m^2M^2(N - m + 1 + \tfrac{16}{3}mM) \text{ flops.}$$

For $d = 5, m = 2, M = 4$ and $N = 128, \mathcal{N}_{svd} \approx 345$ $k$flops.

7.2.2 Application of Selection Matrices to $\boldsymbol{U}_s$

Before going into the computations of the complexity of this stage, it is useful to have a closer look to the dimensions of the selection matrices given in (7.3). To keep the discussion general, let $\boldsymbol{J}_{xi}$ and $\boldsymbol{J}_{yi}$, $i = \{\phi_1, \phi_2, \theta, \gamma\}$ represent one of the selection matrix pairs given in (7.3), and let D_i be the number of rows of $\boldsymbol{J}_{xi}$ and $\boldsymbol{J}_{yi}$, then

$$D_i = \begin{cases} mM, & i = \phi_1 \\ 2(m-1)M, & i = \phi_2 \\ 2mM_x, & i = \theta \\ 2mM_y, & i = \gamma \end{cases} . \tag{7.4}$$

A selection matrix chooses a set of rows from $\boldsymbol{U}_s$. This requires memory accesses only. In fact, the memory accesses are conducted during the subsequent

[1] Note that, since we are working with complex numbers, each operation is multiplied by a factor 4 as compared to those given in [22]

computations only. The selection matrices, therefore, do not physically exist, and the products $\boldsymbol{J}_i\boldsymbol{U}_s$ are not explicitly executed. However each memory access requires some time, thus assuming this accounts about a half of a multiplication, for the eight selection matrices with the dimensions as in (7.4), we have

$$\mathcal{N}_{sel} = 4m^2M(M_x + M_y + (3/2 - \frac{1}{m})M) \text{ flops}$$

For the 4-UCA, $M_x = M_y = 2$, thus $\mathcal{N}_{sel} = 512$ flops.

7.2.3 Generating the Joint eigenvalue Problem

Consider the formulas given in 7.3, and let $\boldsymbol{J}_{xi}$ and $\boldsymbol{J}_{yi}$ for $i = \{\phi_1, \phi_2, \theta, \gamma\}$, represent one of the selection matrix pairs, then we have,

$$\boldsymbol{E}_i = (\boldsymbol{J}_{xi}\boldsymbol{U}_s)^\dagger(\boldsymbol{J}_{yi}\boldsymbol{U}_s). \tag{7.5}$$

It is seen that the $\boldsymbol{E}_i$ are constructed by first computing the left-inverses of the $D_i \times d$ complex matrix $\boldsymbol{J}_{xi}\boldsymbol{U}_s$ and then taking the inner products $(\boldsymbol{J}_{xi}\boldsymbol{U}_s)^\dagger\boldsymbol{J}_{yi}\boldsymbol{U}_s$ of complex matrices. The number of floating point operation required for the inversion is $16d^2D_i$ and for the inner product is $8d^2D_i$. Thus, to construct $\boldsymbol{E}_i$, a total of $24d^2D_i$ floating point operations are required. Considering all the four possible values of $\boldsymbol{E}_i, i = \{\phi_1, \phi_2, \theta, \gamma\}$, the total number of operations required to generate the joint matrix pencil problem becomes $\mathcal{N}_{pen} = 24d^2\sum_i D_i$, substituting for the D_i, we get

$$\mathcal{N}_{eig} = 24d^2m(M_x + M_y + (1 - 1/m)M) \text{ flops.}$$

For $M_x = M_y = 2$ and $d = 5$, $\mathcal{N}_{eig} = 7.2$ $k$flops.

7.2.4 Solving the Joint eigenvalue Problem

The joint diagonalization of the matrix pencil pairs is the final stage in the JAFE algorithm. The problem is to find a single non-singular $d \times d$ matrix $\boldsymbol{T}$ that minimizes the off-diagonal energies of all the $\boldsymbol{E}_i$ matrices in (7.3). To solve this multi-dimensional optimization problem, one might envisage a number of ways [45,86,88] (viz. chapter 3). We stick to a very simple suboptimal method. First, we compute the eigenvalue decomposition (EVD) of $\boldsymbol{E}_{\phi_1}$ as

$$\boldsymbol{E}_{\phi_1} = \boldsymbol{T}^{-1}\boldsymbol{\Phi}_1\boldsymbol{T}, \tag{7.6}$$

then we apply the transformations

$$\boldsymbol{\Phi}_2 \approx \boldsymbol{T}\boldsymbol{E}_{\phi_2}\boldsymbol{T}^{-1}$$

$$\begin{aligned} \boldsymbol{\Theta}_x &\approx \boldsymbol{T}\boldsymbol{E}_\theta\boldsymbol{T}^{-1} \\ \boldsymbol{\Theta}_y &\approx \boldsymbol{T}\boldsymbol{E}_\gamma\boldsymbol{T}^{-1} \end{aligned} \tag{7.7}$$

The eigenvalue decomposition step in (7.6) requires [22] $\mathcal{N}_{eig} = 80d^3$ flops. The expressions in (7.7) involve a single $d \times d$ matrix inversion ($16d^3$ flops) and six $d \times d$ matrix multiplications ($8d^3$ flops each). Thus, the total number of operations amounts to

$$\mathcal{N}_{diag} = 144d^3 \text{ flops.}$$

For $d = 5, \mathcal{N}_{diag} \approx 18$ $k$flops.

7.2.5 The overall Computational complexity of the JAFE Algorithm

The total number of floating point operations needed for the JAFE algorithm is obtained by summing the above components. Note that the complexities were derived assuming a non-adaptive QR processing. Thus, the numbers represent number of flops per $2mM \times N - m + 1$ data block.

$$\begin{aligned} \mathcal{N}_{jafe} = 32m^2M^2 \Big(N - m + 1 + \tfrac{16}{3}mM + \tfrac{6d^2}{8mM}(1 + \tfrac{M_x+M_y}{M} \\ + \tfrac{6d}{mM} - \tfrac{1}{m}) + \tfrac{M_x+M_y}{8M} - \tfrac{1}{8m} + \tfrac{3}{16}\Big) \text{ flops.} \end{aligned} \tag{7.8}$$

Substituting the assumed values for m, M, N, d, M_x and M_y, we get $\mathcal{N}_{jafe} \approx$ 370 $k$flops.

7.3 Real Processing

In this section, I derive the computational complexity of the JAFE algorithm when it is implemented using the unitary-ESPRIT.

7.3.1 Constructing the Real Data and Applying selection Matrices

As described in section 3.8, the complex forward backward averaged data $\boldsymbol{X}$ is converted into equivalent real matrix using an $n \times n$ left π-real matrix $\boldsymbol{Q}_n$ as

$$\mathcal{T}(\boldsymbol{X}) = \boldsymbol{Q}_{mM}^{-1}\boldsymbol{X}\boldsymbol{Q}_{2(N-m+1)}.$$

If $\boldsymbol{Q}_n$ is chosen to be

$$\boldsymbol{Q}_n = \begin{cases} \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I}_n & j\boldsymbol{I}_n \\ \boldsymbol{\Pi}_n & -j\boldsymbol{\Pi}_n \end{bmatrix}, & \text{n even} \\ \\ \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I}_n & 0 & j\boldsymbol{I}_n \\ 0 & \sqrt{2} & 0 \\ \boldsymbol{\Pi}_n & 0 & -j\boldsymbol{\Pi}_n \end{bmatrix}, & \text{n odd} \end{cases} \tag{7.9}$$

this requires only $2mM(N - m + 1)$ additions and no multiplication. The application of the selection matrices to generate the individual real valued shift invariance pairs given in (3.41) requires $(mM - 1)d$ real additions and d real multiplications [27]. For the JAFE with multi-resolution frequency estimation and a two dimensional array there are four shift invariance pairs similar to that of (3.41) requiring a total of $4mM \cdot d$ floating point operations (flops).

7.3.2 SVD and other processing stages

Now since all the computations are implemented using exclusively real matrices the computational complexity drops significantly. First note that as a result of forward backward averaging, the number of sample points is doubled to $2(N - m + 1)$. However, each complex multiplication is now replaced with one real multiplication. Thus, an estimate of computational complexity for the real processing is obtained by dividing the (7.8) by four and replacing the term $N - m + 1$ with $2(N - m + 1)$

$$\mathcal{N}_{real} = 8m^2M^2 \Big(2(N - m + 1) + \tfrac{16}{3}mM + \tfrac{6d^2}{8mM}(1 + \tfrac{M_x+M_y}{M} + \tfrac{6d}{mM} - \tfrac{1}{m}) + \tfrac{M_x+M_y}{8M} - \tfrac{1}{8m} + \tfrac{3}{16} + \tfrac{d}{2mM}\Big) \text{ flops.} \tag{7.10}$$

The last term in (7.10) is due to the additional computation needed for the forward backward averaging [27, 45] and application of the *unitary ESPRIT* selection matrices. Substituting the respective values for m, M, N, d, M_x and M_y, we get $\mathcal{N}_{real} \approx 160$ $k$flops.

7.4 Discussion

In this chapter, I have given estimates of the computational complexity of the JAFE algorithm. I have shown that for complex processing the complexity is $\mathcal{O}(32Nm^2M^2)$ and for real processing it is $\mathcal{O}(16Nm^2M^2)$. These numbers are a bit misleading because in the real processing the number of temporal

samples is doubled as the result of the forward backward averaging. Thus, the accuracy obtained via the real processing is better than that obtained via the complex processing. This means that a complexity reduction of a factor 2 is obtained in addition to the improvement in accuracy. Note that the numbers given above are per processing data block. To obtain the complexity per sample, we need to divide these by N.

In most cases, the sources change location slowly compared to the extent of the data window (viz. section 6.3), the updating of the parameter estimates can be conducted much more slowly than the sampling rate. This further reduces the needed floating point operations per unit time. For instance, in the case $M = 4$,$m = 2$,$N = 128$, if we assume that updating is required every 1 ms, then the number of operations per second are about 370 Mflops for the complex processing and about 160 Mflops for the real processing. These are quite realistic numbers to be processed on today's DSP board.

Chapter 8

EXPERIMENTAL RESULTS

Contents

In this chapter, the behavior of the ESPRIT based joint angle-frequency estimation (JAFE) technique is analyzed using the data supplied by the TNO-FEL.

The objective is to validate both the JAFE algorithm and the data acquisition procedure. The later is achieved using information obtained from a calibration measurement and by verifying the results of JAFE using an other DOA estimation method.

The experiment was conducted at the premises of the TNO-FEL, the Netherlands. The channel behavior fits into a typical rural area scenario. In the following, I present two sets of results corresponding to two experiments. The first one aims in estimating the distortions introduced by antenna coupling and phase error, whereas the second one aims in evaluating the performance of the angle and frequency estimation.

8.1 Calibration

The calibration process is subdivided into two. In the first case the antenna coupling and the phase distortion introduced in each channel is modeled. In the the second calibration setup, phase jitters introduced by the AD (analog to digital) converters are modeled. While the former introduces biases in the DOA estimates the latter introduces some variation in the estimated DOAs.

