Pontifícia Universidade Católica
do Rio de Janeiro

Silvio Fernando Bernardes Pinto

High-Resolution Direction Finding Techniques Exploiting Prior Knowledge

Doctoral Thesis

Thesis presented to the Programa de Pós–graduação em Engenharia Elétrica of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Engenharia Elétrica.

Advisor: Rodrigo C. de Lamare

Rio de Janeiro
May 2018

Pontifícia Universidade Católica
do Rio de Janeiro

Silvio Fernando Bernardes Pinto

High-Resolution Direction Finding Techniques Exploiting Prior Knowledge

Thesis presented to the Programa de Pós–graduação em Engenharia Elétrica of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Engenharia Elétrica. Approved by the undersigned Examination Committee.

Rodrigo C. de Lamare
Advisor
Center for Telecommunications Studies (CETUC) – PUC-Rio

Prof. Mariane Rembold Petraglia
Federal University of Rio de Janeiro – UFRJ

Prof. Tadeu Nagashima Ferreira
Fluminense Federal University – UFF

Prof. Lukas Landau
Pontifical Catholic University of Rio de Janeiro – PUC-Rio

Prof. Alan Conci Kubrusly
Pontifical Catholic University of Rio de Janeiro – PUC-Rio

Prof. Gláucio Lima Siqueira
Pontifical Catholic University of Rio de Janeiro – PUC-Rio

Rio de Janeiro, May the 18th, 2018

Silvio Fernando Bernardes Pinto

Majored in engineer by the Catholic University of Petropolis, Rio de Janeiro, Brazil

Bibliographic data

Bernardes Pinto, Silvio Fernando

High-Resolution Direction Finding Techniques Exploiting Prior Knowledge / Silvio Fernando Bernardes Pinto; advisor: Rodrigo C. de Lamare. – Rio de janeiro: PUC-Rio, Departamento de Center for Telecommunications Studies (CETUC) , 2018.

v., 125 f: il. color. ; 30 cm

Tese (doutorado) - Pontifical Catholic University of Rio de Janeiro, Departamento de Center for Telecommunications Studies (CETUC) .

Inclui bibliografia

1. Electrical Engineering – Teses. 2. Electrical Engineering – Teses. 3. Estimação de parâmetros em alta resolução ;. 4. Estimação de direção;. 5. Técnicas auxiliadas por conhecimento ;. I. Pontifical Catholic University of Rio de Janeiro. Departamento de Center for Telecommunications Studies (CETUC) . II. Título.

CDD: 620.11

To my parents, for their support
and encouragement.

Acknowledgments

I would like to first thank my advisor ...

Then I wish to thank ...

Abstract

Bernardes Pinto, Silvio Fernando; (Advisor). **High-Resolution Direction Finding Techniques Exploiting Prior Knowledge**. Rio de Janeiro, 2018. 125p. Tese de doutorado – Departamento de Center for Telecommunications Studies (CETUC) , Pontifical Catholic University of Rio de Janeiro.

Keywords

High-resolution parameter estimation ; Direction finding; Knowledge-aided techniques;

Resumo

Bernardes Pinto, Silvio Fernando; . **High-resolution Direction Finding Techniques exploiting Prior Knowledge for Closely-spaced Sources**. Rio de Janeiro, 2018. 125p. Tese de Doutorado – Departamento de Center for Telecommunications Studies (CETUC) , Pontifical Catholic University of Rio de Janeiro.

Palavras-chave

Estimação de parâmetros em alta resolução ; Estimação de direção; Técnicas auxiliadas por conhecimento ;

Table of contents

List of figures

List of tables

, .

List of Abreviations

AIC – Akaike Information Theoretic Criterion
AVF – Orthogonal Auxiliary-Vector Filtering
CG – Conjugate Gradient
CMV – Constrained Minimum Variance
CRLB – Cramer-Rao Lower Bound
DOA – Direction of Arrival
DOF – Degrees of Freedom
ESPRIT – Estimation of Signal Parameters via Rotational Invariance Technique
EVD – Eigenvalue Decomposition
FBSS, FB – Forward-Backward Spatial Smoothing
IESPRIT – Iterative ESPRIT
KA – Knowledge-Aided
KA-CC – Knowledge-Aided obtained by Convex Combination
KA-GLC – Knowledge-Aided obtained by General Linear Combination
KR – Khatri-Rao Product
LMS – Least-Mean Squares
MDL – Minimum Description Length Criterion
MIMO – Multiple Input and Multiple Output
ML – Maximum Likelihood
MSE – Minimum Squared Error
MS-KAI – Multi-Step Knowledge-Aided Iterative MUSIC – Multiple Signal Classification
MVDR – Minimum-Variance Distortionless Response
MRA – Minimum Redundancy Array
PR – Probability of Resolution
RLS – Recursive Least Squares
RMSE – Root Mean Squared Error
SMLOF – Stochastic Maximum Likelihood Objective Function
SNR – Signal-to-Noise Ratio
STAP – Space-Time Adaptive Processing
SVD – Singular Value Decomposition

TS – Two-Step
TS-KAI – Two-Step Knowledge-Aided Iterative
UCA – Uniform Circular Array
ULA – Uniform Linear Array
UPA – Uniform Planar Array

1

Introduction

1.1

Overview

As an active area of research in the broad field of signal processing, array signal processing focuses on the problem of estimating signal parameters from data collected over the spatial aperture of an array of sensors, in which the sensors are placed at distinct spatial locations. The estimation task is usually associated with the extraction of desired information from impinging signals in the presence of noise and interference. The sensor array deals with the estimation problem by exploiting the spatial separation of the sensor elements to capture the propagating wavefronts, which emanate from energy-radiating sources. Common signal parameters of interest to be estimated are the signal content itself, the directions of arrival of the signals, and their power. To obtain this information, the sensor array data are processed using statistical and adaptive signal processing techniques. These techniques include parameter estimation and adaptive filtering applied to array signal processing. The fundamental set of principles and techniques for sensor array signal processing is applicable in many areas [1, 2], including wireless communications, radar, sonar, biomedicine, seismology and astronomy.

Two of the most relevant topics within array signal processing are beamforming and direction of arrival (DOA) estimation [1], which present unavoidable challenges when designing wireless communications systems. Both tasks start with the records of radiating wavefronts impinging on sensor arrays at a given instant to form an observation vector which is used to build a sample covariance matrix which becomes the basis of the processes.

In a straightforward way, the first topic, i.e., beamforming or spatial filtering, which is the first topic above mentioned, can be defined as the process of properly weighting signals from a particular direction for emphasizing them and attenuating the interferences, which can be done differently. In the former

approach, the weighting applied to the received signal at each sensor element is fixed and data-independent. In the latter, termed adaptive beamforming, that weighting is continuously adapted to track changes in the system and rejects interference. The adaptive algorithms used to adjust the weights for each sensor element are designed by optimizing certain criteria according to the given properties. One of the most relevant design criteria in practice is the constrained minimum variance (CMV) approach [4], which only involves the knowledge of the array geometry and the angle of the desired signal. The CMV optimality criterion minimizes the total beamformer output power while constraining the array response in the direction of the desired signal to be constant. Due to its simplicity and effectiveness much effort has been devoted over the past few decades to devise efficient adaptive algorithms in order to realize a practical beamformer design [1, 3, 5]. In the class of adaptive algorithms, the least-mean squares (LMS) method [5] as a representative of the low-complexity stochastic gradient techniques makes use of gradient vectors for the iterative computation of the weights and yields an acceptable performance in many applications. However, its efficacy strongly depends on the step size and the eigenvalue spread of the covariance matrix, resulting in an insufficient convergence performance for certain scenarios [3]. An alternative method is the recursive least squares (RLS) algorithm [5], which is independent of the eigenvalue spread and thus achieves fast convergence speed. However, its main drawbacks are numerical instability and a relatively high complexity.

The DOA estimation, the latter of the mentioned most relevant topics within array signal processing, has the purpose of determining the angle of arrival of a given spatially propagating signal relative to the antenna array. To this end, the spatial separation of multiple sensor elements is exploited to obtain the location of the energy-radiating source. The result of the estimation procedure is subsequently used for the beamforming to steer the beam towards this specific direction, in order to capture or radiate maximal power. With the field of applications involving DOA estimation constantly expanding, numerous direction finding techniques have been devised over the past few decades [1, 6]. The most well-known parameter estimation strategies discussed here can be classified into three main categories, namely conventional [7], subspace-based [8–11, 13, 15], and maximum likelihood (ML) methods [14]. The concept of the conventional DOA estimation algorithms relies on the beamforming principle. These techniques successively steer the main beam in all possible look directions and measure the output power [10], which is recorded in the form of a pseudo spectrum over the angle range. The largest peaks in the pseudo spectrum are associated with the DOA estimates.

The most prominent approach within this class is Capon's method [7] based on the CMV criterion [4]. It minimizes the power induced by interfering signals and noise while keeping the gain towards the look direction fixed. Although the implementation of the conventional techniques is simple, they suffer from lack of angular resolution and demand a large number of sensors to improve their accuracies. The class of subspace-based methods exploits a spectral decomposition of the covariance matrix to achieve high-resolution DOA estimates. Among the most relevant techniques are the multiple signal classification (MUSIC) [8], its extension Root-MUSIC [9], the estimation of signal parameters via rotational invariance techniques (ESPRIT) [10], its enhancement, namely Unitary ESPRIT [11, 12], the auxiliary vector filtering (AVF) [13] and conjugate gradient (CG) [15, 16] algorithms developed for direction finding. The MUSIC-type and the ESPRIT-type algorithms exploit the eigen-structure of the covariance matrix, allowing a decomposition of the observation space into a signal subspace and a complementary noise subspace. Specifically, MUSIC scans over the possible angle range and makes use of the orthogonality of the subspaces to obtain a pseudo spectrum with increased resolution. Its extension, termed Root-MUSIC, when applied to uniform linear arrays avoids the exhaustive search for peaks by applying a polynomial rooting technique. The ESPRIT-type algorithms avoid the exhaustive peak search by dividing the sensor array into two identical subarrays and benefit from the uniform displacement of the subarrays. The Unitary ESPRIT uses the fact that the phase factors representing the displacement lie on the unit circle. More recent algorithms like AVF and CG iteratively generate an extended non-eigen-based signal subspace containing the true signal subspace and the scanning vector itself, termed extended Krylov signal subspace. The DOA estimates are determined by the search for the collapse of the extended signal subspace as the scanning vector belongs to it. While the AVF algorithm adopts auxiliary vectors to form the extended signal subspace, the CG method applies residual vectors and can be considered as an extension of the AVF technique. Both approaches provide high-resolution estimates for closely spaced sources at a low signal-to-noise ratio, and a small sample size. ML-type methods, which were some of the first techniques developed for DOA estimation, are based on a parametric approach. They effectively exploit the underlying data model, resulting in sufficiently high accuracy that is superior to the conventional and subspace-based methods, especially in low SNR conditions, or when the number of signal samples or data records are small. However, the efficiency comes at the expense of the computational intensity as a multidimensional search is required, which makes ML-type

methods less attractive than subspace-based algorithms. An iterative approach to limiting the computational effort is the alternating projection technique [14], which transforms the optimization problem into a sequence of one-dimensional optimization problems. Non-linear arrays is a current research field dealing with structures combining two or more ULAs with increasing intersensor spacing for boosting the $M-1$ number of sources that can be resolved with a M element ULA using traditional subspace based methods like MUSIC. One of these structures, termed nested array, makes possible the increase of the degrees of freedom of ULA subspace-based methods from $\mathcal{O}(N)$ physical sensors to $\mathcal{O}(N)^2$. However its initial formulation and later related works are restricted to uncorrelated sources. The inherent saving of sensors at the expense of significant number of samples will be exploited by our MS-KAI-MUSIC method to resolve closely-spaced sources.

1.2 Motivation

As discussed in the previous section, the large computational effort demanded to ensure ML-type methods' better accuracy than conventional and subspace-based ones made the former less popular than the latter. Conventional techniques like Capon's method and its extension root-Capon are subjected to lack of angular resolution and demand a large number of sensors to reach higher resolution. The MUSIC-type techniques, whose variance approaches the CRB for uncorrelated signals, is one of the subspace-based methods that yields high-resolution DOA estimates, however they require an accurate hard peak search, which relies on the size of the search step. Despite its extension Root-MUSIC avoids the costly peak search by using a polynomial rooting technique, both techniques suffer from low levels of angular resolution for closely spaced sources at low signal-to-noise ratios, and at small sample sizes. The ESPRIT-type and its unitary version, also classified as high-resolution subspace methods, take advantage of centro-symmetric array configurations. They prevent long peak searches and are computationally more efficient than MUSIC and Conventional types, however they also deal with loss of resolution as a result of the closeness of the sources. Despite their better performance than conventional and eigenstructure techniques like Capon, MUSIC, ESPRIT and their variations, CG and AVF algorithms, which are based on non-eigenvector bases, are handicapped by computational complexity like in the first two preceding methods. While Capon and MUSIC require long peak searches, Krylov signal subspace for AVF and CG algorithms is built for each search angle. All parame-

ter estimation strategies mentioned previously have drawbacks that constitute a fertile area for research and innovation. Two of them lead to advancements to be pursued. The first is the attenuation of the computational complexity of algorithms, mainly for high-resolution direction finding tasks. The second one is the enhancement of the accuracy of high-resolution direction finding techniques by exploiting prior knowledge, which is the subject of this thesis.