8.1.1 Estimating the Coupling Matrix

To estimate the coupling error, a narrowband signal source with a center frequency of 108.9 MHz was placed at a distance of approximately 260 m from a ULA with M=4 elements and a baseline separation of 109 cm (0.4λ). The antenna array was mounted on a rotating table, such that with a fixed source position, it was possible to generate different angles of arrivals by varying the orientation of the receive array. The received signal was down converted to an IF frequency with a local oscillator operating at 104.9 MHz. After low pass filtering with a cutoff frequency of 10 MHz, the signal was then sampled at 40 MHz (2 times the Nyquist rate to avoid aliasing of the transitional band of the low pass filters). As can be seen from the frequency spectrum of the received signal in Fig. 8.1, the measurement conditions indeed fit into a single source scenario.

The measurements were conducted for DOAs varying from -75 to 75 degrees with steps of 15 degrees. The distortion matrix was estimated with the procedure described in section 6.4.2 making use of the measurements corresponding to the DOA $= [-75 \ -30 \ 0 \ 45 \ 75]$ Degrees. The ESPRIT and MUSIC algorithms were applied to estimate the DOAs for each measurement. The performance improvements achieved with the distortion compensation are summa-

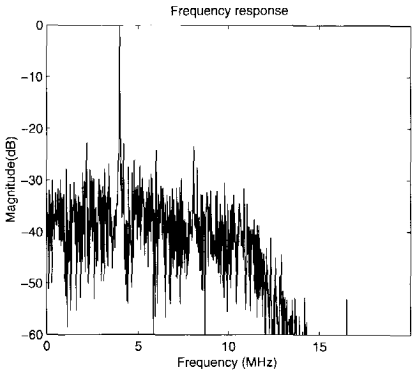


Fig. 8.1: The frequency spectrum of the received signal

rized in figures 8.2 and 8.3 for MUSIC and ESPRIT approaches, respectively. In table 8.1, the expected and the estimated DOAs (before and after distortion compensation) are shown.

From the results, it is seen that the calibration procedure indeed improves the performance of the estimator. However, it does not eliminate the biases completely. This is partly because, while rotating the receive array, the radiation pattern of the immediate surrounding in reference to each antenna element becomes function of DOA. Since the coupling pattern described in section 6.4.2 does not model this effect, the discrepancy is justifiable.

To see the effect of the compensation when more than one sources are involved, the estimation process was conducted for a two source case. Two narrowband signals with IF center frequencies 4 and 5 MHz and DOAs 75 and 60 degrees, respectively, were considered. The MUSIC and ESPRIT results are summarized in table 8.2. Again, it is seen that the compensation does not completely solve the coupling problem. Generally, the radiation pattern and, consequently, the coupling pattern of an antenna array is frequency dependent. Thus, in this case, apart from the problem described in the previous paragraph, frequency dependency of the coupling matrix induces more discrepancy in to the model.

8.1.2 Effect of Channel Phase Distortion

Referring back to section 6.4, we see that the diagonal entries in the coupling matrix account for the relative phase offsets in the different signal processing channels. Thus the effect of unequal signal propagation delays, has been accounted for with the coupling matrix. What the coupling matrix does not

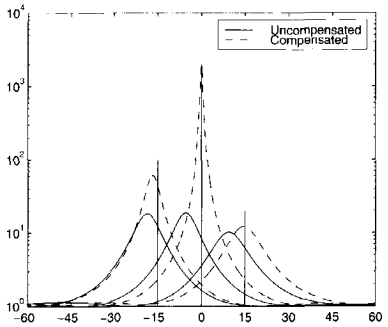


Fig. 8.2: The music Spectra, for three DOAs. The true DOAs are indicated by solid vertical lines

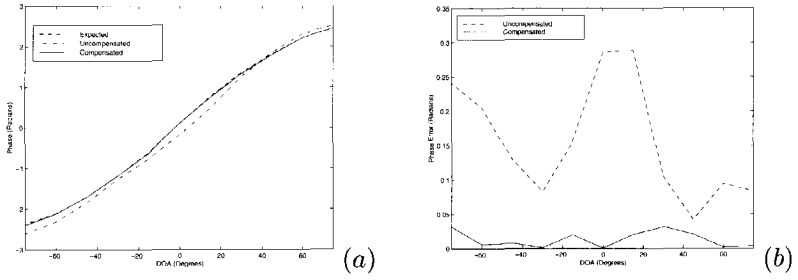


Fig. 8.3: a) Phase behaviors as functions of DOA, and b) Errors in the phase estimates before compensation (dashed line) and after compensation (solid line)

show is the phase jitter in the channels. Due to the quantization process, the AD converters introduce some phase jitter in the channels. This is considered below.

8.1.3 Measuring the Phase Jitter

The four-antenna-setup used to collect data for the JAFE algorithm is shown in Fig. 8.4. To measure the phase jitter, the A/D converters were fed with equal phase signals from a highly stabilized signal generator operating at 109 MHz. The calibration data was collected by measuring the phase delays between the reference channel a and the other channels. The resulting mean phase differ-

Table 8.1: Effect of distortion compensation on DOA estimates (single source)

Expected	Uncompensated		Compensated	
	ESPRIT	MUSIC	ESPRIT	MUSIC
75	87.5	87.5	75.4	74.8
60	71.9	65.1	59.8	59.0
45	47.9	45.4	45.5	43.6
30	27.8	28.0	31.3	30.0
15	12.0	9.4	13.6	14.6
0	-5.6	-5.5	0.0	0.0
-15	-20.7	-18.5	-16.5	-16.7
-30	-35.5	-30.6	-30.3	-30.2
-45	-60.4	-47.1	-44.8	-45.0
-60	-92.5	-67.4	-59.4	-60.3
-75	-92.5	-92.5	-75.9	-76.3

Table 8.2: Effect of distortion compensation on DOA estimates (two sources)

Expected	Uncompensated		Compensated	
	ESPRIT	MUSIC	ESPRIT	MUSIC
75	77.9	78.3	74.6	72.9
60	68.6	70.8	62.6	63.1

ences and their standard deviations corresponding to 50 ms signal segment are summarized in table 8.3. In the table, the second column indicates the delays in μs and the third column in radians. The last column gives the standard deviations of the jitters on the measured phase delays (in radians). Assuming these jitters are zero mean random processes, they can be modeled using an equivalent zero mean additive Gaussian noise. If the phase jitter is required to have negligible influence on the estimation, the equivalent noise σ_μ must be below

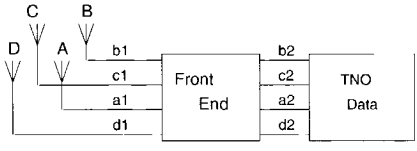


Fig. 8.4: Measurement setup

Table 8.3: Phase offsets measured with respect to the signal in channel a (the reference channel)

Channel	$\Delta\tau(\mu s)$	$\Delta\mu$ (rad)	σ_μ (rad)
B	0.0248	0.157	0.0076
C	0.0276	0.173	0.0131
D	0.0137	0.086	0.0070

the background noise level, i.e., (viz. [47,65]),

$$\sigma_\mu < \sqrt{\sigma_n^2 \left(\frac{1}{M^2 N}\right)}$$

For instance, for $\sigma_n^2 = 0.5$ (normalized with respect to the signal power), $M = 4$ and $N = 64$ we require $\sigma_\mu < 0.022$.

8.2 The ESPRIT based JAFE Estimation Procedure

The data supplied by TNO-FEL was sampled at 40 MHz. This sampling rate is 2 times the Nyquist frequency. It is shown in chapter 5, however, that the (frequency) estimation error is inversely proportional to the sampling period. To exploit this property, the signal was down sampled before the application of the JAFE algorithm. The data is first split into blocks of 513 samples, and then two time shifts of the 4×513 data matrix are augmented into 8×512 matrix as discussed in chapter 5. Following this, the data is down sampled by a factor 4. The basic processing stages involved in the method are summarized in Fig. 8.5.

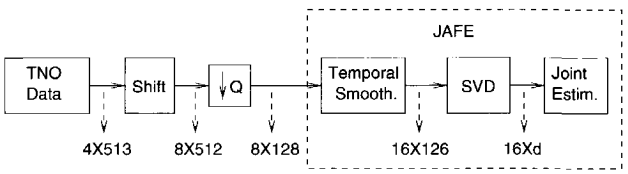


Fig. 8.5: The estimation processing stages

8.2.1 Computational complexity

As discussed in chapter 7, the computational complexity is dominated by the SVD stage. Let the size of the data block at the input of the SVD stage be $2mM \times N_m$, then the overall complexity is given by (7.8), for $d \ll \min(2mM, N_m)$, this simplifies to

$$N_{op} \approx 32m^2M^2(N_m + \frac{16}{3}mM);$$

In our case $2mM = 16, N_m = 126$ and, thus, $N_{op} \approx 350$ kflops/cycle. If we want to update the angle-frequency estimates, say, every 1ms, we need a digital signal processing board that has a speed of 350 Mflops/s.

8.3 Experimental Setup

The TNO supplied data comprises of 16 different measurement scenarios. These are represented by numbers adopted from the data documentation [8] (1A, 1B, and so on). In this section, I give parameter estimation results obtained using ESPRIT based JAFE method. To cross check the results, I also give estimates generated by the MUSIC algorithm. The data were collected using two kinds of array geometries: ADCOCK and ULA. For each of the arrays the angular position of the sources are measured as indicated in Fig. 8.6 and Fig. 8.7, respectively.

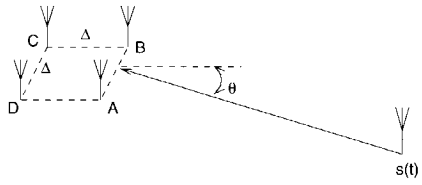


Fig. 8.6: ADCOCK antenna array

8.3.1 Measurement scenario 1 (single source)

In this scenario, a single source (at a center frequency of 109 MHz) is put in a far field (50 m from the array). The data is collected under six different measurement conditions, which are summarized below.

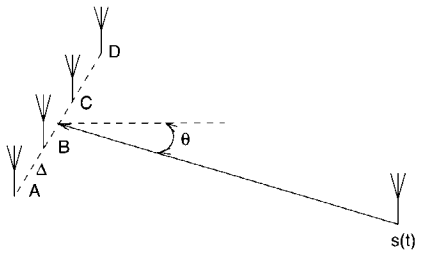


Fig. 8.7: Uniform linear antenna array

- Setup 1A: In this setup, a single source is placed at an angular position of $\theta = 45^o$ from an ADCOCK antenna array. The adjacent antenna elements are put at a distance of $\Delta = 0.246\lambda$ from each other. Here, λ is the wavelength of the signal (at 109 MHz, $\lambda = 2.73$ meters).

- Setup 1B: Same as setup 1A but with $\Delta = 0.357\lambda$

- Setup 1C: Same as setup 1A but with $\Delta = 0.5\lambda$

- Setup 1D: Same as setup 1A but with $\theta = 60^o$

- Setup 1E: Same as setup 1A but with $\theta = 75^o$

- Setup 1F: Same as setup 1A but with the ADCOCK array replaced with ULA, and the antenna elements spaced by $\Delta = 0.348\lambda$.

8.3.2 Measurement scenario 2 (Two distinct sources)

In measurement setups 2, two sources generating signals at center frequencies $f_{c1} = 109$ and $f_{c2} = 110$ MHz and having distinct DOAs of θ_1 and θ_2, respectively, are considered. The angles are measured as indicated in Fig. 8.6 for the ADCOCK arrangement and in Fig. 8.7 for the ULA arrangement. At the RF stage, the center frequencies are shifted back to the baseband (using a local oscillator with $f_o = 108$ MHz) such that $f_1 = 1$ and $f_2 = 2$ MHz. Both signals are transmitted with the same power at approximately the same distance from the antenna array. The remaining details of the measurement setups are specified in table 8.4.