1.3 Contributions

It is known that most of conventional methods for direction finding suffer from poor accuracy when subjected to scenarios characterized by closely spaced sources and or correlated sources. In the last years, some approaches to overcome this lack of accuracy have exploited the prior knowledge of signal directions coming from static users and base stations. However, this concept is restricted to available known directions of arrival. In this context, the main contributions of this thesis can be summarized as follows:

- A new concept of a priori knowledge applied to direction finding, which replaces the traditional available known DOAs so far employed with previous estimates obtained on line. This idea was incorporated to the proposed algorithms and their extensions. As expected, the achieved accuracy of the method is dependent on the accuracy of the initial estimates.
- A new approach to estimating the covariance matrix by a refinement process that iteratively reduces its by-products, which occur in the finite sample region. This concept is complemented by a reshaped covariance matrix analysis, which shows that at the earliest iteration the MSE of the data covariance matrix free of these side effects is less than or equal to the MSE of the original one.
- Formulation of a new method, termed multi-step knowledge-aided iterative (MS-KAI), for increasing the accuracy of existent algorithms based on the combination of the previous concepts. This method, initially employed with ULAs to collect impinging signals to be processed by ESPRIT and CG algorithms, was extended to non-uniform linear arrays of the two-level nested class processed by the MUSIC algorithm. The method can be further extended to other types of non-linear arrays.

  MS-KAI-ESPRIT algorithm performs well when tested under multiple

uncorrelated sources and small number of samples, and despite the inherent degradation in its accuracy resulting from using highly correlated sources, its performance can also be considered good under these conditions.

MS-KAI-CG and its version equipped with forward-backward averaging (MS-KAI-CG-FB) show good performance when applied to *two* uncorrelated and *two* strongly correlated sources, respectively. To this end, both versions require a sufficient number of samples. Preliminary studies have shown that for a number of sources superior to *two*, their performances are not satisfactory, which is a consequence of poor effectiveness of the original CG algorithm in providing MS-KAI-CG and MS-KAI-CG-FB with initial accurate estimates.

MS-KAI-MUSIC applied to a two-level nested array under a scenario of two uncorrelated closely-spaced sources provides a significant gain in terms of probability of resolution when compared to the original MUSIC and the original Nested-MUSIC. In terms of RMSE, its gain is marginal but consistent. In summary, the results lead us to conclude that the proposed techniques have excellent potential for applications with sufficiently large data records in large-scale antenna systems, radar and other large sensor arrays. The heavy computational burden inherent to the MS-KAI method, which is a consequence of nested loops and multiple matrices products needed, is a point to be considered in future works.

1.4 Thesis Outline

This thesis is organized as follows:

Chapter 2

We start this chapter by formulating the problem of source localization using a system model based on uniform linear arrays. Based on such model, we discuss several existing source localization methods and some of their limitations, which leads us to strive for improvements in the accuracy of current techniques or formulate more precise ones. We also examine the fundamentals of prior knowledge and an effective preliminary idea to upgrade the accuracy of conventional algorithms by employing knowledge obtained on line. At the end of this chapter, we examine a non-uniform linear array model composed of two

levels, termed Nested Array, which will be the basis for the algorithm proposed in the last chapter.

Chapter3

In this chapter, we describe a new subspace-based method for DOA estimation for direction-of-arrival (DOA) estimation of signals originating from far-field sources, termed Multi-Step KAI ESPRIT (MS-KAI-ESPRIT), which iteratively reduces the disturbance factors of the estimated data covariance matrix and incorporates prior knowledge which is gradually obtained on line. An analysis of the MSE of the reshaped data covariance matrix is carried out along with comparisons between computational complexities of the proposed and existing algorithms. Simulations focusing on closely-spaced sources, where they are uncorrelated and correlated, illustrate the improvements achieved.

Chapter4

In this chapter, we extend the approach taken for the subspace-based ESPRIT to a Krylov subspace-based method, referred to as multi-step knowledge-aided iterative conjugate gradient (CG) (MS-KAI-CG). We also develop a version of MS-KAI-CG provided with forward-backward averaging, denoted by MS-KAI-CG-FB, that is suitable for correlated signals. Differently from existing knowledge-aided methods, which exploit known DOAs to improve the estimation of the covariance matrix of the input data, the proposed MS-KAI-CG algorithms exploit knowledge of the structure of the forward-backward smoothed covariance matrix and its perturbation terms and the gradual incorporation of prior knowledge, which, similarly to MS-KAI-ESPRIT, is obtained on line. Simulation results employing uncorrelated and correlated closely spaced sources illustrate the improvement achieved by the proposed method and the influence of its iterations on its performance.

Chapter5

This chapter expands even more the concept of the multi-step knowledg-aided method described in the two preceding chapters. Now, the method, which was no longer restricted to ULA based methods, is widened even further in order to embrace a non-uniform linear array like a two-level nested one. Now, differently of the two previous chapters, which are dedicated to ESPRIT and CG, we develop this method applying it to MUSIC algorithm, calling it multi-step knowledge-aided iterative nested MUSIC (MS-KAI-MUSIC). The proposed method refines the augmented sample covariance matrix, which is also obtained by exploiting a difference co-array structure and its properties,

and the gradual incorporation of prior knowledge, which, similarly to MS-KAI-CG and MS-KAI-ESPRIT, is obtained on line. Simulations show that MS-KAI-MUSIC significantly outperforms existing techniques.

Chapter 6

In this chapter, conclusions of this work are presented and future directions for this research topic are discussed.

1.5

Notations

In all expressions and equations of this thesis, lowercase non-bold letters represent scalar values, whereas bold lowercase and upper case letters represent vectors and matrices, respectively. $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^{-1}$ and $(\cdot)^H$ denote the complex conjugate operator, the transpose operator, the matrix inversion operator and the Hermitian transpose operator, respectively. $|\cdot|$, $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the absolutely value of a scalar, the Euclidean norm of a vector and the Frobenius norm of a matrix, respectively. $\mathrm{vec}\,(\cdot)$ stands for the vectorization function, $\odot$ represents the Khatri-Rao product, $\mathbb{E}\,[\cdot]$ denotes the expectations and $\mathrm{Tr}\,(\cdot)$ and $diag\,(\cdot)$ denote the trace and the diagonal entry of a matrix, respectively. An identity matrix of size $(\cdot)$ is represented by $\boldsymbol{I}_{(\cdot)}$.

1.6

Publications

The following papers have been written:

1.6.1

Journals

- Pinto, S.F.B, Lamare, R. C.: 'Multi-Step Knowledge-Aided Iterative MUSIC: Design and Analysis. (Under preparation).
- Pinto, S.F.B, Lamare, R. C.: 'Multi-Step Knowledge-Aided Iterative Conjugate Gradient Algorithms for DOA Estimation', *Digital Signal Processing*. (Submitted).

- Pinto, S.F.B, Lamare, R. C.: 'Multi-Step Knowledge-Aided Iterative ESPRIT: Design and Analysis', *IEEE Transactions on Aerospace and Electronic Systems*, **PP**, (99), February 2018. (IEEE Early Access Articles).

1.6.2

Conferences

- Pinto, S.F.B, Lamare, R. C.: 'Multi-Step Knowledge-Aided Iterative MUSIC for Direction Finding using Nested Arrays', 26th European Signal Processing Conference, September 2018, Rome, Italy. (Submitted).
- Pinto, S.F.B, Lamare, R. C.: 'Multi-Step Knowledge-Aided Iterative Conjugate Gradient for Direction Finding', Proc. IEEE 22nd ITG Workshop on Smart Antennas, March 2018, Bochum, Germany. (Accepted).
- Pinto, S.F.B, Lamare, R. C.: 'Multi-Step Knowledge-Aided Iterative ESPRIT for Direction Finding', Proc. IEEE 22nd International Conference on Digital Signal Processing, London, UK, August 2017, pp. 1-5.
- Pinto, S.F.B, Lamare, R. C.: 'Two-Step Knowledge-aided Iterative ESPRIT Algorithm', Proc. IEEE Twenty First ITG Workshop on Smart Antennas, Berlin, Germany, March 2017, pp. 1-5.
- Pinto, S.F.B, Lamare, R. C: 'Knowledge-Aided Parameter Estimation Based on Conjugate Gradient Algorithms', 35th Brazilian Communications and Signal Processing Symposium, September 2017, Sao Pedro, SP, Brazil, pp.1-5.

2

Literature Review

2.1

Introduction

This Chapter briefly reviews direction finding basics and existing algorithms, which form the basis for the development of new techniques for DOA estimation. In direction finding basics, we will describe the sensor array processing and its main geometries, a discrete-time sensor array model and the concept of beamforming . Thereafter, we revisit some of the most used estimation techniques for DOA estimation. Thus, we shortly examine the maximum likelihood technique, the Capon algorithm, which can be considered the most significant of the conventional class of direction finding methods, and the subspace based methods, including the Krylov subspace-based one, termed Conjugate Gradient (CG). Lastly, we review the theoretical foundations of the use of prior knowledge for sensor array signal processing.

2.1.1

Sensor Array Processing

Sensor array processing aims to process data collected at sensor elements in order to extract useful information, suppress interference and estimate parameters. In order to extract information such as the location of a signal source or the content of the signal, we often have to deal with the presence of the signal and interference. A single sensor with the ability to spatially discriminate signals coming from different directions can carry out this task. Such single-sensor systems can process signals using a continuous spatial extent known as aperture using a parabolic dish. The signals are reflected in the aperture in such a way that signals from the direction in which the dish is pointed are emphasized. The ability of a sensor to perform spatial processing is known a directivity and is governed by the shape and physical characteristics of

its geometric structure. However a single sensor system has several limitations. Since this sensor relies on mechanical pointing for directivity, it can only extract signals from only one direction at a time. Such sensor cannot adapt its response, which would require physically changing the aperture, in order to reject potentially strong interferers.

An array of sensors has the ability to overcome these shortcomings of a single sensor. The signals received by a sensor array can be combined in such a way that a particular direction is emphasized. However, since the direction in which the array points is almost independent of its orientation, the sensors can be combined in distinct ways so as to emphasize different directions and signals of interest. For this reason, multiple signals can be extracted simultaneously through separate signal processing operations per signal. Spatial signals propagate through space and arrive at an array of sensors that spatially samples the waveforms. A processor then collects the data from the sensor array in order to extract information.

The propagation of spatial signals is governed by the wave equation. For electromagnetic signals the wave equation can be deduced from Maxwell's equations, while for sound waves the solution is governed by the basic laws of acoustics. However, in either case, for a propagating wave emitted by a source at $\mathbf{r_0}$, one solution is a single-frequency wave given by

$$s(\mathrm{t},\ \mathbf{r}) = \frac{A}{\parallel \mathbf{r} - \mathbf{r_0} \parallel_2}\ \mathrm{e}^{-j\,2\,\pi\,f_c}\ (\mathrm{t} - \frac{\parallel \mathbf{r} - \mathbf{r_0} \parallel_2}{\mathrm{c}}), \tag{2-1}$$

where A is the complex amplitude, $\|\mathbf{r}\|$ represents the range or distance from the origin, f_c is the carrier frequency of the wave and c is the speed of propagation of the wave. It can be ignored the singularity of the source $s(t, \mathbf{r}_0) = \infty$, and the equation suppresses the dependencies on elevation and azimuth angles, since the wave propagates radially from the source. The signals travel in time, where the spatial propagation is determined by the direct coupling between space and time in order to satisfy the wave propagation. The wavelength of the propagating wave is given by $\lambda = \frac{c}{f_c}$, where c is the speed of the light. Our development relies on further assumptions. We assume that the propagating signals are produced by a point source, which means that the size of the source is small with respect to the distance between the source and the sensors that measure the signals. Another assumption is that the source is in the far field, which means it is at a sufficiently large distance to the the sensor array so that the spherically propagating wave can be reasonably approximated by a plane wave. In order to transmit and receive signals in space,

a designer must employ modulation and demodulation operations that allow the transmission of signals s(t using carrier waves centered at the frequency f_c and the processing of those signals with the extraction of sufficient statistics in discrete-time form $x[i]$, as illustrated in Fig. 2.1.

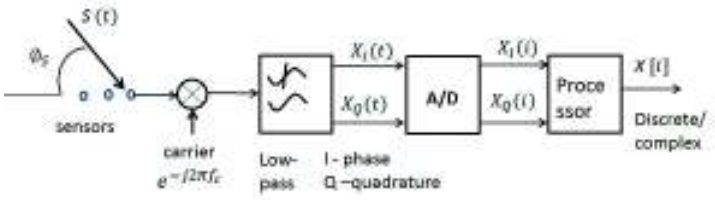

Figure 2.1: Analog-Digital converter

2.1.1.1

Standard Geometries

A sensor array can be organized with M sensors placed according to a particular geometry that affects its radiating properties. The most common array geometries are the uniform linear array (ULA), the uniform circular array (UCA) and the uniform planar array (UPA). In what follows, without loss of generality, we will describe several properties of the first and the latest geometries which are the most often used in practice. They receive signal waveforms whose components at each sensor are delayed replicas of the associated signal waveform.