Table 8.4: Summary of the measurement setup for experiments 2. Angles are in Degrees and frequencies in MHz. The antenna spacings are $\Delta = 0.246\lambda$ for ADCOCK and $\Delta = 0.348\lambda$ for ULA.

setup	Array type	θ_1	θ_2	f_1	f_2
2A	ADCOCK	60	69	1	2
2B	ULA	45	54	1	2
2C	ADCOCK	60	60	1	2
2D	ULA	45	45	1	2

8.3.3 Measurement scenario 3 (Multi-path propagation)

In this scenario, six measurement setups were used. Three using ADCOCK array and three using ULA. For ADCOCK antenna $\Delta = 0.246\lambda$ and for ULA $\Delta = 0.348\lambda$. The details of experimental setups are summarized below.

- Setup 3A: Here, a source placed at an angular position of $\theta = 60^o$ from ADCOCK antenna generates a signal propagating in two artificially generated paths. The first path is the path of sight (direct path with $\theta_1 = 60^o$), and the other path is a secondary- (multi-) path with an angular direction $\theta_2 = 65.5^o$. The ratio of the direct path signal energy to the multi-path signal energy is 20 dB. Angles are measured as in Fig. 8.6.
- Setup 3B: Same as 3A, but with ADCOCK antenna replaced with ULA, and $\theta_1 = 45^o$ and $\theta_2 = 50.5^o$. Angles are measured as in Fig. 8.7.
- Setup 3C: Same as 3A but with direct path to multi-path signal energy ratio of 1 (0 dB).
- Setup 3D: Same as 3B but with direct path to multi-path signal energy ratio of 1 (0 dB).
- Setup 3E: Here, a source with two artificial multi-path propagation is considered. the DOAs are $\theta_1 = 60^o$ for the direct path, $\theta_2 = 65.5^o$ for the first multi-path, and $\theta_3 = 71^o$ for the second multi-path. The ratios of the multi-path signal energies to the direct line signal energy are -10 dB and -20 dB, respectively. Here, an ADCOCK array is used.
- Setup 3F: Same as 3E, but with ADCOCK antenna replaced with ULA, and $\theta_1 = 45^o$, $\theta_2 = 50.5^o$ and $\theta_3 = 56^o$.

8.4 Results

Experimental results for ESPRIT based JAFE and MUSIC methods are summarized below. The values in the table are given after the biases due to coupling have been accounted for, and the results are based on 100 trials.

Table 8.5: Single Source

Experimental Setup		Setup 1A	Setup 1B	Setup 1C	Setup 1D	Setup 1E	Setup 1F
DOA (Deg)	Mean	45.49	45.77	44.86	59.2	75.2	45.47
	RMSE	0.90	0.57	0.44	0.58	0.89	0.98
	CRB**	0.042	0.028	0.021	0.060	0.12	0.030
Freq. (MHz)	Mean	1.000	0.999	0.999	1.000	1.000	1.001
	RMSE	0.0027	0.0038	0.0026	0.0028	0.0025	0.0022
	CRB	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018
Estimation based on MUSIC method							
	Mean	46.2	46.4	45.12	58.6	74.4	45.2
DOA (Deg)	RMSE	0.92	0.79	0.58	0.84	0.98	0.86

Table 8.6: Experimental Setup 2. Column #1 correspond to the 109 MHz signal source and column #2 to the 110 MHz signal source.

Experimental Setup		Setup 2A		Setup 2B		Setup 2C		Setup 2D	
		#1	#2	#1	#2	#1	#2	#1	#2
DOA (Deg.)	Mean	62.1	79.3	44.7	58.4	59.6	60.0	45.1	46.5
	MSE	1.85	1.84	2.4	1.96	0.93	1.27	2.45	2.62
Freq. (MHz)	Mean	1.002	1.997	1.006	2.001	0.999	2.003	1.003	1.998
	MSE	0.007	0.005	0.0055	0.0058	0.0057	0.0069	0.0053	0.0116
Results of MUSIC algorithm									
DOA (Deg.)	Mean	62.3	79.6	45.0	58.4	59.1	59.7	45.5	47.4
	MSE	2.45	2.44	1.23	1.81	1.93	2.27	2.05	1.92

8.5 Remarks

Setups 3A and 3B

The results for these setups are summarized in table 8.7. As the multi-path signal was too weak to be detected, only one DOA was estimated. More precisely, the worst SNR of the direct path signal is about 2 dB, and that of the secondary

path (which is 20 dB below the direct path) is -18 dB.[1] This signal level is too weak to be detected by the implemented averaging level.

Table 8.7: Experimental results for setup 4

Setup	θ_o	$\hat{\theta}$	σ_θ	f_o	$\hat{f}$	σ_f
3A	15	16.7	0.53	1	1.000	0.0023
3B	45	44.6	0.69	1	1.001	0.0027

Setups 3C and 3D

Here a two-path scenario with equal signal energy were considered. The results show that, apart from the intended artificial multi-path, there was an additional signal path from the direction -6.5 degrees. Moreover, the two paths (θ_1, θ_2)=(45,50.5) degrees were too close to be separated. Consequently, a rank 2 estimation gave the angle pairs (48.02,-6.47) degrees instead of the expected (45,50.5). This shows that the two close paths were approximated by a single average path.

Setups 3E and 3F

Here a three-path propagation is considered. Because of the fact that the multipath components were deep in the noise, only a single path was detected.

8.6 Discussion

As seen from table 8.5, the angle estimates are far above the Cramer Rao Bound (CRB). This is partly caused by the relatively large phase jitter. If we convert the phase jitters in table 8.3 into corresponding errors in DOA, we get an RMS DOA error in the order of 0.5 Degrees at $\Delta = 0.25$. This means that the measured RMS error is highly influenced by the phase jitter apart from the noise.

One behavior that draws attention is the variation of the SNR between the antennas. As can be seen from table 8.8, The SNR is not the same for all the

[1] Theoretically the source can be lifted to above the noise if we process a data block of size $10^{3.6} \approx 4000$ samples.

antenna elements. Taking into account the free space attenuation, this is expected. However, it does not explain the 7 dB SNR difference between antenna elements A and C. In fact, the experimental results obtained by interchanging the positions of the antenna elements relative to the signal direction, show that the poor SNR at position C is mainly the result of antenna shadowing. When two antenna elements are placed on a straight line connecting the two elements and the source, the element nearest to the source acts as a sink absorbing some of the electromagnetic energy from the signal propagating to the next element. This, in turn, reduces the strength of the signal received at the second antenna element. Referring to Fig. 8.6, one can clearly see that antenna element A is shadowing element C, explaining the poor SNR at the latter element.

An unexpected behavior is seen between the errors at $\theta = 45$ and $\theta = 60$ degrees. That is, at 45 degrees, the estimation error is seen to be larger than that at 60 degrees. This can be explained (see table 8.8) by the slightly better SNR in setup 1D as compared to setup 1A, which is a consequence of shadowing. Referring to Fig. 8.6, it is seen that the antenna shadowing on C is much sever at 45 degrees (setup 1A) than at 60 degrees (setup 1D).

Table 8.8: SNR (in dB) at the four antennas for setups 1A and 1D

Setup	A	B	C	D
1A	7.6	9.6	2.2	7.2
1D	7.9	9.0	2.8	7.4

In the second part of table 8.5, estimation results corresponding to the MUSIC algorithm is presented. The results do not only confirm the above discussion regarding the errors in setups 1A and 1D, but also indicate that the biases seen in the estimates consistently agree with that of the JAFE techniques.

Finally, note that the results confirm that the estimation accuracy improves with an increased antenna separation. This is means that it is always best to put the antenna elements as far from each other as possible.

Referring to table 8.6, it is noticed that the results in 2A (corresponding to the signal #2) are far removed from the expectation. It looks that either there has been an unintentional multi-path propagation, or the bias due to the phase distortions is not correctly compensated.

If the first speculation is correct, then the collected data matrix has to bee rank deficient. And, if this was true, a rank 2 matrix pencil solution would be in-

consistent over the 100 Monte-Carlo trials. According to the results, however, both the ADCOCK and the ULA arrays consistently predict the second source to be at an angular displacement of about 17 and 14 Degrees from the first one. This makes the first reason a less likely cause for the observed offset. The consistency of the results, in fact, strongly support the second speculation, namely, inaccurate bias representation. Note that the coupling matrix was computed for a signal at 109MHz. In section 6.4, it has been shown that the coupling matrix could be highly dependent on the signal frequency. Thus, for proper compensation, the coupling matrix has to be determined at all relevant frequency values, or the antenna should be designed to have constant frequency behavior over the band of interest.

8.7 Conclusion and Remarks

8.7.1 Calibration

- If high-resolution DOA estimation algorithms are required to achieve their theoretically expected performances, it is mandatory to appropriately compensate the distortions caused by non-ideal array behaviors. Here, a simple off-line compensation method was employed. For better results, it may be useful to fine-tune the distortion matrix using on-line blind compensation algorithms such as those described in [19,59,75,95]. Moreover, the solution should consider the effect of frequency on the coupling matrix.

- The A/D converters introduces phase jitters that have the same effect as an additive noise. Their effect is masked under the background noise if the equivalent noise level is below the background noise level.

- The biases on the estimated DOAs are functions the signal frequency, the array geometry and the antenna response. This means that, a calibration data should make possible the correct determination of these dependencies.

8.7.2 Experiments

- The estimation of the parameters when a single source is present in the channel is quite accurate. However, the errors on the DOA estimates are seen to be much higher than the CRB. This was caused partly by a relatively large phase jitters in the A/D converters.

- It is seen that, between the antenna inputs, the SNR varied considerably. By interchanging the positions of the antenna elements in the array, it has been verified that this large amplitude variation was caused by the shadowing effect. Notice that the effect of unequal antenna gain is automatically accounted for by the coupling matrix.

- The parameter estimation involving multi-path propagation suffered from either too low (multi-path) signal energy or too close multi-path propagations. The former is not, actually, a problem since the method correctly detects the direct path DOA. The latter, however, resulted in a DOA estimate averaged over the two close angles. The scenario best fits into a local scatterer case (spatial modulation) instead of multi-path. Of course, at a higher SNR and with more number of antenna elements, a better result is expected.

Chapter 9

CONCLUSIONS

In this thesis I have presented a complete description of the JAFE (Joint Angle-Frequency Estimation) algorithm, ranging from the basic concepts to behavioral analysis, from mathematical models to physical implementation issues. In doing so, I have achieved all the intended goals. That is, I was able to estimate the DOAs and center frequencies of narrowband far field signals

- in a robust and reliable fashion, i.e., the solution works under the presence of multiple signals with multipath propagation, and under a considerable model mismatch.
- using minimal signal information (only the narrowband and far field assumptions have been made)
- with a complexity that allows its implementation on the available signal processing systems.

JAFE is not a completely new concept. It has been previously addressed by a few people. Compared to these results, I have presented the algorithm in a more formal and elegant fashion, and have addressed some issues that have not been previously considered. In the following, after presenting a summary of the main results, I discuss the unresolved issues and give some pointers towards further development.

9.1 Main contributions

The JAFE algorithm (or any ESPRIT type algorithm) involves three basic steps, namely, SVD (singularvalue decomposition) of the data matrix, diagonalization of a set of EVD (eigenvalue decomposition) problems and the transformation of the eigenvalues into actual signal parameters. The first step, which is equivalent to finding the EVD of the data covariance matrix, has been studied in the literature for the case of white Gaussian noise contaminated data. In my context, however, since some data stacking techniques have been employed, the noise is no more white. In this thesis, I have presented a performance analysis for the latter case. Moreover, I have shown how the estimation error is affected by the different data stacking parameters. The optimal values of these parameters have been derived, and the results were verified using simulation results. One interesting observation is that the data stacking parameters that maximize the number of identifiable signal parameters are different from those which minimize the parameter estimation errors. This means that, depending on the interest, one may trade between these two optimum values.

Generally, array geometries contain or can be designed to contain multiple shift invariance properties. In such cases the MR-ESPRIT algorithm may be used to improve the accuracy of the ESPRIT algorithm. MR-ESPRIT is a new extension of the ESPRIT algorithm that uses a multi-resolution spatio-temporal sampling rates to significantly improve the performance of the ESPRIT algorithm without raising the computational complexity. Or, it can be employed to attain a certain performance level with a reduced computational complexity. With the a behavioral analysis, I was able to derive bounds on the attainable accuracy improvements. These bounds are shown to be dependent, among other things, on noise level and array measurement errors. By employing a self calibrating MR-ESPRIT, I was able to mitigate (to a great extent) the sensitivity to the array measurement errors.

One way of analyzing the robustness of a given algorithm is to see how it behaves under non-ideal situations. To this end, I have considered different practical cases where the data models do not perfectly fit into the underlying assumptions. These include, effects of frequency hopping sources, effects of moving sources and moving receivers, effects of frequency and angle spreading, and effects of antenna coupling.

For the case of frequency hopping signals, I have shown that, for a given SNR and number of antenna elements, there exists a lower limit on the size of a frequency hop which guarantees a specified performance level. Moreover, I have presented simulation analysis showing how the model error caused by frequency hopping introduces biases in the parameter estimates. For the case

of modulated signals, I have shown how the algorithm can be applied to estimate the bandwidths of the modulated signals. It is seen that the method gives useful estimates of the bandwidths when the signal energies are symmetrically distributed in their frequency bands, which is universally true for almost all communication signals.

Generally, antenna elements in an array do not behave independent of each other. The reflected radiation from one element couples to its neighbors, as do currents that propagate along the surface of the array. Under these conditions, unless the distortions introduced by the non-ideal (practical) array behavior are properly accounted for, any high resolution algorithm can not attain the expected performance level. To this end, I have presented simulation and measurement results demonstrating that array modeling is an essential element of any algorithm that relies on data collected via array of antenna elements.

Finally, I have added a new value to the JAFE algorithm by validating it using actual field measurement data. The results show that actual physical phenomena such as antenna coupling, channel distortion and phase jitter in the analog to digital converters degrade the expected performance significantly. I have shown that, if the algorithm is required to achieve the expected performance, these non ideal behaviors must be adequately modeled and compensated for.

9.2 Unresolved Issues and Further Development

In this thesis, I have made effort to address all issues arising in implementing the JAFE algorithm. However, because of shortage of time and direction of focus, there are some issues that need further investigation. In the following, I discuss these issues and give some indications of their significance.

Supposing that there are sufficient numbers of spatial and temporal data samples, independent occurrences of multiple DOAs and multiple frequencies have been successfully resolved using certain data stacking techniques. However, when the multiplicities occur simultaneously (for instance like two sources having the same DOA and center frequency), the data stacking techniques fail to give meaningful results. Solutions to these cases should exploit additional signal behaviors such as constant modulus and finite alphabet. An interesting extension to this work could be improving the sensitivity of the JAFE algorithm in such instances by exploiting more signal behaviors than those considered in this thesis. Preliminary investigations show that the combination of the JAFE and the analytic constant modulus algorithms may give better results in cases where there are many multipaths, and when the signals are closely separated in space and frequency.

In the performance analysis, I have derived explicit expressions for the dependencies of the estimation errors on the data stacking parameters. These expressions were derived assuming that all the sources have a certain minimum spatial (angular) separation. However, one would expect that the spatio-temporal smoothing process should improve the performance of the algorithm tremendously in the particular cases when the signals are closely separated. Thus, finding similar dependencies for the cases where the sources are closely separated in space is a quite interesting and challenging problem.

One important stage in the JAFE algorithm is the joint diagonalization (JD) of a set of EVD problems. Apart from JAFE, JD of a set of equal dimensioned matrices appears in numerous applications. It has also been studied by many different researchers leading to various JD techniques. However, all of the known methods are sub-optimal. Finding a robust optimal JD solution is an open research topic. I believe an optimal analytical solution is a great and challenging research opportunity.

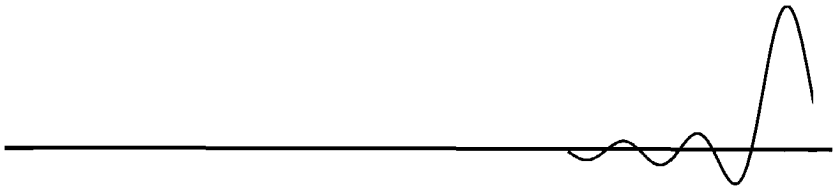
Finally, it is worthwhile mentioning that I have used JAFE to estimate angle and frequencies of slow frequency hopping signals. The data models considered do not apply to fast frequency hopping and (direct) spread spectrum signals. For such cases further analysis is required. However, it should be noted that frequency estimation gives meaning only when the signal bandwidths are sufficiently narrow. If all signals occupy the whole frequency band of interest, frequency estimation does not make any sense.

9.3 Final remarks

Apart from the applications mentioned in the thesis, the JAFE algorithm can be used to solve some interesting problems in modern communications systems. For instance, it is well known that residual carrier frequency in OFDM (Orthogonal Frequency Division Multiple-access) system causes non-orthogonality of the frequency channels. In these cases, MR-JAFE (multi-resolution JAFE) can be implemented to estimate the residual frequency. Moreover, JAFE based beamformer can be used to minimize cross-talks between adjacent bands, to implement SDMA (spatial division multiple access) systems, for location sensitive billing and so forth.

As a final remark, it is worthwhile to mention that, with the ever increasing demand on the communication channels and quality of services, the future direction to increase the bandwidth efficiency should couple the different multiple access algorithms with high resolution parameter estimation techniques, for instance, as the OFDM-JAFE combination mentioned in the previous paragraph.

Appendix A



PROPERTIES AND PROOFS

Contents

A.1 The Sensitivities of computing the DOA based on the sin, cos and tan transformations

Consider the three transformations given in (2.28), (2.29) and (2.30):

$$\alpha = \text{asin}(\frac{\mu_x}{2\pi}) \quad \alpha = \text{acos}(\frac{\mu_y}{2\pi}) \quad \text{and} \quad \alpha = \text{atan}(\frac{\mu_x}{\mu_y})$$

Assume that μ_x and μ_y are corrupted by independent noise, with variances $\sigma^2_{\mu_x} = \sigma^2_{\mu_y} = \sigma^2_{\mu}$. Moreover, let σ_s, σ_c and σ_t be the standard deviations of the DOA estimations based on the sin, cos and tan transformations, respectively. Then, using the transformation rule discussed in [39], for each of the above equations, we get

$$\sigma^2_s = \frac{\sigma^2_{\mu}}{4\pi^2 \cos^2 \alpha}$$

$$\begin{aligned} \sigma_c^2 &= \frac{\sigma_\mu^2}{4\pi^2 \sin^2 \alpha} \\ \sigma_t^2 &= \sigma_\mu^2 \end{aligned}$$

It is seen from these that, while the first two transformations become quite unstable around $\pm\frac{\pi}{2}$, and $[0, \pi]$, respectively, the third transformation remains stable at all DOA values.

A.2 Proof of theorem 4.1

Consider the Gaussian, circulant noise contaminated data model $\boldsymbol{X}_{m,L}$ given in (3.20). Let $\boldsymbol{Y}_{m,L} = \boldsymbol{X}_{m,L} - \boldsymbol{N}_{m,L}$ be the noise free data,

$$\begin{aligned} \boldsymbol{R}_{xx} &:= \frac{1}{L(N-m)} \boldsymbol{X}_{m,L} \boldsymbol{X}_{m,L}^H \\ \boldsymbol{R}_{yy} &:= \frac{1}{L(N-m)} \boldsymbol{Y}_{m,L} \boldsymbol{Y}_{m,L}^H \\ \boldsymbol{R}_{nn} &:= \frac{1}{L(N-m)} \boldsymbol{N}_{m,L} \boldsymbol{N}_{m,L}^H. \end{aligned}$$

In the following I use a perturbation analysis to derive the covariance of the errors on the eigenvectors of $\boldsymbol{R}_{xx}$. In the derivation I shall make the assumption that the noise free data $\boldsymbol{Y}_{m,L}$ is deterministic, and hence $E\{\boldsymbol{R}_{yy}\} = \boldsymbol{R}_{y,y}$. Now, let $\bar{\boldsymbol{R}}_{xx} = E\{\boldsymbol{R}_{xx}\}$ and

$$\boldsymbol{V} = \boldsymbol{R}_{xx} - \bar{\boldsymbol{R}}_{xx}.$$

$\boldsymbol{V}$ is a perturbation on $\bar{\boldsymbol{R}}_{xx}$. From its definition, it is seen that $\boldsymbol{V}$ is Hermitian and $E\{\boldsymbol{V}\} = 0$. Let $\bar{\boldsymbol{R}}_{xx} = \boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{U}^H$ be the EVD of $\bar{\boldsymbol{R}}_{xx}$, $\boldsymbol{u}_v$ be the v-th column of $\boldsymbol{U}$ (the v-th eigenvector of $\bar{\boldsymbol{R}}_{xx}$) and $\Delta\boldsymbol{u}_v$ represent the perturbation on $\boldsymbol{u}_v$ due to $\boldsymbol{V}$. In [96], it has been sown that, for the above model, the first order approximation of $\Delta\boldsymbol{u}_v$, $v \leq d$ (d is the number of signals in the channel) is given by

$$\Delta\boldsymbol{u}_v \approx \sum_{\substack{j=1 \\ j\neq v}}^{mM_L} \frac{\boldsymbol{u}_j^H \boldsymbol{V} \boldsymbol{u}_v}{\sigma_v^2 - \sigma_j^2} \boldsymbol{u}_j.$$

From this, it follows that

$$E\{\Delta\boldsymbol{u}_v \Delta\boldsymbol{u}_w^H\} \approx \sum_{\substack{j=1 \\ j\neq v}}^{mM_L} \sum_{\substack{i=1 \\ i\neq w}}^{mM_L} \frac{E\{\boldsymbol{u}_j^H \boldsymbol{V} \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{V} \boldsymbol{u}_i\}}{(\sigma_v^2 - \sigma_j^2)(\sigma_w^2 - \sigma_i^2)} \boldsymbol{u}_j \boldsymbol{u}_i^H. \tag{A.1}$$