Uniform Linear Array

Consider a plane wavefront with the waveform *s(t)* impinging on a ULA of M sensors at an angle θ, as illustrated in the Fig. 2.2.

The incident angle on the array is known as the direction of arrival (DOA) of the signal. Since in a ULA all elements are equally spaced, the signal is given by any two successive sensors with a time delay given by $\tau = \frac{d \sin(\theta)}{c}$, where d is the spacing between sensor elements and d $\sin(\theta)$ is the distance in propagation for the signal *s(t)* to reach successive sensors. As a result, the delay of the mth with respect to the first element in the sensor array is given by $\tau_m = (m-1)\,\frac{d \sin(\theta)}{c}$.

By multiplying the signal *s(t)* by the carrier $\mathrm{e}^{-j\,2\,\pi\,f_c}$ and using the first sensor as a reference, we obtain $\mathrm{s}_1(t) = \; s(t)\, e^{-j\,2\,\pi\,f_c\,t}$.

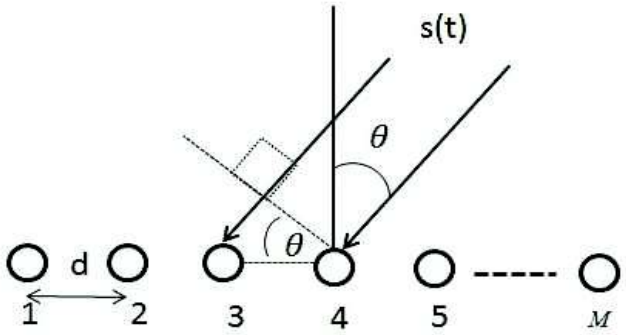

Figure 2.2: Uniform Linear Array

The generalization of the above result to the m^{th} sensor yields $\mathrm{s}_n(t) = s(t)e^{-j\,2\,\pi\,f_c\,\tau_n} = s(t)e^{-j\,2\,\pi\,f_c\,(m-1)\,\frac{d\,\sin(\theta)}{c}} = s(t)e^{-j\,2\,\pi\,(n-1)\,\frac{d\,\sin(\theta)}{\lambda}}$, where $\lambda = \frac{\mathrm{c}}{\mathrm{f_c}}$, and we can write the received signal for the ULA as

$$\mathbf{x}\,[i] \;=\; \begin{bmatrix} 1 \\ e^{-j\,2\,\pi\,\frac{d\,\sin(\theta)}{\lambda}} \\ \vdots \\ e^{-j\,2\,\pi\,(M-1)\,\frac{d\,\sin(\theta)}{\lambda}} \end{bmatrix} s\,[i] \;+\; \mathbf{n}\,[i] \;=\; \mathbf{a}\,(\theta)\;s\,[i] \;+\; \mathbf{n}\,[i] \qquad (2\text{-}2)$$

where the [Mx1] vector $\mathbf{a}\,(\theta)$ is the steering vector of the ULA and $\mathbf{n}\,[i]$ is the noise vector.

Uniform Planar Array

Let us now consider a signal that impinges on a UPA of $\mathrm{M} = \mathrm{M}_1\,\mathrm{x}\,\mathrm{M}_2$ sensors at an azimuth angle ϕ and at an elevation angle θ as depicted in the Fig. 2.3. The distances between the sensors are d_1 and d_2.

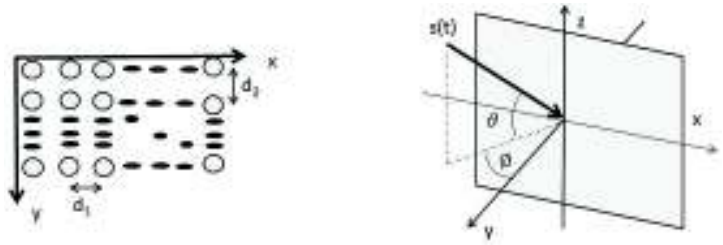

Figure 2.3: Uniform Planar Array

In this situation, the azimuth and elevation angles of the waveform *s(t)* are the DOAs of the signal. The delay introduced in the $(\mathrm{m}_1,\ \mathrm{m}_2)^{\mathrm{th}}$ sensor is given by $\tau_{m_1,\,m_2} = \frac{m_1\,d_1}{c}\sin(\theta)\cos(\phi) + \frac{m_2\,d_2}{c}\sin(\theta)\sin(\phi)$ where m_1 and m_2 are the indexes of the sensors.

After multiplication by $\mathrm{e}^{-j\,2\,\pi\,f_c}$, the signal $s(t)$ that arrives at the $(\mathrm{m}_1,\,\mathrm{m}_2)^{\mathrm{th}}$ sensor is given by $\mathrm{s}_{(m_1,\,m_2)}(t) = s(t)\;e^{-j\,2\,\pi\,f_c\,\tau_{m1,\,m2}} = s(t)\;e^{-j\,2\,\pi\,\frac{m_1\,d_1}{\lambda}\sin(\theta)\cos(\phi)\,+\,\frac{m_2\,d_2}{\lambda}\sin(\theta)\sin(\phi)}$. The received signal for the UPA can be writen as

$$\begin{aligned}\mathbf{x}\,[i] &= \begin{bmatrix} 1 \\ e^{-j\,2\,\pi\,\frac{m_1\,d_1}{\lambda}\sin(\theta)\cos(\phi)\,+\,\frac{m_2\,d_2}{\lambda}\sin(\theta)\sin(\phi)} \\ \vdots \\ e^{-j\,2\,\pi\,\frac{(M_1-1)\,d_1}{\lambda}\sin(\theta)\cos(\phi)\,+\,\frac{(M_2-1)\,d_2}{\lambda}\sin(\theta)\sin(\phi)} \end{bmatrix} s\,[i]\;+\;\mathbf{n}\,[i] \\ &= \mathbf{a}\,(\theta,\,\phi)\;s\,[i]\;+\;\mathbf{n}\,[i] \end{aligned} \tag{2-3}$$

where the [Mx1] vector $\mathbf{a}\,(\theta,\,\phi)$ is the steering vector of the UPA and $\mathbf{n}\,[i]$ is the noise vector.

Other geometries such as sparse arrays and volumetric arrangements (3D) can be used by system designers. Besides, it is also possible to employ or form distributed arrays composed of single sensor and subarrays, which offer an advantage in terms of manufacturing. UPAs are used for large systems in radar.

Concerning the spacing of elements d and $(d_1,\;d_2)$, the former for ULAs and the latter for UPAs, it is customary to set them to $\frac{\lambda}{2}$ to avoid coupling effects. However, the use of compact antenna arrays is often associated with closely spaced elements $(\frac{\lambda}{2},\;\frac{\lambda}{4},\;\frac{\lambda}{8}, etc)$, which requires extra care with the couple effects.

2.1.1.2

Discrete-time Models

This subsection aims to deal with discrete-time models for arbitrary sensor arrays that have the objective of extracting a desired signal embedded in interference. The models are general and can be used for any sensor array geometry, as illustrated in Fig.2.4.

The schematic above considers an arbitrary sensor array that aims to extract the desired signal $\mathrm{s_d(t)}$ in the presence of $(\mathrm{D}-1)$ interfering signals. The sensor array has M elements and processes the incoming signal, resulting in

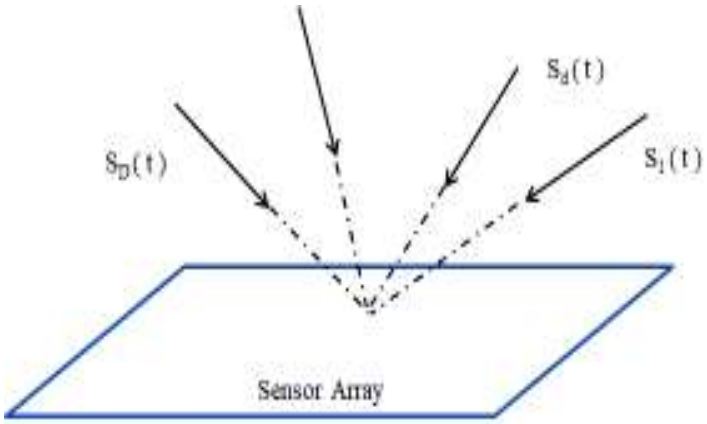

Figure 2.4: Sensor array geometry for discrete-time models

the following discrete model:

$$\begin{aligned}\mathbf{x}[i] &= s_d[i]\,\mathbf{a}(\theta_d,\,\phi_d) \;+\; \sum_{k=1,\,k\neq d}^{D} s_k[i]\,\mathbf{a}(\theta_k,\,\phi_k) + \;\mathbf{n}[i] \\ &= \mathbf{A}(\theta,\,\phi)\;\mathbf{s}[i] \;+\; \mathbf{n}[i] \end{aligned} \tag{2-4}$$

where the [Mx1] vectors $\mathbf{a}(\theta_d,\,\phi_d)$ correspond to the array responses of the signals and the [MxD] matrix $\mathbf{A}(\theta,\,\phi)$ contains all the steering vectors. The discrete-time desired signal $s_d[i]$ is perturbed by the $(D-1)$ interfering signals $s_k[i]$, $k \neq d$, $k = 1,\,2,\,\cdots,D$, and these D impinging signals are represented by the [Dx1] vector $\mathbf{s}[i]$. The [Mx1] noise vector $\mathbf{n}[i]$ is assumed to be drawn from complex Gaussian random variables with zero mean and variance σ_n^2.

A particular case of this model for a ULA is described by

$$\mathbf{x}[i] = s_d[i]\,\mathbf{a}(\theta_d) \;+\; \sum_{k=1,\,k\neq d}^{D} s_k[i]\,\mathbf{a}(\theta_k) + \;\mathbf{n}[i] \tag{2-5}$$

$$= \mathbf{A}(\theta)\;\mathbf{s}[i] \;+\; \mathbf{n}[i] \tag{2-6}$$

where $\mathbf{a}(\theta_k)$ are the steering vectors of a ULA for $k = 1,\,2,\,\cdots,D$. As described by these models, a sensor array system depends on the structure of the array, the DOAs θ_k (and ϕ_k) and the processing required to extract desired information for retrieving $s_d[i]$. Except as otherwise specified in this work, our study on beamforming algorithms to extract $s_d[i]$ from $\mathbf{x}[i]$ will be done by assuming the knowledge of θ_k (and ϕ_k). Then, we will develop direction finding methods to estimate θ_k (and ϕ_k).

2.1.1.3

Beamforming

In many applications, the desired information to be extracted from an array of sensors is the content of a spatially propagating signal from a certain direction. The content may be a message contained in the signal, such as communications applications, or merely the existence of the signal, as in radar and sonar systems. to this end, we want to linearly combine the signals from all the sensors with a certain weighting, so as to examine signals arriving from a specific angle. This operation is known as beamforming and is illustrated in Fig. 2.5.

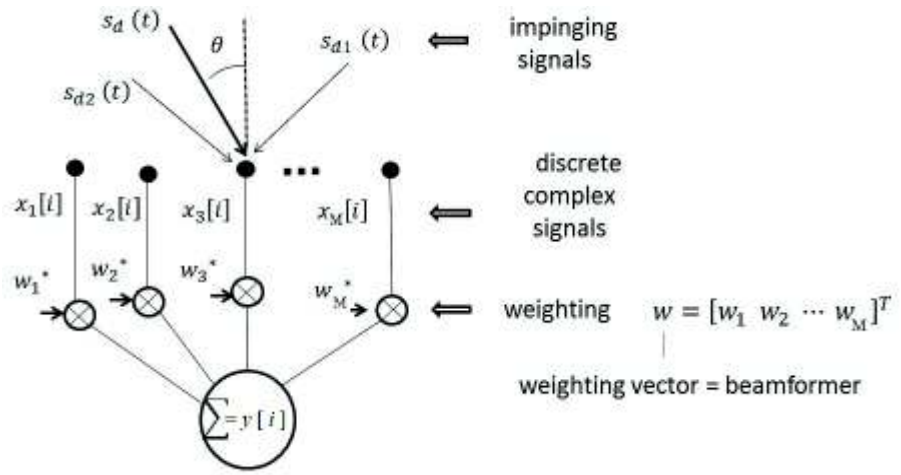

Figure 2.5: Beamforming

The weighting process emphasizes signals from a particular direction while attenuating those from other directions and can be thought as forming a beam. In this sense, a beamformer is a spatial filter that places nulls in the direction of interfering signals. Beamforming can also be viewed as an electronic steering since the weights w_m, $m = 1, 2, \cdots, M$ are applied using digital signal processing following the reception of the signal for the purpose of steering the array in a particular direction.

The output is formed by a weighted combination of signals from the N elements of the sensor array, as follows:

$$\begin{aligned} \mathrm{y}[i] &= \sum_{m=1}^{M} w_m^* x_m[i] \\ &= \mathbf{w}^H \mathbf{x}[i] \end{aligned} \tag{2-7}$$

where $\mathbf{w} = [w_1 \ w_2 \ \cdots \ w_M]^T$ is the [Mx1] vector of beamforming weights, which must be computed by an algorithm.

An important concept for beamforming techniques is that of beam response, which is the response of a beamformer $\mathtt{w}$ as a function of the θ. This beam response can be computed by applying it to a set of array steering vectors $\mathbf{a}(\theta)$ from all possible angles as given by

$$\mathrm{B}(\theta) = \mathbf{w}^H \mathbf{a}(\theta), \tag{2-8}$$

where $\theta_{min} \leq \theta \leq \theta_{max}$

Optimum beamformer

An improved concept of the previously described weighting process, which is classified optimum, consider methods that employ the statistics of the data to derive the beamforming weights. In particular, we refer to optimum beamformers as those that optimize a certain criterion based on the knowledge of the statistics of the data and we refer to adaptive beamformers as the techniques that employ estimates of the statistics of the data.

The optimum beamformer can be derived based on a signal model for a ULA, containing interference in addition to the desired signal and noise, as given in equation 2-6, as follows

$$\mathbf{x}[i] = s_d[i]\,\mathbf{a}(\theta_d) + \sum_{k=1,\,k\neq d}^{D} s_k[i]\,\mathbf{a}(\theta_k) + \mathbf{n}[i], \tag{2-9}$$

where $s_d[i]$ is the desired signal and, as a result, its signal vector $\mathbf{s}[i]$. Moreover, the desired power received by the element is expressed by $E[|s_d[i]|^2] = \sigma_s^2$.

The interference-plus-noise component of the signal is defined as $\mathbf{x}_{j+n}[i] = \mathbf{j}[i] + \mathbf{n}[i]$, which are both modeled as zero-mean stochastic processes. The interference has spatial correlation according to the angles of the contributing interferers, while the noise is spatially uncorrelated and has zero mean and variance σ_n^2. As a result, the correlation matrix is given by

$$\mathbf{R} = E[\mathbf{x}[i]\,\mathbf{x}^H[i]] = \sigma_s^2\,\mathbf{a}(\theta_d)\,\mathbf{a}^H(\theta_d) + \mathbf{R}_j + \mathbf{R}_n, \tag{2-10}$$

where $\mathbf{R}_j = E\left[\mathbf{j}[i]\,\mathbf{j}^H[i]\right]$ and $\mathbf{R}_n = E[\mathbf{n}[i]\,\mathbf{n}^H[i]]$ are the interference and the noise correlation matrices, respectively. The interference-plus-noise correlation

matrix is the sum of these latter two matrices, that is to say $\mathbf{R}_{j+n} = \mathbf{R}_j + \mathbf{R}_n$, where $\mathbf{R}_n = E[\mathbf{n}[i]\,\mathbf{n}^H[i]] = \sigma_n^2\mathbf{I}$, since the noise is assumed spatially uncorrelated.

The signal-to-interference-plus-noise ratio (SINR) at the beamformer output is described by

$$\mathrm{SINR}_{\mathrm{out}} = \frac{E\left[\mathbf{w}^H\,\mathbf{s}[i]\;\;\mathbf{s}^H[i]\,\mathbf{w}\right]}{E\left[\mathbf{w}^H\,\mathbf{x}[i]\,\mathbf{x}^H[i]\,\mathbf{w}\right]} = \frac{\mathbf{w}^H\,\mathbf{R}_s\,\mathbf{w}}{\mathbf{w}^H\,\mathbf{R}_{j+n}\,\mathbf{w}} \tag{2-11}$$

where $\mathbf{R}_s = E\left[\mathbf{s}[i]\,\mathbf{s}^H[i]\right]$ is the desired signal correlation matrix.

By minimizing the interference-plus-noise power at the beamformer output, that is, $\mathbf{R}_{j+n}\;\mathbf{w}$, while keeping the desired signal response $\mathbf{w}^H\;\mathbf{R}_s\;\mathbf{w}$. This results in the optimization problem, as follows:

$$\min\;\mathbf{w}^H\,\mathbf{R}_{j+n}\,\mathbf{w} \quad \text{subjected to} \quad \mathbf{w}^H\mathbf{a}\left(\theta_d\right) = \text{ constant} \tag{2-12}$$

The solution to this optimization problem is known as minimum-variance distortionless response (MVDR) beamformer. The MVDR beamformer maximizes the SINR by matching the response of signals impinging on the array from a direction $\mathbf{a}\left(\theta_d\right)$ and can be considered an optimum spatial filter. As in practice $\mathbf{R}_{j+n}$ is difficult to be obtained, the standard MVDR is formulated by replacing $\mathbf{R}_{j+n}$ with $\mathbf{R}$ and employing a similar optimization.

The MVDR beamformer optimization with $\mathbf{R}$ is given by

$$\min\;\mathbf{w}^H\,\mathbf{R}\,\mathbf{w} \quad \text{subjected to} \quad \mathbf{w}^H\mathbf{a}\left(\theta_d\right) = 1 \tag{2-13}$$

and its solution can be obtained with the method of Lagrange multipliers, which results in the desired weight vector

$$\mathbf{w}_o = \frac{\mathbf{R}^{-1}\,\mathbf{a}\left(\theta_d\right)}{\mathbf{a}^H\left(\theta_d\right)\;\mathbf{R}^{-1}\,\mathbf{a}\left(\theta_d\right)} \tag{2-14}$$

The minimum variance associated to the preceding solution can be obtained by substituting $\mathbf{w}_o$ into the cost function [5]

$$J\left(\mathbf{w}_o\right) = \mathbf{w}_o^H\;\mathbf{R}\;\mathbf{w}_o, \tag{2-15}$$

resulting in the power spectrum in Eq.(2-16), also known as Capon's spatial spectrum.

$$J\left(\mathbf{w}_o\right) = \frac{1}{\mathbf{a}^H\left(\theta_d\right)\;\mathbf{R}^{-1}\,\mathbf{a}\left(\theta_d\right)} \;=\; P_{Cap}\left(\theta_d\right) \tag{2-16}$$

2.1.2

Maximum Likelihood estimation

Maximum Likelihood-type method (ML) is one of the first techniques developed for DOA estimation. As mentioned in 1.1, ML is based on a parametric approach. Despite its computational intensity that makes it less attractive then other conventional methods, ML estimation plays an essential role in DOA estimation by providing the Cramer-Rao lower bound (CRLB), which is a measure against which any algorithm can be compared.

Let us consider the signal model given in Eq. (2-6) as follows:

$$\mathbf{x}[i] = s_d[i]\,\mathbf{a}(\theta_d) \; + \sum_{k=1,\, k\neq d}^{D} s_k[i]\,\mathbf{a}(\theta_k) + \; \mathbf{n}[i] \tag{2-17}$$

$$= s_d[i] \; \mathbf{a}(\theta_d) \; + \; \mathbf{j}[i] + \; \mathbf{n}[i]\,, \tag{2-18}$$

where it is assumed that the interference $\mathbf{j}[i]$ has known statistics, the steering vectors $\mathbf{a}(\theta_d)$ are assumed linearly independent, and the noise is obtained from a complex Gaussian random process of mean equal to zero and variance σ_n^2 with the samples of the noise being statistically independent.

The derivation of the ML estimator employs the joint probability density function (pdf) of the data $\mathbf{x}[i]$ given the parameter θ_d, as described by

$$\mathrm{p}_{\mathrm{x}|\theta_\mathrm{d}}(\mathbf{x}[i] \mid \theta_d) = \frac{1}{\pi^N \det(\mathbf{R})} \exp\big(- \; (\mathbf{x}[i] - s_d[i]\,\mathbf{a}(\theta_d))^H \; \mathbf{R}^{-1} (\mathbf{x}[i] \; - \; s_d[i]\,\mathbf{a}(\theta_d))\big) \tag{2-19}$$

where N is the number of snapshots. The ML estimator computes

$$\hat{\theta}_d \; = \; \arg \max_{\theta_d} \; \mathrm{p}_{\mathrm{x}|\theta_\mathrm{d}}(\mathbf{x}[i] \mid \theta_d) \tag{2-20}$$

which is equivalent to maximizing the logarithm of $\mathrm{p}_{\mathrm{x}|\theta_\mathrm{d}}(\mathbf{x}[i] \; , \; \theta_d)$, as follows:

$$\begin{aligned}\hat{\theta}_d &= \arg \max_{\theta_d} \ln \mathrm{p}_{\mathrm{x}|\theta_\mathrm{d}}\left(\mathbf{x}\left[i\right] , \theta_d\right) \\ &= \arg \min_{\theta_d} \Big(\left(\mathbf{x}\left[i\right] - s_d\left[i\right]\mathbf{a}\left(\theta_d\right)\right)^H \mathbf{R}^{-1}\left(\mathbf{x}\left[i\right] - s_d\left[i\right]\mathbf{a}\left(\theta_d\right)\right)\Big) \\ &= \arg \max_{\theta_d} \left[\frac{|\mathbf{a}^H\left(\theta_d\right) \mathbf{R}^{-1}\mathbf{x}\left[i\right]|^2}{\mathbf{a}^H\left(\theta_d\right) \mathbf{R}^{-1}\mathbf{a}\left(\theta_d\right)}\right]\end{aligned} \tag{2-21}$$

The function

$$P_{ML}\left(\theta_d\right) = \frac{|\mathbf{a}^H\left(\theta_d\right) \mathbf{R}^{-1}\mathbf{x}\left[i\right]|^2}{\mathbf{a}^H\left(\theta_d\right) \mathbf{R}^{-1}\mathbf{a}\left(\theta_d\right)}, \tag{2-22}$$

which is inside the braces of (2-21), is the ML estimate of the incoming data $\mathbf{x}\left[i\right]$. The DOA estimate is the angle associated with the maximum of $P_{ML}\left(\theta_d\right)$. Despite the optimality of the ML approach, it is impractical in most situations because the algorithm requires a search to find the peaks of $P_{ML}\left(\theta_d\right)$ in a dense grid with small angular spacing.

2.1.2.1

Cramer-Rao lower bound

As mentioned in 2.1.2, the measure of the ML estimation procedure is the CRLB, which is very useful to give a designer how accurate a DOA estimator can be. From estimation theory, we know that the variance C of an unbiased estimate of $\hat{\theta}_d$ is greater than or equal to its CRLB (C_{CR}), that is to say

$$C\left(\hat{\theta}_d\right) \geq C_{CR}\left(\hat{\theta}_d\right) \triangleq J_{dd}^{-1} \tag{2-23}$$

where J_{dd}^{-1} is the d^{th} diagonal entry of the inverse of the Fisher information matrix J whose $(i,j)^{th}$ element is given by

$$Jij = -\mathbb{E}\left[\frac{\delta^2}{\delta\theta_i\,\delta\theta_j}\left[\ln \mathrm{p}_{\mathrm{x}|\theta_\mathrm{d}}\left(\mathbf{x} \mid \theta\right)\right]\right], \tag{2-24}$$

where $\mathbb{E}\left[\cdot\right]$ means the expectation operator.

Using the signal model for a ULA with m sensors and computing the preceding partial derivatives, we can obtain the CRLB [17] for $D = 1$ signal.

$$\begin{aligned}C\left(\hat{\theta}_d\right) &\geq \frac{6\,\sigma_n^2}{\sigma_s^2\, m\left(m^2 - 1\right)\left(k\,d\right)^2 \sin\theta_d} \\ &= \frac{6}{SNR\, m\left(m^2 - 1\right)\left(k\,d\right)^2 \sin\theta_d}\end{aligned} \tag{2-25}$$

where $SNR = \frac{\sigma_s^2}{\sigma_n^2}$ contributes to decreasing the CRLB as it increases. The number of sensors also contributes to a reduction of the CRLB. More specific CRLB in matrix forms taking into account the variation of the number of snapshots are studied in [18].

2.1.3

Capon algorithm

In 2.1.1.3 it was discussed the MVDR beamformer which is identical to the Capon's method. In order to estimate the DOA, we compute the Eq. 2-16 over the whole range of θ_d and locate its peaks as described by

$$\hat{\theta}_d \;=\; \arg \max_{\theta_d} \; P_{Cap}\left(\theta_d\right), \quad d = 1, 2, \cdots, D \tag{2-26}$$

where the cost of the required search depends on the angle spacing.

Among the disadvantages of Capon's method are that it requires a search, it fails for correlated signals and provides poor resolution for closely spaced sources. These drawbacks can be seen in a scenario in which a ULA with 10 sensors inter-element spaced by $\Delta = \frac{\lambda_c}{2}$ receives two impinging uncorrelated complex Gaussian signals with equal power. The sources are separated by 2.8^o, at $(88.6^o, 91.4^o)$, and the number of available snapshots is N=40. The probability of resolution obtained using 100 trials is depicted in Fig. 2.6.

In order to have better comparisons, we use the same scenario for the performance analysis of all algorithms.

2.1.3.1

Root-Capon

An approach that circumvent the need for a search and improve the performance of the Capon's method is to employ a root version of the algorithm. This is limited to the case of a ULA. In the root-Capon algorithm, we consider the spatial spectrum given by

$$P_{root-Cap}\left(\theta_d\right) = \frac{1}{\mathbf{a}^H\left(\theta_d\right)\; \mathbf{R}^{-1}\, \mathbf{a}\left(\theta_d\right)} \;=\; \frac{1}{C\left(\theta_d\right)} \tag{2-27}$$

where $C\left(\theta_d\right) \;=\; \mathbf{a}^H\left(\theta_d\right)\; \mathbf{R}^{-1}\, \mathbf{a}\left(\theta_d\right)$ is the null spectrum.