Putting $\boldsymbol{V} = \boldsymbol{R}_{xx} - \bar{\boldsymbol{R}}_{xx}$ into above and noting that, for $x \neq y$, $\boldsymbol{u}_x^H \bar{\boldsymbol{R}}_{xx} \boldsymbol{u}_y = 0$, the above can be re-written as

$$E\{\Delta \boldsymbol{u}_v \Delta \boldsymbol{u}_w^H\} \approx \sum_{\substack{j=1 \\ j \neq v}}^{mM_L} \sum_{\substack{i=1 \\ i \neq w}}^{mM_L} \frac{E\{\boldsymbol{u}_j^H \boldsymbol{R}_{xx} \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{R}_{xx} \boldsymbol{u}_i\}}{(\sigma_v^2 - \sigma_j^2)(\sigma_w^2 - \sigma_i^2)} \boldsymbol{u}_j \boldsymbol{u}_i^H. \quad \text{(A.2)}$$

Let $\boldsymbol{x}_k$ be the k-th column of $\boldsymbol{X}_{m,L}$, then we have

$$\boldsymbol{R}_{xx} = \frac{1}{L(N-m+1)} \sum_{k=1}^{L(N-m+1)} \boldsymbol{x}_k \boldsymbol{x}_k^H.$$

Thus putting this in place of $\boldsymbol{R}_{xx}$ in (A.2), the numerator term is expressed as

$$E\{\boldsymbol{u}_j^H \boldsymbol{R}_{xx} \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{R}_{xx} \boldsymbol{u}_i\} \approx \frac{1}{L^2(N-m+1)^2} \sum_{k,n=1}^{L(N-m+1)} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{x}_k^H \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{x}_n \boldsymbol{x}_n^H \boldsymbol{u}_i\}. \quad \text{(A.3)}$$

It is well known that, for non zero mean Gaussian random variables z_1, z_2, z_3 and z_4,

$$E\{z_1 z_2 z_3 z_4\} = E\{z_1 z_2\} E\{z_3 z_4\} + E\{z_1 z_3\} E\{z_2 z_4\} \\ E\{z_1 z_4\} E\{z_2 z_3\} - 2E\{z_1\} E\{z_2\} E\{z_3\} E\{z_4\}.$$

Using this property and noting $E\{\boldsymbol{x}_k\} = \boldsymbol{y}_k$, we get

$$\begin{aligned} &\sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{x}_k^H \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{x}_n \boldsymbol{x}_n^H \boldsymbol{u}_i\} \\ &\quad = \sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{x}_k^H \boldsymbol{u}_v\} E\{\boldsymbol{u}_w^H \boldsymbol{x}_n \boldsymbol{x}_n^H \boldsymbol{u}_i\} \\ &\quad + \sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{u}_w^H \boldsymbol{x}_n\} E\{\boldsymbol{x}_k^H \boldsymbol{u}_v \boldsymbol{x}_n^H \boldsymbol{u}_i\} \quad \text{(A.4)} \\ &\quad + \sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{x}_n \boldsymbol{u}_i\} E\{\boldsymbol{u}_w^H \boldsymbol{x}_n \boldsymbol{x}_k \boldsymbol{u}_v\} \\ &\quad - 2 \sum_{k,n} \boldsymbol{u}_j^H \boldsymbol{y}_k \boldsymbol{y}_k \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{y}_n \boldsymbol{y}_n \boldsymbol{u}_i. \end{aligned}$$

The first term may be written as $\boldsymbol{u}_j^H E\{\sum_k \boldsymbol{x}_k \boldsymbol{x}_k^H\} \boldsymbol{u}_v \boldsymbol{u}_w^H E\{\sum_n \boldsymbol{x}_n \boldsymbol{x}_n^H\} \boldsymbol{u}_i$. But $E\{\sum_k \boldsymbol{x}_k \boldsymbol{x}_k^H\} = L(N-m+1)\bar{\boldsymbol{R}}_{xx}$. Thus, the first term in (A.4) can be written more compactly as $L^2(N-m+1)^2 \boldsymbol{u}_j^H \bar{\boldsymbol{R}}_{xx} \boldsymbol{u}_v \boldsymbol{u}_w^H \bar{\boldsymbol{R}}_{xx} \boldsymbol{u}_i$. Since $j \neq v$ and $i \neq w$, this term reduces to zero.

Consider the second term. Let $\boldsymbol{w}_k$ be the k-th noise vector (which is assumed to be Gaussian and circulant) such that $\boldsymbol{x}_k = \boldsymbol{y}_k + \boldsymbol{w}_k$. Putting this into the second term, and noting that all the odd number moments in the noise are zero and, for all k, n, $E\{\boldsymbol{w}_k \boldsymbol{w}_n^T\} = 0$, we obtain

$$\sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{u}_w^H \boldsymbol{x}_n\} E\{\boldsymbol{x}_k^H \boldsymbol{u}_v \boldsymbol{x}_n^H \boldsymbol{u}_i = \sum_{k,n} \boldsymbol{u}_j^H \boldsymbol{y}_k \boldsymbol{y}_k \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{y}_n \boldsymbol{y}_n \boldsymbol{u}_i. \quad \text{(A.5)}$$

Re arranging the terms, the above can be compactly written as

$$L^2(N-m+1)^2 \boldsymbol{u}_j^H \bar{\boldsymbol{R}}_{yy} \boldsymbol{u}_v \boldsymbol{u}_w^H \bar{\boldsymbol{R}}_{yy} \boldsymbol{u}_i \quad \text{(A.6)}$$

For $j \neq v$ or $i \neq w$ this vanishes. Note that the fourth term in (A.4) also reduces to (A.6). Thus, it also vanishes. This means that the only remaining term in (A.4) is the third one, hence,

$$\sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{x}_k^H \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{x}_n \boldsymbol{x}_n^H \boldsymbol{u}_i\} = \sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{x}_n \boldsymbol{u}_i\} E\{\boldsymbol{u}_w^H \boldsymbol{x}_n \boldsymbol{x}_k \boldsymbol{u}_v\}.$$

Putting $\boldsymbol{x}_k = \boldsymbol{y}_k + \boldsymbol{w}_k$ in the above, we get sixteen terms. Of these, all the odd order moments in the noise are zero and, since we have assumed circulant noise, $E\{\boldsymbol{w}_k \boldsymbol{w}_n^T\} = 0$, for all n, k. After eliminating all these terms, we obtain

$$\begin{aligned}
&\sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{x}_k \boldsymbol{x}_n \boldsymbol{u}_i\} E\{\boldsymbol{u}_w^H \boldsymbol{x}_n \boldsymbol{x}_k \boldsymbol{u}_v\} \\
&\quad = \sum_{k,n} \boldsymbol{u}_j^H \boldsymbol{y}_k \boldsymbol{y}_n^H \boldsymbol{u}_i \boldsymbol{u}_w^H \boldsymbol{y}_n \boldsymbol{y}_k^H \boldsymbol{u}_v \\
&\quad + \sum_{k,n} \boldsymbol{u}_j^H \boldsymbol{y}_k \boldsymbol{y}_n^H \boldsymbol{u}_i E\{\boldsymbol{u}_w^H \boldsymbol{w}_n \boldsymbol{w}_k^H \boldsymbol{u}_v\} \\
&\quad + \sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{w}_k \boldsymbol{w}_n^H \boldsymbol{u}_i\} \boldsymbol{u}_w^H \boldsymbol{y}_n \boldsymbol{y}_k^H \boldsymbol{u}_v \\
&\quad + \sum_{k,n} E\{\boldsymbol{u}_j^H \boldsymbol{w}_k \boldsymbol{w}_n^H \boldsymbol{u}_i\} E\{\boldsymbol{u}_w^H \boldsymbol{w}_n \boldsymbol{w}_k^H \boldsymbol{u}_v\}
\end{aligned} \quad \text{(A.7)}$$

Note that the first term is equal to (A.6), which is identical to zero. Thus, we are left with the last three terms only. Consider the second term (henceforth denoted by $t2$). Let $\boldsymbol{A}_m \in \mathbb{C}^{mM_L, d}$ be the extended array steering matrix, the d-vector $\boldsymbol{\rho}$ be the complex amplitudes of the d signals and $\boldsymbol{B} := \text{diag}\{\boldsymbol{\rho}\}$. The spatio-temporally smoothed noise free data matrix $\boldsymbol{Y}_{m,L} \in \mathbb{C}^{mM_L, L(N-m+1)}$ can then be expressed as

$$\boldsymbol{Y}_{m,L} = \boldsymbol{A}_m \boldsymbol{B} \left[\begin{array}{cccc} \boldsymbol{F}_s & \boldsymbol{\Theta} \boldsymbol{F}_s & \cdots & \boldsymbol{\Theta}^{L-1} \boldsymbol{F}_s \end{array} \right] =: \boldsymbol{A}_m \boldsymbol{S}$$

For $p, q \in [0, L]$ and $g, h \in [0, N-m]$ let k and n be written as

$$\begin{aligned} k &= p(N-m+1)+k' \\ n &= q(N-m+1)+n'. \end{aligned}$$

Also, let $\boldsymbol{\Phi}^n := \text{diag}\{\phi_i^n\}_{i=1}^d$ and $\boldsymbol{\Theta}^n := \text{diag}\{\theta_i^n\}_{i=1}^d$. Then, the k-th and n-th columns of $\boldsymbol{Y}_{m,L}$, denoted by $\boldsymbol{y}_k$ and $\boldsymbol{y}_n$, are given by

$$\begin{aligned} \boldsymbol{s}_k &= \boldsymbol{A}_m \boldsymbol{\Phi}^{k'} \boldsymbol{\Theta}^p \boldsymbol{\rho} \\ \boldsymbol{s}_n &= \boldsymbol{A}_m \boldsymbol{\Phi}^{n'} \boldsymbol{\Theta}^q \boldsymbol{\rho} \end{aligned}$$

Let $\boldsymbol{D} = \boldsymbol{\rho}\boldsymbol{\rho}^H$, then we have

$$t2 = \sum_{k',p,n',q} \boldsymbol{u}_j^H \boldsymbol{A}_m \boldsymbol{\Phi}^{k'} \boldsymbol{\Theta}^p \boldsymbol{D} \boldsymbol{\Phi}^{-n'} \boldsymbol{\Theta}^{-q} \boldsymbol{A}_m^H E\{\boldsymbol{u}_w^H \boldsymbol{w}_n \boldsymbol{w}_k^H \boldsymbol{u}_v\}. \quad \text{(A.8)}$$

From the spatio-temporal smoothing process we see that the factor

$$E\{\boldsymbol{u}_w^H \boldsymbol{w}_n \boldsymbol{w}_k^H \boldsymbol{u}_v\},$$

is a function of $k-n$ only. More precisely, let $\boldsymbol{Z}^h$ be a Toeplitz matrix with all the elements equal to zero, except for those unity valued entries on the h-th parallel to the diagonal and σ_n^2 be the variance of the noise. Let $L_o = \min(L, M_L)$ and $m_o = \min(m, N-m+1)$, then

$$E\{\boldsymbol{u}_w^H \boldsymbol{w}_n \boldsymbol{w}_k^H \boldsymbol{u}_v\} = \begin{cases} \sigma_n^2 \boldsymbol{Z}^{(k'-n')M_L+(p-q)}, & |k'-n'| < m_o \text{ and } |p-q| < L_o \\ 0, & \text{otherwise.} \end{cases} \quad \text{(A.9)}$$

Making change of variables and, for $r \in [-L_o, L_o]$ and $h =\in [-m_o, m_o]$, setting $n' = k' - h$ and $q = p - r$, (A.8) can be expressed as

$$t2 = \sigma_n^2 \sum_{k',p,r,h} \boldsymbol{u}_j^H \boldsymbol{A}_m \boldsymbol{\Phi}^{k'} \boldsymbol{\Theta}^p \boldsymbol{D} \boldsymbol{\Phi}^{h-k'} \boldsymbol{\Theta}^{r-p} \boldsymbol{A}_m^H \boldsymbol{u}_i \boldsymbol{u}_w^H \boldsymbol{Z}^{hM_L+r} \boldsymbol{u}_v, \quad \text{(A.10)}$$