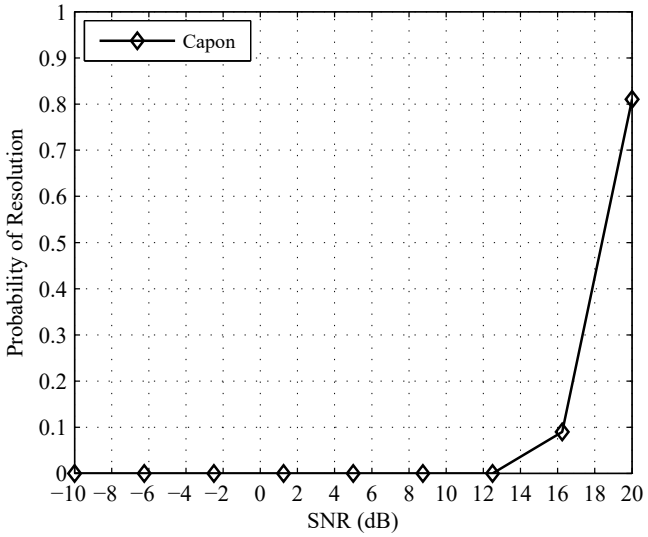

Figure 2.6: Probability of resolution versus SNR, 2 uncorrelated sources, 2.8 degrees, ULA-10 sensors, 40 snapshots, 100 runs

The algorithm constructs the null spectrum and computes the its z_d roots, $d = 1,\ 2, \cdots, D$, inside and closest to the unit circle. Then, we have

$$\hat{\theta}_d = \sin^{-1}\left[\left(\frac{\lambda}{2\ \pi\ d}\right) \arg\ z_d\right] \tag{2-28}$$

2.1.4

MUSIC

The Multiple signal Classification (MUSIC) algorithm proposed by [8] in 1979 is a popular high resolution based on eigenstructure technique. The main idea behind this DOA algorithm is that of performing eigenvalue decomposition of the correlation matrix, separating it into two subspaces: signal subspace and the noise subspace. Since the signal subspace is spanned by the array steering vector of the received signals, this makes the steering vector orthogonal to the noise subspace. Therefore, the product of the array steering vector and the noise subspace is a null for a particular angle of arrival.

Let us consider the problem of model order selection, i.e., of estimating the number of signals D in a ULA with M elements. For this aim, we can take into account the received data model in Eq. (2-6). Without loss of generality, we do not consider interferers, which results in the model

$$\mathbf{x}[i] = \sum_{d=1}^{D} s_d[i]\,\mathbf{a}(\theta_d) + \mathbf{n}[i] \tag{2-29}$$

$$= \mathbf{A}(\theta)\ \mathbf{s}[i] + \mathbf{n}[i] \tag{2-30}$$

where the [MxD] array manifold $\mathbf{A}(\theta_d)$ contains the [Mx1] steering vectors $\mathbf{a}(\theta_d)$ of the D signals and the vector $\mathbf{s}[i]$ has the D signals. The [Mx1] noise vector $\mathbf{n}[i]$ has spatially uncorrelated samples of a complex Gaussian random process of zero mean and covariance matrix $\sigma_n^2\,\mathbf{I}$. The correlation matrix $\mathbf{R}$ of the received data is given by

$$\mathbf{R} = E\left[\mathbf{x}[i]\ \mathbf{x}^H[i]\right] = \mathbf{A}\ E\left[\mathbf{s}[i]\ \mathbf{s}^H[i]\right]\ \mathbf{A}^H + E\left[\mathbf{n}[i]\ \mathbf{n}^H[i]\right]$$
$$= \mathbf{A}\,\mathbf{R}_s\,\mathbf{A}^H + \sigma_n^2\,\mathbf{I} \tag{2-31}$$

where the [MxM] $\mathbf{R}$ matrix has full rank, σ_n^2 is the noise power and the $\mathbf{R}_s$ has dimensions [DxD]. The matrix $\mathbf{R}$ can be estimated by taking sample averages as given

$$\hat{\mathbf{R}}[l] = \frac{1}{i}\sum_{l=1}^{i}\mathbf{x}[l]\,\mathbf{x}^H[l] \tag{2-32}$$

and

$$\mathbf{R}_s = E\left[\mathbf{x}[i]\ \mathbf{x}^H[i]\right] = diag\left\{\sigma_1^2, \cdots, \sigma_D^2\right\} \tag{2-33}$$

$\mathbf{R}$ has M eigenvalues $[\lambda_1, \lambda_2, \cdots, \lambda_M]$ and M associated eigenvectors, making a subspace $\bar{\boldsymbol{E}} = [\bar{e}_1, \cdots, \bar{e}_M]$. Sorting the M eigenvalues from the largest to the smallest, the subspace $\bar{\boldsymbol{E}}$ can be decomposed into two subspaces:

$$\bar{\boldsymbol{E}} = [\bar{\boldsymbol{e}}_1, \cdots, \bar{\boldsymbol{e}}_D\ \ \bar{\boldsymbol{e}}_{D+1}, \cdots, \bar{\boldsymbol{e}}_M]$$
$$= \left[\bar{\boldsymbol{E}}_{\boldsymbol{S}}\ \ \bar{\boldsymbol{E}}_{\boldsymbol{N}}\right] \tag{2-34}$$

$\bar{\boldsymbol{E}}_{\boldsymbol{N}}$ is the M x [M − D] noise subspace composed of the eigenvectors associated with noise, whereas $\bar{\boldsymbol{E}}_{\boldsymbol{S}}$ is the [M x D] signal subspace composed of the eigenvectors associated with the arriving signal.

Due to the orthogonality of the noise subspace and the array steering vector at the angles of arrival $(\theta_1, \theta_2, \cdots, \theta_D)$, the matrix product $\mathbf{a}^H(\theta_d)\,\bar{\boldsymbol{E}}_{\boldsymbol{N}}\,\bar{\boldsymbol{E}}_{\boldsymbol{N}}^{\boldsymbol{H}}\,\mathbf{a}(\theta_d)$ is zero for this angles. The reciprocal of this matrix product creates sharp peaks at the angles of arrival. Thus the MUSIC spatial spectrum is given as

$$\boldsymbol{P}_{MU}(\theta_d) = \frac{1}{\|\mathbf{a}^H(\theta_d)\,\bar{\boldsymbol{E}}_{\boldsymbol{N}}\,\bar{\boldsymbol{E}}_{\boldsymbol{N}}^{\boldsymbol{H}}\,\mathbf{a}(\theta_d)\,\|} \tag{2-35}$$

In order to estimate the DOA, we compute $\boldsymbol{P}_{MU}(\theta_d)$ over the whole range of θ_d and locate the peaks as described by

$$\hat{\theta}_d = \arg \max_{\theta_d} P_{MU}(\theta_d), \quad d = 1, 2, \cdots, D \tag{2-36}$$

Despite MUSIC performs better than Capon algorithm for closely-spaced sources in terms of probability and RMSE, its performance in this scenario is poor, as can be seen in Figs. 2.7 and 2.8.

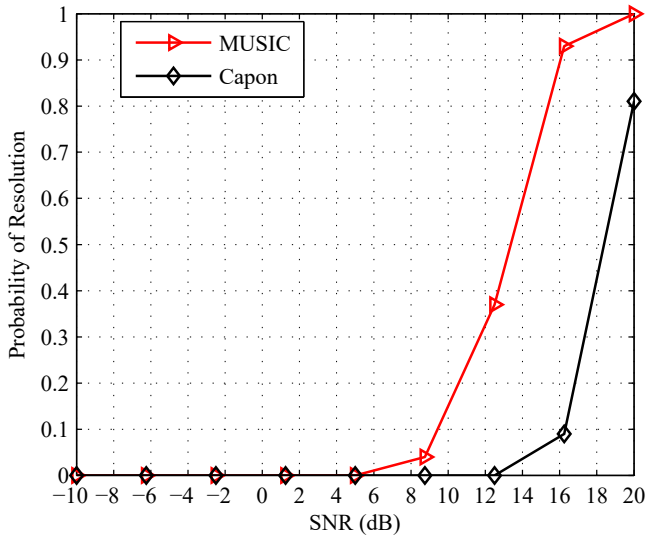

Figure 2.7: fig/Probability of resolution versus SNR 2 uncorrelated sources, 2.8 degrees, ULA 10 sensors, 40 snapshots, 100 runs

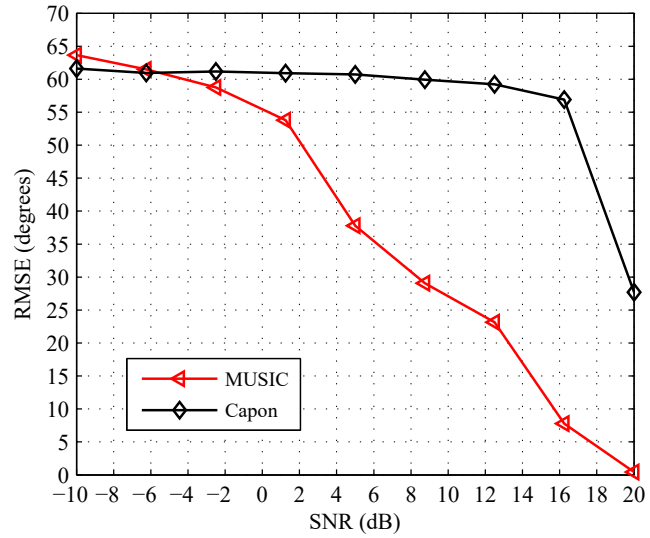

Figure 2.8: RMSE versus SNR, 2 uncorrelated sources, 2.8 degrees, ULA 10 sensors, 40 snapshots, 100 runs

Root-MUSIC

Instead of plotting the spatial spectrum against the angles and searching for the peaks, this variant of MUSIC algorithm involves finding the roots of a polynomial. Starting with the spatial spectrum of MUSIC in (2-35), and defining $\boldsymbol{C} = \bar{\boldsymbol{E}}_{\boldsymbol{N}}\, \bar{\boldsymbol{E}}_{\boldsymbol{N}}^{\boldsymbol{H}}$, its denominator can be writen as

$$\boldsymbol{P}_{MU}\left(\theta_d\right) = \frac{1}{\|\mathbf{a}^H\left(\theta_d\right)\, \boldsymbol{C}\, \mathbf{a}\left(\theta_d\right)\|} \tag{2-37}$$

The m^{th} element $a_n\left(\theta_d\right)$ of the array steering vector is defined as

$$a_m\left(\theta_d\right) = e^{-j\, k\, d\, m\, \sin(\theta_d)}, \quad m = 0, 1, \cdots, M-1 \tag{2-38}$$

The denominator, thus can be rewritten as

$$\begin{aligned} \mathbf{a}^H\left(\theta_d\right)\, \boldsymbol{C}\, \mathbf{a}\left(\theta_d\right) &= \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} e^{-j\, k\, d\, m\, \sin(\theta_d)}\, \boldsymbol{C}_{m\,n}\, e^{j\, k\, d\, n\, \sin(\theta_d)} \\ &= \sum_{l=-(N-1)}^{N-1} \boldsymbol{C}_l\, e^{j\, k\, d\, l\, \sin(\theta_d)}, \end{aligned} \tag{2-39}$$

where $\boldsymbol{C}_{nm}$ is the entry in the m^{th} row and n^{th} column of $\boldsymbol{C}$ and $\boldsymbol{C}_l = \sum_{m-n=l} \boldsymbol{C}_{m\,n}$ is the sum of the elements along the l^{th} diagonal of $\boldsymbol{C}$. Letting $z = e^{j\, k\, d\, \sin(\theta_d)}$, Eq.(2-39) simplifies to

$$D\left(z\right) = \sum_{l=-M+1}^{M-1} \boldsymbol{C}_l\, z^l \tag{2-40}$$

The roots of $D\left(z\right)$ that lie closest to the unit circle correspond to the poles of the MUSIC spatial spectrum. These $2\left(M-1\right)$ roots can be written as

$$z_d = e^{j\, \arg(z_d)}, \quad d = 1, 2, \cdots, 2\left(M-1\right) \tag{2-41}$$

Choosing those roots inside the unit circle whose magnitude are $|z_d| \simeq 1$, and comparing $e^{j\, \arg(z_d)}$ to $e^{-j\, k\, d\, \sin(\theta_d)}$, gives

$$\hat{\theta}_d = -\sin^{-1}\left\{\frac{\arg\, z_d}{k\, d}\right\}, \quad k = \frac{2\,\pi}{\lambda} \tag{2-42}$$

2.1.5

ESPRIT

The estimation of signal parameters via rotational invariance (ESPRIT) algorithm [10], reduces the computational and storage requirements of MUSIC and avoids an exhaustive search. The key idea is to decompose an M element

array into two identical centro-symetric subarrays, each of them with S elements. The objective of ESPRIT algorithm is to estimate the angle of arrival by determining the rotation operator Φ. The separation distance between the two sensors is assumed to be $\frac{\lambda}{2}$. In this case, the first element in the original array is the first element of the first subarray whereas the second element of the original array is the first element in the second subarray. Fig.2.9 shows a M element linear array and one of the possible configurations.