Rearranging the terms, we obtain

$$t2 = \sigma_n^2 \sum_{r,h} \boldsymbol{u}_w^H \boldsymbol{Z}^{rM_L+h} \boldsymbol{u}_v \boldsymbol{u}_j^H \boldsymbol{A}_m \sum_{k',p} \left(\boldsymbol{\Phi}^{k'} \boldsymbol{\Theta}^p \boldsymbol{D} \boldsymbol{\Phi}^{-k'} \boldsymbol{\Theta}^{-p} \right) \boldsymbol{\Phi}^h \boldsymbol{\Theta}^r \boldsymbol{A}_m^H \boldsymbol{u}_i$$

Note that

$$\sum_{k',p} \boldsymbol{\Phi}^{k'} \boldsymbol{\Theta}^p \boldsymbol{D} \boldsymbol{\Phi}^{-k'} \boldsymbol{\Theta}^{-p} = L(N-m+1) \boldsymbol{R}_{ss},$$

where $\boldsymbol{R}_{ss} \in \mathbb{C}^{d,d}$ is the signal covariance matrix. Thus, we have

$$t2 = L(N-m+1)\sigma_n^2 \sum_{r,h} \boldsymbol{u}_w^H \boldsymbol{Z}^{hM_L+r} \boldsymbol{u}_v \boldsymbol{u}_j^H \boldsymbol{A}_m \boldsymbol{R}_{ss} \boldsymbol{\Phi}^h \boldsymbol{\Theta}^r \boldsymbol{A}_m^H \boldsymbol{u}_i.$$

Since $\boldsymbol{U}_s$ and $\boldsymbol{A}_m$ span the same column space, there exists an invertible matrix $\boldsymbol{T} \in \mathbb{C}^{d,d}$ such that $\boldsymbol{U} = \boldsymbol{A}_m \boldsymbol{T}^{-1}$. Let $\boldsymbol{Q} = \boldsymbol{T}^{-1}$, then noting that $\boldsymbol{R}_{ss} = \boldsymbol{Q}\boldsymbol{\Sigma}^2\boldsymbol{Q}^H$, for $x \leq d$ it follows that,

$$\begin{aligned} \boldsymbol{u}_x \boldsymbol{A}_m \boldsymbol{R}_{ss} &= \bar{\sigma}_x^2 \boldsymbol{q}_x^H \\ \boldsymbol{A}_m^H \boldsymbol{u}_x &= \boldsymbol{t}_x, \end{aligned}$$

where $\bar{\sigma}_x$ is the x-th noise free eigenvalue, and $\boldsymbol{q}_x^H$ and $\boldsymbol{t}_x$, are the x-th row and column of $\boldsymbol{Q}^H$ and $\boldsymbol{T}^H$, respectively. Putting these in to above, we get

$$t2 = L(N-m+1)\bar{\sigma}_j^2 \sigma_n^2 \sum_{r,h} \boldsymbol{u}_w^H \boldsymbol{Z}^{hM_L+r} \boldsymbol{u}_v \boldsymbol{q}_j^H \boldsymbol{\Phi}^h \boldsymbol{\Theta}^r \boldsymbol{t}_i.$$

In a similar way, for the third term in (A.1) (denoted by $t3$), we get

$$t3 = L(N-m+1)\bar{\sigma}_w^2 \sigma_n^2 \sum_{r,h} \boldsymbol{u}_j^H \boldsymbol{Z}^{-hM_L-r} \boldsymbol{u}_i \boldsymbol{q}_w^H \boldsymbol{\Phi}^{-h} \boldsymbol{\Theta}^{-r} \boldsymbol{t}_v,$$

and the fourth term becomes

$$t4 = \sigma_n^4 \sum_k \sum_{r,h} \boldsymbol{u}_w^H \boldsymbol{Z}^{hM_L+r} \boldsymbol{u}_v \boldsymbol{u}_j^H \boldsymbol{Z}^{-hM_L-r} \boldsymbol{u}_i.$$

Since $\boldsymbol{Z}^{hM_L+r}$ is independent of k, this reduces to

$$t4 = L(N-m+1)\sigma_n^4 \sum_{r,h} \boldsymbol{u}_w^H \boldsymbol{Z}^{hM_L+r} \boldsymbol{u}_v \boldsymbol{u}_j^H \boldsymbol{Z}^{-hM_L-r} \boldsymbol{u}_i$$

Combining all the tree terms, we obtain

$$\begin{aligned} E\{\boldsymbol{u}_j^H \boldsymbol{R}_{xx} \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{R}_{xx} \boldsymbol{u}_i\} \approx \\ = \sum_{r=-L_o}^{L_o} \sum_{h=-m_o}^{m_o} \Bigg(& \frac{\bar{\sigma}_j^2 \sigma_n^2}{L(N-m+1)} (\boldsymbol{u}_w^H \boldsymbol{Z}^{rM_L+h} \boldsymbol{u}_v \boldsymbol{q}_j^H \boldsymbol{\Phi}^h \boldsymbol{\Theta}^r \boldsymbol{t}_i) \\ & + \frac{\bar{\sigma}_w^2 \sigma_n^2}{L(N-m+1)} (\boldsymbol{u}_j^H \boldsymbol{Z}^{-rM_L-h} \boldsymbol{u}_i \boldsymbol{q}_w^H \boldsymbol{\Phi}^{-h} \boldsymbol{\Theta}^{-r} \boldsymbol{t}_v) \\ & + \frac{\sigma_n^4}{L(N-m+1)} (\boldsymbol{u}_w^H \boldsymbol{Z}^{rM_L+h} \boldsymbol{u}_v \boldsymbol{u}_j^H \boldsymbol{Z}^{-rM_L-h} \boldsymbol{u}_i) \Bigg) \end{aligned} \quad \text{(A.11)}$$

Putting this into (A.1), and noting that $t2$ is non zero only for $i, j \leq d$, (4.2) follows. This concludes the proof.

A.3 Proof of lemma 4.1

For white noise, $\boldsymbol{Z}^{hM_L+r}$ in (A.11), is non zero only for $r = 0$ and $h = 0$. In this case, $\boldsymbol{Z} = \boldsymbol{I}$ and (A.11) reduces to

$$\begin{aligned} E\{\boldsymbol{u}_j^H \boldsymbol{R}_{xx} \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{R}_{xx} \boldsymbol{u}_i\} \approx \\ &= \frac{\bar{\sigma}_j^2 \sigma_n^2}{L(N-m+1)} (\boldsymbol{u}_w^H \boldsymbol{u}_v \boldsymbol{q}_j^H \boldsymbol{t}_i) \\ &+ \frac{\bar{\sigma}_w^2 \sigma_n^2}{L(N-m+1)} (\boldsymbol{u}_j^H \boldsymbol{u}_i \boldsymbol{q}_w^H \boldsymbol{t}_v) \\ &+ \frac{\sigma_n^4}{L(N-m+1)} (\boldsymbol{u}_w^H \boldsymbol{u}_v \boldsymbol{u}_j^H \boldsymbol{u}_i) \end{aligned} \tag{A.12}$$

Putting this into (A.1) and noting that, for $i, j > d$, the first term is zero; for $i \neq j$, $\boldsymbol{u}_j^H \boldsymbol{u}_i = \boldsymbol{q}_j^H \boldsymbol{t}_i = 0$; and for $v \neq w$, $\boldsymbol{u}_w^H \boldsymbol{u}_v = \boldsymbol{q}_w^H \boldsymbol{t}_v = 0$, we obtain

$$E\{\boldsymbol{u}_j^H \boldsymbol{R}_{xx} \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{R}_{xx} \boldsymbol{u}_i\} = 0$$

for $w \neq v$ or $i \neq j$, and

$$\begin{aligned} E\{\boldsymbol{u}_j^H \boldsymbol{R}_{xx} \boldsymbol{u}_v \boldsymbol{u}_w^H \boldsymbol{R}_{xx} \boldsymbol{u}_i\} \approx \frac{1}{L(N-m+1)} \times \\ \left(\sum_{\substack{j=1 \\ j \neq v}}^{d} \frac{\bar{\sigma}_j^2 \sigma_n^2 + \bar{\sigma}_w^2 \sigma_n^2 + \sigma_n^4}{(\sigma_w^2 - \sigma_j^2)^2} + \sum_{j=d+1}^{mM_L} \frac{\bar{\sigma}_w^2 \sigma_n^2 + \sigma_n^4}{(\sigma_w^2 - \sigma_j^2)^2} \right) \end{aligned} \tag{A.13}$$

for $w = v$ and $i = j$. Since $\sigma_x^2 = \bar{\sigma}_x^2 + \sigma_n^2$, we may write

$$\begin{aligned} \bar{\sigma}_j^2 \sigma_n^2 + \bar{\sigma}_w^2 \sigma_n^2 + \sigma_n^4 &= \sigma_w^2 \sigma_j^2 - \bar{\sigma}_w^2 \bar{\sigma}_j^2 \\ \bar{\sigma}_w^2 \sigma_n^2 + \sigma_n^4 &= \sigma_w^2 \sigma_n^2 \end{aligned}$$

Placing these into (A.13), we obtain (4.3). This concludes the proof.

BIBLIOGRAPHY

[1] T. W. Anderson. Asymptotic theory for principal component analysis. *Annals Maths. Stat.*, 34:122–148, 1963.

[2] Rainer Bachl. The forward-backward averaging technique applied to TLS-ESPRIT processing. *IEEE Trans. SP*, 43:2691–2699, November 1995.

[3] A.J. Barabell. Improving the resolution performance of eigenstructure-based direction finding algorithms. In *Proc. ICASSP'79*, pages 336–339, 1983.

[4] M. S. Bartlett. Smoothing periodograms from time series with continuous spectra. *Nature*, 161:686–687, May 1969.

[5] G. Bienvenu. Influence of the spatial coherence of the background noise on high resolution passive methods. In *Proc. ICASSP'79*, pages 306–309, Washington, DC, April 2-4 1979.

[6] G. Bienvenu and L. Kopp. Adaptivity to background noise spatial coherence for high resolution passive methods. In *Proc. ICASSP'80*, pages 307–310, Denver, CO, 1980.

[7] G. Bienvenu and L. Kopp. Optimality of high resolution array processing using the eigensystem approach. *IEEE Trans. on Acoust. Speech and Signal processing*, ASSP-31:1235–1247, 1989.

[8] Laurens Bierens. ADBF experimental data, March 1998. This document describes the data acquisition scenarios used by the TNO-FEL in the ADBF experiment.

[9] Y. Bresler and A. Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Trans. on Acoust. Speech and Signal processing*, ASSP-34:1081–1089, October 1986.

[10] D. R. Brillinger. *Time Series: Data Analysis and Theory.* New York: Holt, Rhinehart and Winston, 1975.

[11] R. Bro, N. D. Sidiropoulos, and G. B. Giannakis. Optimal joint azimuth-elevation and signal-array response estimation using parallel factor analysis. In *32-nd Asilomar Cnof. on Signals, Systems and Computers*, Monterey, CA, November 1998.

[12] J. Burg. Maximum entropy spectral analysis. In *37-th Meeting Society Exploration Geophysicists*, 1967.