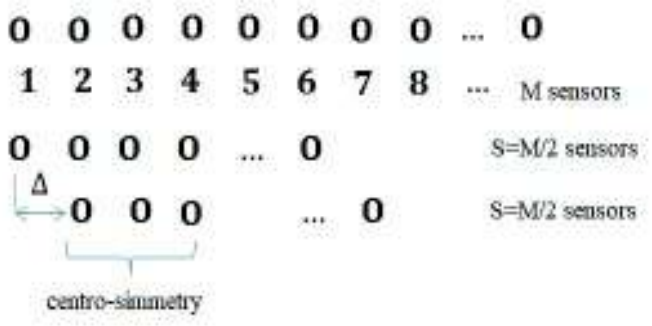

Figure 2.9: ULA decomposition in ESPRIT algorithm

Let us consider D signals hitting the subarrays. Also letting $\mathbf{x}_1[i]$ and $\mathbf{x}_2[i]$ be the received signal in the two subarrays, corrupted by additive Gaussian noise $\mathbf{n}_1[i]$ and $\mathbf{n}_2[i]$, respectively.

$$\begin{aligned} \mathbf{x}_1[i] &= \mathbf{A}\, s[i] + \mathbf{n}_1[i] \\ \mathbf{x}_2[i] &= \mathbf{A}\, \Phi\, s[i] + \mathbf{n}_2[i] \end{aligned} \tag{2-43}$$

where $\mathbf{x}_1[i]$, $\mathbf{x}_2[i]$, $\mathbf{n}_1[i]$ and $\mathbf{n}_1[i]$ are [Sx1] matrices. $\mathbf{A}$ is the [SxD] steering matrix and the variable Φ is a [DxD] diagonal matrix called rotation operator

$$\Phi = diag\left\{e^{j\psi_1},\, e^{j\psi_2}, \cdots, e^{j\psi_D}\right\} \tag{2-44}$$

where $\psi_d = -2\,k\Delta\,\sin\theta_d,\ d = 1, 2, \cdots, D$ and Δ measured in wavelengths. From Eq.(2-43), correlation matrices $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$ of the signals in the two subarrays can be estimated as

$$\begin{aligned} \mathbf{R}_{11} &= E\left[\mathbf{x}_1[i]\ \mathbf{x}_1^H[i]\right] \\ \mathbf{R}_{22} &= E\left[\mathbf{x}_2[i]\ \mathbf{x}_2^H[i]\right] \end{aligned} \tag{2-45}$$

By eigen-decomposing $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$, we can obtain two signal subspaces Υ_1 and Υ_2 respectively. Defining a [2Dx2D] matrix $\mathbf{C}$ from the two subspaces such

that

$$\mathbf{C} = \begin{bmatrix} \Upsilon_1^H \\ \Upsilon_2^H \end{bmatrix} \begin{bmatrix} \Upsilon_1 & \Upsilon_2 \end{bmatrix} = \Upsilon_C \, \Lambda \, \Upsilon_C^H \tag{2-46}$$

Υ_C is a [2Dx2D] matrix obtained by eigenvalue decomposition of $\mathbf{C}$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{2D}$ and $\Lambda = diag\ \{\lambda_1\ \lambda_2 \cdots \lambda_{2D}\}$ By partitioning Υ_C into four [DxD] submatrices such that

$$\Upsilon_C = \begin{bmatrix} \Upsilon_{11} & \Upsilon_{12} \\ \Upsilon_{21} & \Upsilon_{22}, \end{bmatrix} \tag{2-47}$$

the rotation operator can be estimated as $\Phi = -\Upsilon_{12}\, \Upsilon_{22}^{-1}$. From D eigenvalues of Φ, angles of arrival can be estimated as

$$\hat{\theta}_d = \sin^{-1} \left\{ \frac{\arg\ \lambda_d}{k\, \Delta} \right\}, \quad k = \frac{2\,\pi}{\lambda} \tag{2-48}$$

ESPRIT performs better than MUSIC and Capon algorithms for closely-spaced sources in terms of probability and RMSE, as can be seen in Figs. 2.10 and 2.11.

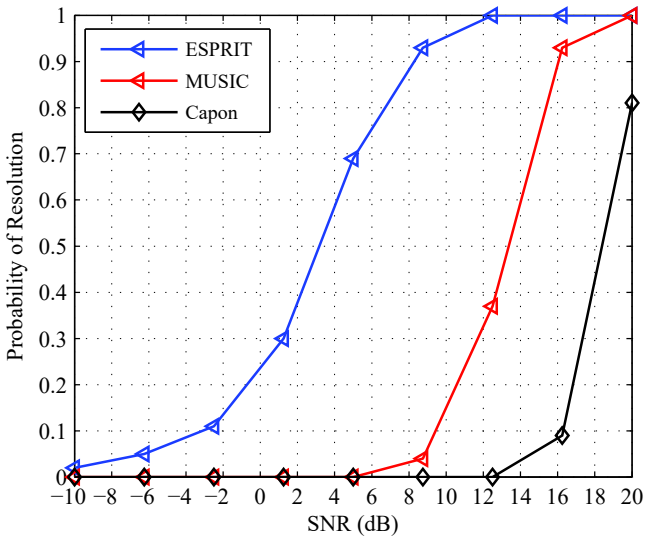

Figure 2.10: Probability of resolution versus SNR 2 uncorrelated sources, 2.8 degrees, ULA 10 sensors, 40 snapshots, 100 runs

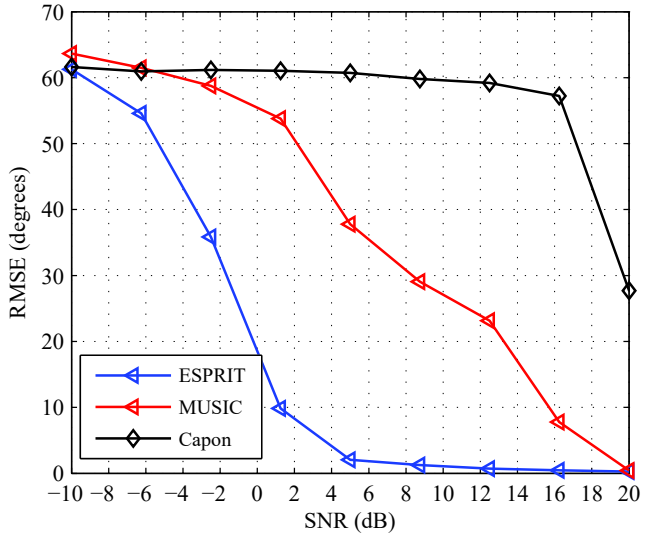

Figure 2.11: RMSE versus SNR, 2 uncorrelated sources, 2.8 degrees, ULA 10 sensors, 40 snapshots, 100 runs

2.1.6

Conjugate gradient

Let us now assume that P narrowband signals from far-field sources are impinging on a uniform linear array (ULA) of M ($M > P$) sensor elements with the unknown directions $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_P]^T$. The ith data snapshot of the $(M \times 1)$-dimensional array output vector can be modeled as

$$\boldsymbol{x}(i) = \boldsymbol{A}(\boldsymbol{\Theta})\boldsymbol{s}(i) + \boldsymbol{n}(i), \qquad i = 1, 2, \ldots, N, \tag{2-49}$$

where $\boldsymbol{s}(i) = [s_1(i), \ldots, s_P(i)]^T \in \mathbb{C}^{P\times 1}$ represents the zero-mean source data vector, $\boldsymbol{n}(i) \in \mathbb{C}^{M\times 1}$ is the vector of white circular complex Gaussian noise with zero mean and variance σ_n^2, and N denotes the number of available snapshots. The matrix $\boldsymbol{A}(\boldsymbol{\Theta}) = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_P)] \in \mathbb{C}^{M\times P}$ contains the array steering vectors $\boldsymbol{a}(\theta_j)$ corresponding to the nth source, which can be expressed as

$$\boldsymbol{a}(\theta_n) = [1, e^{j2\pi \frac{\Delta}{\lambda_c} \sin\theta_n}, \ldots, e^{j2\pi(M-1)\frac{\Delta}{\lambda_c}\sin\theta_n}]^T, \tag{2-50}$$

where $n = 1, \ldots, P$, Δ denotes the interelement spacing of the ULA and λ_c is the signal wavelength.

Using the fact that $\boldsymbol{s}(i)$ and $\boldsymbol{n}(i)$ are modeled as uncorrelated linearly independent variables, the $M \times M$ signal covariance matrix is calculated by

$$\boldsymbol{R} = \mathbb{E}\left[\boldsymbol{x}(i)\boldsymbol{x}^H(i)\right] = \boldsymbol{A}(\boldsymbol{\Theta})\boldsymbol{R}_{ss}\boldsymbol{A}^H(\boldsymbol{\Theta}) + \sigma_n^2\boldsymbol{I}_M, \tag{2-51}$$

where $\boldsymbol{R}_{ss} = \mathbb{E}[\boldsymbol{s}(i)\boldsymbol{s}^H(i)]$, which is diagonal if the sources are uncorrelated

and non-diagonal for partially correlated sources, and $\mathbb{E}[\boldsymbol{n}(i)\boldsymbol{n}^H(i)] = \sigma_n^2 \boldsymbol{I}_M$ with $\boldsymbol{I}_M$ being the $M \times M$ identity matrix. Since the true signal covariance matrix is unknown, it must be estimated and a widely-adopted approach is the sample average formula given by

$$\hat{\boldsymbol{R}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}(i)\boldsymbol{x}^H(i), \tag{2-52}$$

whose estimation accuracy is dependent on N.

The CG method [15] is used to minimize a cost function, or analogously, to solve a linear system of equations by approaching the optimal solution step by step via a line search along successive directions, which are sequentially determined at each direction [19]. As a result of the application of the CG algorithm to direction finding, we have a system of equations that is iteratively solved for $\mathbf{w}$ at each search angle:

$$\mathbf{R}\mathbf{w} = \mathbf{b}(\theta), \tag{2-53}$$

where $\mathbf{R}$ is the covariance matrix and $\mathbf{b}(\theta)$ is the initial vector defined as

$$\mathbf{b}(\theta) = \frac{\mathbf{R}\,\mathbf{a}(\theta)}{\|\mathbf{R}\,\mathbf{a}(\theta)\|}, \tag{2-54}$$

where $\mathbf{a}(\theta)$ is the search vector.

The extended signal subspace of rank P is obtained by means of the CG algorithm summarized in Table 2.1. The set of orthogonal residual vectors

Table 2.1: Conjugate Gradient Algorithm

$$\mathbf{w}_0 = 0,\ \mathbf{d}_1 = \mathbf{g}_{cg,0} = b,\ \rho_0 = \mathbf{g}_{cg,0}^H \mathbf{g}_{cg,0}$$

for i=1 to P do:

$$\begin{aligned}
\mathbf{v}_i &= \mathbf{R}\,\mathbf{d}_i \\
\alpha_i &= \rho_{i-1} \,/\, \mathbf{d}_i^H \mathbf{v}_i \\
\mathbf{w}_i &= \mathbf{w}_{i-1} + \alpha_i \mathbf{d}_i \\
\mathbf{g}_{cg,i} &= \mathbf{g}_{cg,i-1} - \alpha_i \mathbf{v}_i \\
\rho_i &= \mathbf{g}_{cg,i}^H \mathbf{g}_{cg,i} \\
\beta_i &= \rho_i \,/\, \rho_{i-1} = \|\mathbf{g}_{cg,i}\|^2 \,/\, \|\mathbf{g}_{cg,i-1}\|^2 \\
\mathbf{d}_{i+1} &= \mathbf{g}_{cg,i} + \beta_i \mathbf{d}_i
\end{aligned}$$

end for

form $\mathbf{G}_{cg,P+1}(\theta)$ (2-55)

compute $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ (2-57)

find $\hat{P}$ largest peaks of $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ to obtain estimates $\hat{\theta}_l$ of the DOA, $l = 1, 2, \cdots, P$

$$\mathbf{G}_{cg,P+1}(\theta) = [\mathbf{g}_{cg,0}(\theta), \mathbf{g}_{cg,1}(\theta), \ldots, \mathbf{g}_{cg,P}(\theta)], \tag{2-55}$$

where $\mathbf{b}(\theta)$= $\mathbf{g}_0(\theta)$ generates the well-known extended Krylov subspace comprised of the true signal subspace of dimension P and the search vector itself. All the residual vectors are normalized except for the last one. If $\theta \in \{\theta_1, \ldots \theta_P\}$, the initial vector $\mathbf{b}(\theta)$ lies in the true signal subspace space spanned by the
$[\mathbf{g}_{cg,0}(\theta), \mathbf{g}_{cg,1}(\theta), \ldots, \mathbf{g}_{cg,P-1}(\theta)]$ basis vectors of the extended Krylov subspace. Therefore, the rank of the generated signal subspace drops from P+1 to P and we have

$$\mathbf{g}_{cg,P}(\theta) = 0, \tag{2-56}$$

where $\mathbf{g}_{cg,P}$ is the last unnormalized residual vector.

In order to exploit this behavior, the proposed KA-CG algorithm makes use of the spectral function defined in [13]:

$$\mathcal{P}_{\mathcal{K}}(\theta^{(n)}) = \frac{1}{\|\mathbf{g}_{cg,P}^{H}(\theta^{(n)})\mathbf{G}_{cg,P+1}(\theta^{(n-1)})\|^2}, \tag{2-57}$$

where $\theta^{(n)}$denotes the search angle in the whole angle range $\{-90^o, \ldots, 90^o\}$ with $\theta^{(n)} = n\Delta^o - 90^o$, where Δ^o is the search step and $n = 0, 1, \ldots, 180^o/\Delta^o$. The matrix $\mathbf{G}_{cg,P+1}(\theta^{(n-1)})$ contains all residual vectors at the $(n-1)$th angle and $\mathbf{g}_{cg,P}(\theta^{(n)})$ is the last residual vector calculated at the current search step n. If $\theta^{(n)} \in \{\theta_1, \ldots, \theta_P\}$, $\mathbf{g}_{cg,P}(\theta^{(n)}) = 0$ and we can expect a peak in the spectrum. Taking into account that $\hat{\boldsymbol{R}}$ in (2-32) is only a sample average estimate, which is unknown in practical applications, $\mathbf{g}_{cg,P}(\theta^{(n)})$ and $\mathbf{G}_{cg,P+1}(\theta^{(n-1)})$ become approximations. Hence the spectral function in (2-57) can just provide very large values but they do not tend to infinity as for the original covariance matrix. In this specific scenario, in which the sources are separated by 2.8 degrees CG performs better than MUSIC and Capon algorithms for closely-spaced sources in terms of probability and RMSE, as can be seen in Figs. 2.10 and 2.11. In spite of performing worse than ESPRIT in this particular scenario, CG performance tends to overcome ESPRIT one as the space between the sources becomes smaller and the number of snapshots increases, as can be seen in the next sections.

2.1.7

Prior knowledge-based direction of arrival estimation

The problem of obtaining more accurate estimates by using prior-knowledge is closely related to the way of collecting it. The traditional approach to deal with

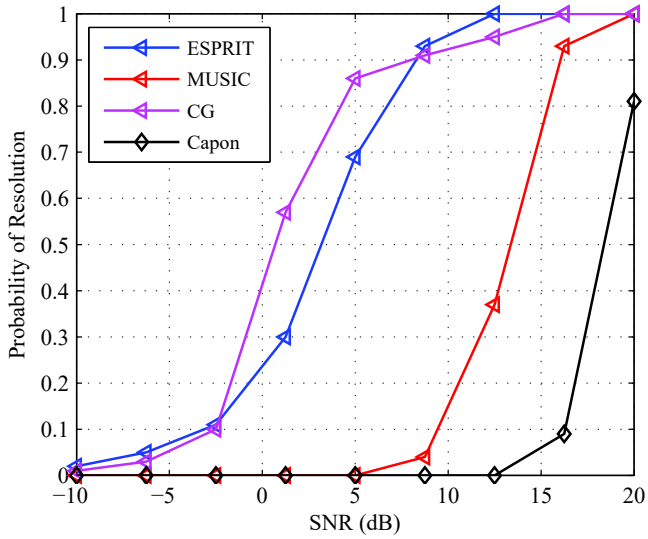

Figure 2.12: Probability of resolution versus SNR 2 uncorrelated sources, 2.8 degrees, ULA 10 sensors, 40 snapshots, 100 runs

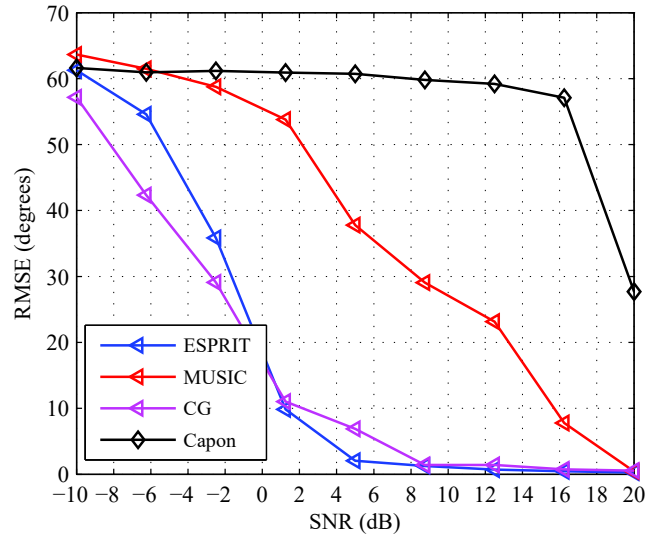

Figure 2.13: RMSE versus SNR, 2 uncorrelated sources, 2.8 degrees, ULA 10 sensors, 40 snapshots, 100 runs

this question makes use of available known DOAs to form a known covariance matrix to be optimally combined with the data covariance matrix, i.e., in a minimum mean squared error sense, resulting in an enhanced matrix which can be the basis for several DOA estimation methods. A new preliminary approach, which we have proposed in [30], instead of using that accessible known DOAs, makes use of previous estimates. As expected, in order to have good performance, the method requires that these previous estimates are sufficiently accurate. To achieve this goal, most methods may require a significant number of samples to obtain accurate statistical estimates. The

concept of prior-knowledge obtained on line applied to DOA estimation will be exploited in the methods discussed in the next chapters.

2.1.7.1

Problem formulation

Despite the numerous parameter estimation techniques developed over the last decades and their specific properties, advantages and drawbacks, their estimation accuracy depends on the $(M \times M)$ dimensional signal covariance matrix of the sensor array data vector $\boldsymbol{x}(i)$, which is defined for the ith snapshot as

$$\boldsymbol{R} = \mathbb{E}\left[\boldsymbol{x}(i)\boldsymbol{x}^H(i)\right], \qquad i = 1, \ldots, N, \tag{2-58}$$

where the superscript H and $\mathbb{E}[\cdot]$ denote the conjugate transpose and the statistical expectation respectively, and N is the number of available snapshots. In practice, the true signal covariance matrix in (2-58) is unknown, but can be estimated via the widely used sample-average formula given by

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}(i)\boldsymbol{x}^H(i). \tag{2-59}$$

Applying the covariance matrix estimate in (2-59), the estimation accuracy is essentially determined by the data record size N. Thus, in applications where the number of available sensors M is small, the increase in the number of snapshots become more significant.

In practical scenarios with low signal-to-noise ratio (SNR), stationary and non-stationary sources whose DOAs are to be estimated, the knowledge of the directions of strong consistent users can be effectively exploited in order to increase the estimation accuracy of non-stationary sources, which enter the system. The knowledge of previously estimated DOAs can be exploited in the form of a known covariance matrix $\boldsymbol{C}$. Knowledge-aided (KA) signal processing techniques, which make use of *a priori* knowledge of key parameters of interest such as the existence of strong interferers, cognitive users and geographical localization of users [20] have recently gained significant attention [21–25]. In KA techniques, the key issues are how to obtain *a priori* knowledge about the parameters of interest and how to exploit them. Prior work on KA algorithms has considered the design of space-time adaptive processing (STAP) techniques [21]- [22], [24, 25] and beamforming algorithms [23]. These methods have shown superior performance to conventional approaches that do not rely on KA techniques when the limited sample support is used in highly non-stationary environments.

Let us assume the same system model of 2.1.6. The main idea of prior knowledge-based direction of arrival estimation is to replace the sample covariance matrix with an enhanced one that linearly combines the original sample matrix with a known rank deficient matrix. In order to obtain an enhanced covariance matrix estimate $\tilde{\boldsymbol{R}}$, we assume that the *a priori* knowledge matrix $\boldsymbol{C}$ is nonrandom, according to [28], and perform a linear combination of $\boldsymbol{C}$ and the sample covariance matrix $\hat{\boldsymbol{R}}$ by applying the weight factors α and β, which are formulated as

$$\tilde{\boldsymbol{R}} = \alpha \boldsymbol{C} + \beta \hat{\boldsymbol{R}}, \tag{2-60}$$

where the combination factors are constrained to $\alpha > 0$ and $\beta > 0$, and $\boldsymbol{C}$ is restricted to be positive semidefinite to ensure that $\tilde{\boldsymbol{R}}$ is also positive semidefinite. Then it is possible to find optimal estimates of the weight factors α and β, which efficiently combine $\boldsymbol{C}$ and $\hat{\boldsymbol{R}}$ depending on the scenario. One of the most common criteria is the minimization of the parameters in a mean squared error (MSE) sense, that is

$$\begin{aligned} \min_{\alpha,\beta} \quad & \text{MSE} = \mathbb{E}\left[\|\tilde{\boldsymbol{R}} - \boldsymbol{R}\|_F^2\right] \\ & \text{s.t. } \tilde{\boldsymbol{R}} = \alpha \boldsymbol{C} + \beta \hat{\boldsymbol{R}}, \end{aligned} \tag{2-61}$$

where $\|\cdot\|_F$ denotes the Frobenius matrix norm. Note that the optimization problem is solved by minimizing the MSE with respect to the two parameters α and β, which as expected depend on each other and the unknown true covariance matrix $\boldsymbol{R}$. Another widely used criterion to reduce the complexity of the optimization problem, which can be considered a special case of the function (2-61), is the optimization described by (2-62):

$$\begin{aligned} \min_{\alpha} \quad & \text{MSE} = \mathbb{E}\left[\|\tilde{\boldsymbol{R}} - \boldsymbol{R}\|_F^2\right] \\ & \text{s.t. } \tilde{\boldsymbol{R}} = \alpha \boldsymbol{C} + (1-\alpha) \hat{\boldsymbol{R}} \end{aligned} \tag{2-62}$$

with α being restricted to $\alpha \in (0, 1)$ to ensure the positive semi-definiteness of $\tilde{\boldsymbol{R}}$. Both types of optimization are briefly discussed and applied to the simulations in the next topics.

2.1.7.2

Computation of the optimal weight factors

Assuming the knowledge of the DOAs of k signals that are impinging on the array from the known directions $\bar{\theta} = [\theta_1, \ldots, \theta_k]^T$, the *a priori* covariance matrix $\boldsymbol{C}$ can be calculated by

$$\boldsymbol{C} = \sum_{l=1}^{k} \boldsymbol{a}(\theta_l)\boldsymbol{a}^H(\theta_l)\sigma_l, \tag{2-63}$$

where $\boldsymbol{a}(\theta_l)$ is the array steering vector of the lth known DOA and σ_l is the power of the lth signal.

Let α_0 and β_0 denote the optimal values α and β that satisfy (7) and (8). The estimates $\hat{\alpha_0}$ and $\hat{\beta}_0$, of α_0 and β_0, obtained from the available data, can be compactly expressed by means of two approaches, as follows:

KA-General Linear Combination

In this case, the estimates given in (2-64) and (2-65) are the two weight factors

$$\hat{\beta}_{\mathrm{o}} = \frac{\hat{\gamma}}{\hat{\rho} + \hat{\gamma}}, \tag{2-64}$$

$$\hat{\alpha}_{\mathrm{o}} = \hat{\nu}(1 - \hat{\beta}_{\mathrm{opt}}), \tag{2-65}$$

to be applied to 2-60 and $\hat{\gamma}$, $\hat{\nu}$, and $\hat{\rho}$ are defined as

$$\hat{\gamma} = \|\hat{\nu}\boldsymbol{C} - \hat{\boldsymbol{R}}\|_F^2, \tag{2-66}$$

$$\hat{\nu} = \frac{\mathrm{Tr}\{\boldsymbol{C}^H\hat{\boldsymbol{R}}\}}{\|\boldsymbol{C}\|_F^2}, \tag{2-67}$$

$$\hat{\rho} = \frac{1}{N^2}\sum_{i=1}^{N}\|\boldsymbol{x}(i)\|_F^4 - \frac{1}{N}\|\hat{\boldsymbol{R}}\|_F^2. \tag{2-68}$$

KA-Convex Combination

In this case, the estimate given in (2-69) is the the sole weight factor

$$\hat{\alpha_0} = \frac{\hat{\rho}}{\hat{\rho} + \|\hat{\boldsymbol{R}} - \boldsymbol{C}\|_F^2}, \tag{2-69}$$

to be applied to 2-62, and $\hat{\rho}$ is defined as (2-68).

2.1.7.3

Knowledge-aided methods

Despite the existence of KA methods combined with classical algorithms for parameter estimation like MUSIC and ESPRIT, KA methods have not been combined with high-resolution source localization algorithms like CG so far. In order to fulfill the expectations of works in this specific field of research, as a preliminary study, we formulated a new knowledge-aided parameter estimation technique, termed as KAv-CG [30], that combines the (CG) algorithm [15,16,27] and *a priori* knowledge of the directions of arrivals of source signals. Since KAv-CG follows the same approach employed in MUSIC and ESPRIT, without loss of generality it can give an idea of how the KAv-algorithm class works.

Knowledge-aided CG algorithm based on available known DOA

The KAv-CG algorithm, which we developed for complex-valued data, follows a similar approach to [12] and considers the general case [18], where the knowledge-aided covariance $\boldsymbol{C}$ (2-63) is rank deficient and the noise power is assumed to be unknown. This new KAv-CG method also replaces the original sample covariance matrix with an enhanced covariance matrix, which is a combination of the original weighted sample covariance matrix and a weighted knowledge-aided covariance matrix, similarly to the existing KAv applied to subspace-based methods [12]. KAv-CG can be summarized as shown in Table 2.2.