[13] J. Capon. High-resolution frequency-wavenumber spectrum analysis. In *Proc. IEEE*, volume 57, pages 2408–2418, August 1969.

[14] O. Caspary, P. Nus, and T. Cecchin. The source number estimation based on gerschgorin radii. In *ICASSP 98*, pages 1993–1996, Seatle, Washington, May 1998.

[15] J. A. Caszow. A high-resolution direction-of-arrival algorithm for narrow-band coherent and incoherent sources. *IEEE Trans. Acoust., Speech, Signal Processing*, 36:965–979, October 1988.

[16] R. T. Compton. *Adaptive Antennas.* Prentice Hall, 1988.

[17] Lieven de Lathauwer. *Signal Processing Based on Multilinear Algebra.* PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium, 1997.

[18] J. L. Dornstetter and D. Verhulst. The digital cellular sfh 900 system. In *Proc. the IEEE International Conf. on Communications*, pages 36.3.1–36.3.5, June 1986.

[19] B. Friedlander and A. J. Weiss. Eigenstructure methods for direction finding with sensor gain and phase uncertainties. In *ICASSP*, pages 2681–2684, New York City, April 1988.

[20] G. Golub and C. van Loan. An analysis of the total least squares problem. *SIAM Journal Numerical Analysis*, 17(6):883–893, December 1980.

[21] G. Golub and C. van Loan. Total least squares. In *Matrix Computations*, pages 420–425, Boltimore,MD, 1983. Johns Hopkins University Press.

[22] Gene H. Golub and Charles F. van Loan. *Matrix Computations.* North Oxford Academic Ltd, Oxford, England, 1986.

[23] M. Haardt, K. Hüper, J. B. Moore, and A. Nossek. Simultaneous schur decomposition of several matrices to achieve automatic pairing in multidimensional harmonic retrieval problems. In *Signal Processing VIII: Theories and Applications (Proc. of EUSIPCO-96)*, volume I, pages 531–534, Trieste, Italy, September 1996.

[24] M. Haardt and J. A. Nossek. Unitary esprit: How to obtain increased estimation accuracy with a reduced computational burden. *IEEE Trans. Signal Processing*, 43:1232–1242, May 1995.

[25] M. Haardt and J. A. Nossek. Structured leastsquares to improve the performance of ESPRIT-type high-resolution techniques. In *Proc. ICASSP'96*, pages 2805–2808, Atlanta, Georgia USA, May 1996.

[26] M. Haardt and J. A. Nossek. 3-D unitary ESPRIT for joint angle and carrier estimation. In *Proc. ICASSP'97*, pages 255–258, Munich, Germany, April 1997.

[27] Martin Haardt. *Efficient One-, Two-, and Multidimensional High-Resolution Array Signal Processing*. PhD thesis, Munich University of Technology, Munich, Germany, 1997.

[28] Martin Haardt, Michael D. Zoltowski, Cherian P. Mathews, and Josef A. Nossek. 2D unitary ESPRIT for efficient 2D parameter estimation. In *Proc. ICASSP'95*, pages 2096–2099, Detroit, MI USA, May 1995.

[29] Simon Haykin. *An Introduction to Analog and Digital Communications*. John Wiley & Sons, 1986.

[30] R. D. Hill, R. G. Bates, and S. R. Waters. On centrohermitian matrices. *SIAM J. Matrix Anal. Appl.*, 11:128–133, January 1990.

[31] Yingbo Hua and Tapan K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. Signal Processing*, 38:814–824, May 1990.

[32] K. C. Huarng and C. C. Yeh. Adaptive beamforming with conjugate symmetric weights. *IEEE Trans. Antennas and Propagation*, 39:926–932, July 1991.

[33] K. C. Huarng and C. C. Yeh. A unitary transformation method for angle of arrival estimation. *IEEE Trans. SP*, 39:975–977, April 1991.

[34] J. E. Hudson. *Adaptive Array Principles*. Institution of Electrical Engineers, London, 1981.

[35] S. Van Huffel and J. Vandewalle. The use of total least squares techniques for identification and parameter estimation. *IFAC, Identification Syst. Parameter Estimation*, 2-c:1167–1172, 1985.

[36] D. J. Jeffries and D. R. Farrier. Asymptotic results for eigenvector methods. *Proc. Inst. Elect. Eng.*, 132:589–594, June 1985.

[37] J.Pierre and M. Kaveh. Experimental performance of calibration and direction finding algorithms. In *ICASSP*, pages 1365–1368, Toronto, Canada, 1991.

[38] Mostafa Kaveh and Arthur J. barabell. The statistical performance of MUSIC and minimum-norm algorithms in resolving plane waves. *IEEE Trans. on Acoust. Speech and Signal processing*, ASSP-34:331–341, April 1986.

[39] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, NJ, 1993.

[40] K. Konstantinides and K. Yao. Statistical analysis of effective singular values in matrix rank determination. *IEEE Trans. ASSP*, 36:757–763, May 1988.

[41] Hamid Krim and Mats Viberg. Two decades of array signal processing research. *IEEE Signal Processing Magazine*, pages 67–94, July 1996.

[42] S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao. State space and SVD based approximation methods for the harmonic retrieval problem. In *Jornal of Opt. Soc. Amer.*, volume 73, pages 1799–1811, December 1983.

[43] A. Lee. Centro-hermitian and skew-centrohermitian matrices. *Linear Algebra and its Applications*, 29:205–210, 1980.

[44] Harry Lee and Fu Li. An eigenvector technique for detecting the number of emitters in a cluster. *IEEE Trans. SP*, 42:2380–2388, September 1994.

[45] Aweke N. Lemma. Joint azimuth and elevation angles, and frequency estimation. Technical Report 2/1997, Delft University of Technology, Delft, The Netherlands, September 1997.

[46] Aweke N. Lemma, Ed F. Deprettere, and Alle-Jan Van der Veen. Experimental analysis of antenna coupling for high-resolugion doa estimation algorithms. In *Proc. IEEE Workshop on Signal Proc. Advances in Wireless Com. (SPAWC99)*, Annapolis USA, May 1999.

[47] Aweke N. Lemma, A.-J. van der Veen, and Ed F. Deprettere. Joint angle-frequency estimation using multi-resolution ESPRIT. In *Proc. ICASSP'98*, Seattle, Washington, May 1998.

[48] Aweke N. Lemma, A.-J. van der Veen, and Ed F. Deprettere. On the multi-resolution ESPRIT algorithm. In *9-th SP Workshop on SSAP*, Portland, Oregan, USA, September 1998.

[49] Aweke N. Lemma, A.-J. van der Veen, and Ed F. Deprettere. Analysis of joint angle-frequency estimation using ESPRIT. *Submit. IEEE Trans. Signal Processing*, April 1999.

[50] Aweke N. Lemma, A.-J. van der Veen, and Ed F. Deprettere. Multi-resolution ESPRIT algorithm. *IEEE Trans. Signal Processing*, 47, June 1999.

[51] Aweke N. Lemma, A.-J. van der Veen, and Ed F. Deprettere. Slow frequency hopping signals and the jafe algorithm. *Submit. IEEE Trans. Signal Processing*, August 1999.

[52] D. A. Linebarger, R. D. deGroat, and E. M. Dowling. Efficient direction finding methods employing forward/backward averaging. *IEEE Trans. SP*, 42:2136–2145, August 1994.

[53] Hui Liu and Guanghan Xu. Closed-form blind symbol estimation in digital communications. *IEEE Trans. Signal Processing*, 43:2714–2723, November 1995.

[54] V. H. MacDonald and P. M. Schultheiss. Optimum passive bearing estimation in a spatially incoherent noise environment. *Jour. Acoust. Society Am.*, 46(1):37–43, 1969.

[55] C. P. Mathews, M. Haardt, and M. D. Zoltowski. Performance analysis of closed-form, ESPRIT based 2-D angle estimator for rectangular arrays. *IEEE Signal Processing Letters*, 3:124–126, April 1996.

[56] C. P. Mathews and M. D. Zoltowski. Eigen structure techniques for 2-D angle estimation with uniform circular arrays. *IEEE Trans. Signal Processing*, 42:2395–2407, Sept. 1994.

[57] B. Ottersten, R. Roy, and T. Kailath. Signal waveform estimation in sensor array processing. In *Proc. 23rd Asilomar Conf. on Signals, Systems and Computers*, pages 787–791, Pacific Grove, CA, November 1989.

[58] Björn Ottersten, Mats Viberg, and Thomas Kailath. Performance analysis of the total least squares ESPRIT algorithm. *IEEE Trans. Signal Processing*, 39:1122–1135, May 1991.

[59] A. Paulraj and T. Kailath. Direction of arrival estimation by eigenstructure method with unknown sensor gain and phase. In *ICASSP*, pages 640–643, Tempa, Florida, March 1985.

[60] A. Paulraj, R. Roy, and T. Kailath. Estimation of signal parameters via rotational invariance techniques-ESPRIT. In *19th Asilomar Conf. on Signals, Systems and Computers*, pages 83–89, Pacific Grove, CA, July 1985.

[61] S. U. Pillai and B. H. Kwon. Forward/backward spatial smoothing techniques for coherent signal identification. *IEEE Trans. Acoustics, Speech and Signal Processing*, ASSP-37:8–15, January 1989.

[62] S. U. Pillai and B. H. Kwon. Forward/backward spatial smoothing techniques for coherent signal identification. *IEEE Trans. ASSP*, 37:8–15, January 1989.

[63] V. F. Pisarenko. The retrieval of harmonics from a covariance function. *Geophysics J. Roy Astron. Society*, 33:347–366, 1973.

[64] B. Rao and K. V. S. Hari. Weighted subspace methods and spatial smoothing: Analysis and comparison. *IEEE Trans. SP*, 41:788–803, February 1993.

[65] Bhaskar D. Rao and K. V. S. Hari. Performance analysis of ESPRIT and TAM in determining the direction of arrival of plane waves in noise. *IEEE Trans. Acous. Speech and Signal Processing*, 37:1990–1995, December 1989.

[66] D.V. Bhaskar Rao and K.V.S. Hari. Performance analysis of root-MUSIC. *IEEE Trans. on Acoust. Speech and Signal processing*, 37:1939–1949, December 1989.

[67] R. Roy and T. Kailath. ESPRIT and total least squares. In *Proc. 21st Asilomar Conf. on Signals, Systems and Computers*, pages 297–301, November 1987.

[68] R. Roy and T. Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. In L. Auslander, F. A. Grünbaum, J. W. Helton, T. Kailath, P. Khargonekar, and S. Mitter, editors, *Signal Processing part II: Control Theory and Applications*, pages 369–411. Springer-verlag, 1990.

[69] R. H. Roy, A. Paulraj, and T. Kailath. A subspace rotation approach to estimation of parameters of cisoids in noise. *IEEE Trans. Acoustics, Speech and Signal Processing*, 34:1340–1342, Oct. 1986.

[70] R.H. Roy. *ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques*. PhD thesis, Stanford University, Stanford, CA, 1987.

[71] Richard Roy and Thomas Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. on Acoust. Speech and Signal processing*, ASSP-37:984–995, July 1989.

[72] R.O. Schmidt. Multiple emitter location and signal parameters estimation. In *Proc. RADC Spectrum Estimation Workshop*, Griffis AFBS, NY, October 1979.

[73] R.O. Schmidt. *A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation*. PhD thesis, Stanford University, Stanford, CA, 1981.

[74] F. C. Schweppe. Sensor array data processing for multiple signal sources. *IEEE Trans. on IT*, 14:294–305, 1968.

[75] C.M.S. See. A method for array calibration in parametric sensor array processing. In *ICCS*, pages 915–919, Singapore, 1994.

[76] C.M.S. See. Sensor array calibration in the presence of mutual coupling, gain and phase mismatch. *Electronics Letters*, 30(5):373–374, March 1994.

[77] Abhijit A. Shah and Donald W. Tufts. Determination of the dimension of a signal subspace from short data records. *IEEE Trans. SP*, 42:2531–2535, September 1994.

[78] T. J. Shan, M. Wax, and T. Kailath. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Trans. Acoustics, Speech and Signal Processing*, ASSP-33:806–811, August 1985.

[79] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and cramer-rao bound. *IEEE Trans. on Acoust. Speech and Signal processing*, ASSP-37:720–741, May 1989.

[80] P. Stoica and A. Nehorai. Performance comparison of subspace rotation and MUSIC methods for direction estimation. *IEEE Tran. Signal Processing*, 39:446–453, February 1991.

[81] P. Stoica and K.C. Sharman. Maximum likelihood methods for direction-of-arrival estimation. *IEEE Trans. on Acoust. Speech and Signal processing*, ASSP-38:1132–1143, July 1990.

[82] A. Lee Swindlehurst and T. Kailath. A performance analysis of subspace-based methods in the presence of model errors, part I: The MUSIC algorithm. *IEEE Trans. SP*, 40:1758–1774, July 1992.

[83] A. Lee Swindlehurst and T. Kailath. Algorithms for azimuth/elevation direction finding using regular array geometries. *IEEE Trans. AES*, 29:145–156, January 1993.

[84] A. Lee Swindlehurst, Björn Ottersten, Richard Roy, and Thomas Kailath. Multiple invariance ESPRIT. *IEEE Trans. Signal Processing*, 40:867–881, April 1992.

[85] A.-J. van der Veen, E.F. Deprettere, and A. Lee Swindlehurst. Subspace-based signal analysis using singular value decomposition. *Proceedings of the IEEE*, 81:1277–1308, September 1993.

[86] A.-J. van der Veen, P.B. Ober, and E.F. Deprettere. Azimuth and elevation computation in high resolution DOA estimation. *IEEE Trans. Signal Processing*, 40:1828–1832, July 1992.

[87] A.-J. van der Veen and A. Paulraj. An analytical constant modulus algorithm. *IEEE Trans. Signal Processing*, 44:1136–1155, May 1996.

[88] A.-J. van der Veen, Michaela C. Vanderveen, and A. Paulraj. Joint angle and delay estimation using shift-invariance techniques. *IEEE Signal Processing Letters*, 4(5):142–145, May 1997.

[89] A.J van der Veen and E.F. Deprettere. Parallel VLSI matrix pencil algorithm for high resolution direction finding. *IEEE Trans. Signal Processing*, 39:383–394, February 1991.

[90] A.J. van der Veen, M.C. Vanderveen, and A. Paulraj. Joint angle and delay estimation using shift-invariance techniques. *IEEE Tr. Signal Processing*, 46(2):405–418, February 1998.

[91] Mats Viberg and Björn Ottersten. Sensor array processing based on subspace fitting. *IEEE Trans. Signal Processing*, 39:1110–1121, May 1991.

[92] Mats Viberg, Björn Ottersten, and Thomas Kailath. Detection and estimation in sensor arrays using weighted subspace fitting. *IEEE Trans. Signal Processing*, 39:2436–2449, November 1991.

[93] Mats Viberg, Björn Ottersten, and Arye Nehorai. Performance analysis of direction finding with large arrays and finite data. *IEEE Trans. Signal Processing*, 43:469–477, February 1995.

[94] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. ASSP*, 33:387–392, April 1985.

[95] A. J. Weiss and B. Friedlander. Direction finding in the presence of mutual coupling. In *Conf. 22nd Asilomar Conf. on Signals, Syst. Comput.*, pages 598–602, Pacific Grove, CA, November 1988.

[96] J. H. Wilkinson. *The Algebraic Eigenvalue Problem.* Oxford University Press, London, 1965.

[97] K. T. Wong and M. D. Zoltowski. Closed-form multi-dimensional multi-invariance ESPRIT. In *Proc. ICASSP'97*, pages 3489–3493, Munich, FRG, April 1997.

[98] H.-T. Wu, J.-F. Yang, and F.-K. Chen. Source number estimators using transformed gerschgorin radii. *IEEE Trans. SP*, 43:1325–1330, June 1995.

[99] G. Xu, R. H. Roy, and T. Kailath. Detection of number of sources via exploitation of centro-symmetry property. *IEEE Trans. Signal Processing*, 42:102–112, January 1994.

[100] Jiankan Yang and A. Lee Swindlehurst. The effects of array calibration errors on DF-based signal copy performance. *IEEE Trans. SP*, 43:2724–2732, November 1995.

[101] M. Zatman. How narrow is narrowband? In *IEEE Proc. Radar, Sonar, Navig.*, volume 145, pages 85–91, April 1998.

[102] M.D Zoltowski, M Haardt, and C.P. Mathews. Closed-form 2-D angle estimation with rectangular arrays in element space or beam space via unitary esprit. *IEEE Trans. Signal Processing*, 44:316–328, February 1996.

[103] M.D Zoltowski, G. M. Kautz, and S. D. Silverstein. Beamspace root-MUSIC. *IEEE Trans. Signal Processing*, 41:344–364, January 1993.

[104] Michael D. Zoltowski and Cherian P. Mathews. Real-time frequency and 2-D angle estimation with sub-nyquist spatio-temporal sampling. *IEEE Trans. Signal Processing*, SP-42:2781–2794, October 1994.

[105] Michael D. Zoltowski and Demosthenis Stavrinides. Sensor array signal processing via a procrustes rotations based eigenanalysis of the ESPRIT data pencil. *IEEE Trans. Signal Processing*, 37:832–861, June 1989.

SAMENVATTING

Signalen komen voor in ons dagelijkse leven in verschillende vormen. De veelheid van de verschillende facetten daarbij is zo overweldigend, dat velen meestal onopgemerkt blijven. In bijna alle gevallen wordt de nuttige informatie omvat door signaalparameters zoals frequentie, fase, amplitude, positie, enz. Om op een toereikende manier de nuttige informatie te extraheren, moet men daarom in staat zijn deze parameters te schatten van gemeten (verkregen) gegevens.

In dit proefschrift behandel ik schattingen van aankomstrichtingen (DOAs) en centrale frequenties van smalbandige lange-afstand radiosignalen, op basis van gemeten gegevens verzameld door middel van sensor arrays. Dit probleem is belangrijk bij verschillende toepassingen, zoals radioastronomie, communicatie, radar, sonar, akoestische systemen, enz. De gebruikte oplossingsmethoden worden gezamenlijk gerefereerd als hoekfrequentie schatting (JAFE) algoritmen.

De oudste signaalparameter schattingstechniek dateert van de tweede wereldoorlog en hield niets anders in dan een Fourier spectrumanalyse op gegevens verzameld door sensor arrays. Gedurende de laatste tientallen jaren zijn verschillende hoge-resolutie signaalparameter schattingstechnieken naar voren gekomen. Van deze technieken is het ESPRIT algoritme (schatting van signaalparameters via een rotatie-invariante techniek) gekozen als onderliggende oplossingsmethode voor dit werk.

Het ESPRIT algoritme valt in een klasse van schatters aangeduid met "subspace fitting" methoden. Subspace fitting methoden, dus ook het ESPRIT algoritme, schatten signaalparameters door de gegevens te fitten op een onderliggend geparameteriseerd model. Sinds de ontwikkeling van het ESPRIT algoritme in 1985, is het één van de belangrijkste parameter schattingsalgoritmen geworden. De belangrijkste reden hiervoor is dat het algoritme rekentechnisch veel aantrekkelijker is vergeleken met andere hoge-resolutie DOA schattingstechnieken.

Een matrix met gegevens verzameld van een geometrisch goed gedefinieerde antenne-array die smalbandige lange-afstand signalen opvangt, heeft een rijke structuur. In JAFE wordt de al bestaande structuur verder verrijkt door een aantal "data-stacking" technieken toe te passen. De "data-stacking" technieken worden gekarakteriseerd door een aantal parameters. De "data-stacking" technieken beïnvloeden het aantal te onderscheiden bronnen en de prestaties van het algoritme. Gebruik makende van perturbatie analyse leid ik optimale waarden af voor deze parameters. Ik loon aan dat het algoritme alleen de theoretisch bepaalde fout haalt wanneer de "data-stacking" parameters dicht bij hun optimale waarde worden gekozen. Het kiezen van optimale "data-stacking" parameters kan echter beperkt worden door het rekenwerk.

Om dit complexiteit-nauwkeurigheid uitwisselingsprobleem op te lossen heb ik het zogeheten multi-resolutie ESPRIT (MR-ESPRIT) algoritme geïmplementeerd. De essentie van het ESPRIT algoritme is dat het signaal parameters schat door het benutten van structuren in de matrix met gegevens afkomstig van de sensor arrays. In de meeste gevallen bevatten de array geometrieën redundanties, of kunnen ontworpen worden zodanig dat ze redundanties bevatten. In het MR-ESPRIT algoritme worden deze redundanties benut om de nauwkeurigheid van de parameterschatting te vergroten. De gedragsanalyse levert ons grenzen voor de te verkrijgen nauwkeurigheid.

Een manier om de robuustheid van een algoritme te evalueren is door zijn gedrag te evalueren wanneer het veronderstelde model en het werkelijke gedrag niet met elkaar overeenstemmen. Met dit doel pas ik het JAFE algoritme toe op gevallen waarbij de werkelijke gegevensstructuur sterk afwijkt van het veronderstelde model. Het algoritme kan bijvoorbeeld op elegante wijze gebruikt worden om "frequency hopping" bronnen (bronnen met een variabele frequentie) te detecteren en te volgen, en om de bandbreedte te schatten van gemoduleerde signalen.

Het ontwikkelde algoritme wordt verder gevalideerd, gebruik makende van praktische meetgegevens. De experimentele resultaten tonen aan dat bij het afleiden van het model het belangrijk is om op juiste wijze rekening te houden met het niet-ideale array en front-end systeemgedrag. Dit omvat o.a. antenne koppeling, fase jitter in analoog naar digitale omzetters, en niet-uniforme antenne versterking.

ABOUT THE AUTHOR

Aweke N. Lemma was born in Arba Minch, Ethiopia, on September 7, 1965. He received the B.Sc. degree, with a first class honor, from Addis Ababa University (AAU), Addis Ababa, Ethiopia in 1988, the M.Sc. degree with great distinction from the Eindhoven University of Technology (EUT), Eindhven, The Netherlands, in 1993 and the Chartered Designer Degree from the Delft University of Technology (DUT), Delft, The Netherlands, in 1996.

In 1988, he became a Graduate Assistant and a Lecturer at the AAU, where he has worked for four years in the department of Electrical Engineering. In 1994, he became a research assistant at the DUT, where he has been conducting researches in digital signal processing. Particularly, in time-varying filter banks design, system modeling using wavelet networks, low bit-rate speech coding, and statistical and array signal processing.

Aweke Negash Lemma is married to Gelila challa and has two children, Bethel and Jonathan.