In order to develop KAv-CG, we make use of the same system model of 2.1.6 also used in 2.1.7.1, in which it is assumed the knowledge of the DOAs of k signals that are impinging on the array from the known directions $\bar{\theta} = [\theta_1, \ldots, \theta_k]^T$.

The aim of the proposed KAv-CG algorithm is to exploit *a priori* knowledge in the form of the enhanced signal covariance matrix $\tilde{\boldsymbol{R}}$ in (2-60) and process it using CG algorithm. As can be seen in [18], one can calculate the *a priori* covariance matrix $\boldsymbol{C}$ in (2-63) by means of the steering vectors in (2-50) based on the known directions of impinging signals. The proposed alternative method is composed of two stages. The first stage encompasses three substeps.The first substep is to calculate the *a priori* covariance matrix $\boldsymbol{C}$, using the steering vectors of known DOAs. The second substep is to compute the weight

Table 2.2: Proposed KAv-Conjugate Gradient Algorithm

Inputs:

$M,\ d,\ \lambda,\ N,\ P$

Received vectors $\boldsymbol{x}(1),\ \boldsymbol{x}(2),\cdots,\ \boldsymbol{x}(N)$

Prior knowledge $\rightarrow$ known DOAs: $\theta_1, \theta_2, \cdots;\ \theta_k, \quad 1 \leq k < P$

Outputs:

Estimates $\hat{\theta}_{k+1},\ \hat{\theta}_{k+2},\cdots,\ \hat{\theta}_P$

First stage:

compute $\boldsymbol{C}$ (2-63), for the k known DOAs

compute $\hat{\alpha_0}$ (2-69) for convex combination or $\hat{\beta}_0$ (2-64) and $\hat{\alpha_0}$ (2-65) for general linear combination

compute $\tilde{\boldsymbol{R}}$ (2-61) or (2-62) according to the combination in use as previously mentioned

Second stage:

$$\mathbf{w}_0 = 0,\ \mathbf{d}_1 = \mathbf{g}_{cg,0} = b,\ \rho_0 = \mathbf{g}_{cg,0}^H \mathbf{g}_{cg,0}$$

for i=1 to P do:

$$\begin{aligned}
\mathbf{v}_i &= \tilde{\mathbf{R}}\,\mathbf{d}_i \\
\alpha_i &= \rho_{i-1} \,/\, \mathbf{d}_i^H \mathbf{v}_i \\
\mathbf{w}_i &= \mathbf{w}_{i-1} + \alpha_i \mathbf{d}_i \\
\mathbf{g}_{cg,i} &= \mathbf{g}_{cg,i-1} - \alpha_i \mathbf{v}_i \\
\rho_i &= \mathbf{g}_{cg,i}^H \mathbf{g}_{cg,i} \\
\beta_i &= \rho_i \,/\, \rho_{i-1} = \|\mathbf{g}_{cg,i}\|^2 \,/\, \|\mathbf{g}_{cg,i-1}\|^2 \\
\mathbf{d}_{i+1} &= \mathbf{g}_{cg,i} + \beta_i \mathbf{d}_i
\end{aligned}$$

end for

form $\mathbf{G}_{cg,P+1}(\theta)$ (2-55)

compute $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ (2-57)

find $\hat{P}$ largest peaks of $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ to obtain estimates $\hat{\theta}_k$ of the DOA

DOAs $= \{\theta_1, \cdots,\ \theta_k, \hat{\theta}_{k+1}, \cdots, \hat{\theta}_P\}$

factors according to the combination to be applied. For KA-General Linear Combination (KA-GLC), the factors are $\hat{\alpha}_{\mathrm{o}}$ (2-65) and $\hat{\beta}_{\mathrm{o}}$ and (2-64). In the case of KA-Convex Combination (KA-CC), $\hat{\alpha_0}$ is given by (2-69). Our proposed KA-CG algorithm makes use of the latter approach. In both cases, the covariance matrix to be applied to the first stage of the KA-CG algorithm is obtained by the sample average formula given in (2-59). The third substep is to calculate the enhanced covariance matrix $\tilde{\boldsymbol{R}}$ in (2-61) or (2-62) according to the chosen combination. The last stage includes two substeps. The first is to

compute the estimates by processing the CG algorithm described in 2.1.6 and 2.1. For this purpose, we use the enhanced signal covariance matrix $\tilde{\boldsymbol{R}}$ instead of the sample covariance matrix $\hat{R}$ (2-32). The last substep is to form the solution set containing the subset including the known DOAs and the subset encompassing the estimates of the unknown DOAs.

Simulations

The simulations of MUSIC, ESPRIT, CG and their knowledge-aided versions, termed KAv-ESPRIT, KAv-MUSIC and KAv-CG, respectively, are based upon a scenario with $P = 2$ two uncorrelated closely-spaced signals at $(89.05, 90.95)^\circ$ impinging on a ULA with $M = 12$ sensors equally spaced by half wavelength. The sample matrix (2-59) has been computed with 180 snapshots and the simulated curves are obtained by averaging the results over 200 independent trials. The a priori covariance matrix $\boldsymbol{C}$ has been obtained using the steering vectors of the second DOA, supposed to be known. In order to assess the accuracy in terms of probability of resolution, we take into account the criterion [29], [13], in which two sources with DOA θ_1 and θ_2 are said to be resolved if their respective estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are such that both $\left|\hat{\theta}_1 - \theta_1\right|$ and $\left|\hat{\theta}_2 - \theta_2\right|$ are less than $|\theta_1 - \theta_2|/2$.

Figure 2.14 depicts the probability of resolution of CG, KAv-CG, MUSIC, KAv-MUSIC, ESPRIT and KAv-ESPRIT. It can be noticed tha the original versions are outperformed by their KAv versions, respectively. It can be noticed that KAv-CG outperforms KAv-ESP in $\simeq [0.3 \;\; 7.0\, dB]$. The gap between the corresponding versions make clear the potential of the original versions to be exploited in terms of probability of resolution. The KAv-versions of CG, MUSIC and ESPRIT can be viewed as upper bounds of their original versions, i.e, CG, MUSIC and ESPRIT.

Knowledge-aided CG algorithm based on DOAs obtained on line

The previous approaches to process knowledge make use of available known DOAs to yield the*a priori* knowledge matrix $\boldsymbol{C}$, responsible for upgrading the sample covariance matrix. Now, we will discuss a new approach to knowledge-aided parameter estimation based on CG algorithm called KA-CG as mentioned in 2.1.7. This approach, is summarized in Table 2.3 and, similarly to KAv-CG, has been presented in [30]. The main idea of this algorithm,

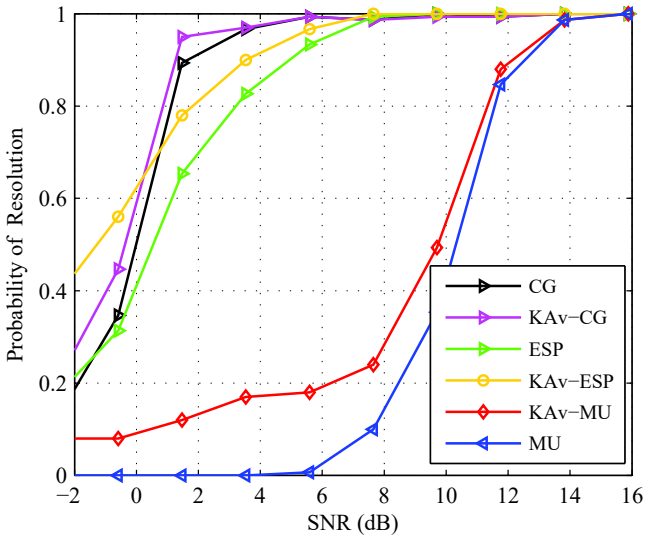

Figure 2.14: Probability of resolution of CG, MUSIC, ESPRIT and their knowledge-aided versions versus SNR with $M = 12$, $N = 180$, $P = 2$, $L = 200$ runs, unknown uncorrelated sources at $(89.05, 90.95)^\circ$

which is composed of three stages, is to process the CG algorithm described in Table 2.1 twice. The first time in which CG is processed aims to obtain initial estimates, from which some are considered as if they were true DOAs. These estimates are the basis for computing the a priori knowledge matrix $\boldsymbol{C}$ in the second stage. Thus, differently from KAv versions, which make use of available known DOAs to form $\boldsymbol{C}$, this procedure can be considered knowledge acquisition on line since that matrix is calculated by using steering vectors of previous estimates. The last time in which CG is processed has the purpose of computing the final DOA estimates making use of the enhanced covariance matrix estimate $\tilde{\boldsymbol{R}}$ (2-61) or (2-62).

Simulations

Let us now evaluate the performance of our proposed Knowledge-Aided Conjugate Gradient (KA-CG) algorithm for direction finding and localization techniques. Specifically, we evaluate the probability of resolution of two adjacent signals. For this purpose, we compare the KA-CG, the KA-ESPRIT and the KA-MUSIC, where the *a priori* covariance matrices $\boldsymbol{C}$ (2-63) are based on estimates, to their original versions and also to their KAv versions, in which $\boldsymbol{C}$ is constructed with known DOAs. All experiments are based upon a scenario identical to the simulations used for MUSIC, KAv-MUSIC, ESPRIT and KAv-ESPRIT, i.e., with $P = 2$ two uncorrelated closely-spaced signals at $(89.05, 90.95)^\circ$ impinging on a ULA with $M = 12$ sensors equally

Table 2.3: Proposed KA-Conjugate Gradient Algorithm

First stage:

$$\mathbf{w}_0 = 0,\ \mathbf{d}_1 = \mathbf{g}_{cg,0} = b,\ \rho_0 = \mathbf{g}_{cg,0}^H \mathbf{g}_{cg,0}$$

for i=1 to P do:

$$\begin{aligned}
\mathbf{v}_i &= \mathbf{R}\,\mathbf{d}_i \\
\alpha_i &= \rho_{i-1} \,/\, \mathbf{d}_i^H \mathbf{v}_i \\
\mathbf{w}_i &= \mathbf{w}_{i-1} + \alpha_i \mathbf{d}_i \\
\mathbf{g}_{cg,i} &= \mathbf{g}_{cg,i-1} - \alpha_i \mathbf{v}_i \\
\rho_i &= \mathbf{g}_{cg,i}^H \mathbf{g}_{cg,i} \\
\beta_i &= \rho_i \,/\, \rho_{i-1} = \|\mathbf{g}_{cg,i}\|^2 \,/\, \|\mathbf{g}_{cg,i-1}\|^2 \\
\mathbf{d}_{i+1} &= \mathbf{g}_{cg,i} + \beta_i \mathbf{d}_i
\end{aligned}$$

end for

form $\mathbf{G}_{cg,P+1}(\theta)$ (2-55)

compute $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ (2-57)

find $\hat{P}$ largest peaks of $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ to obtain estimates $\hat{\theta}_l$ of the DOA

Second stage:

compute $\boldsymbol{C}$ (2-63), for $\theta_l = \hat{\theta}_l$, $l < P$

compute $\hat{\alpha_0}$ (2-69) for convex combination or $\hat{\beta}_0$ (2-64) and $\hat{\alpha_0}$ (2-65) for general linear combination

compute $\tilde{\boldsymbol{R}}$ (2-61) or (2-62) according to the combination in use as previously mentioned

Last stage:

Repeat the first stage to obtain enhanced estimates of DOA making use of $\tilde{\boldsymbol{R}}$ instead of $\hat{\mathbf{R}}$

spaced by half wavelength. The sample matrix (2-59) is computed with 180 snapshots and the simulated curves are obtained by averaging the results over 200 independent trials. In order to assess the accuracy in terms of probability of resolution, we take into account the same criterion described in 2.1.7.3.

In our experiment, we compare the probability of resolution of the KA-CG, KA-ESPRIT and KA-MUSIC, to their KAv versions, in which the *a priori* covariance matrices $\boldsymbol{C}$ (2-63) are obtained from the steering vector of the second DOA, which is supposed to be known. The covariance matrices $\boldsymbol{C}$ (2-63) $\boldsymbol{C}$ for KA versions were obtained from the steering vector of the estimate of the second DOA. The results depicted in Fig.2.15 show the best performance of each KAv version, in which $\boldsymbol{C}$ is obtained by the configuration with one known DOA, over its KA version, where $\boldsymbol{C}$ is calculated using one of the estimates.

Each KAv-version can be considered an upper bound of its KA version. Thus, it can be noticed that the small area limited by KA-CG and KAv-CG shows that the former already exploits its potential close to the effective optimal performance. The gap available to improvements is situated within [-1.9, 1.6] SNR(dB) where the probability of resolution is lower than 0.88. Differently from the previous KA-CG case, there is a larger area limited by KA-ESPRIT and KAv-ESPRIT that is available to enhancements. It can also be seen that their effective optimal performance (KAv-ESPRIT) is outperformed by both KAv-CG and KA-CG. The area limited by KA-MUSIC and KAv-MUSIC shows that most of the potential to be exploited is situated at the lower levels of the probability of resolution and that the potential of improvement of the KA-MUSIC is poor at higher ones.

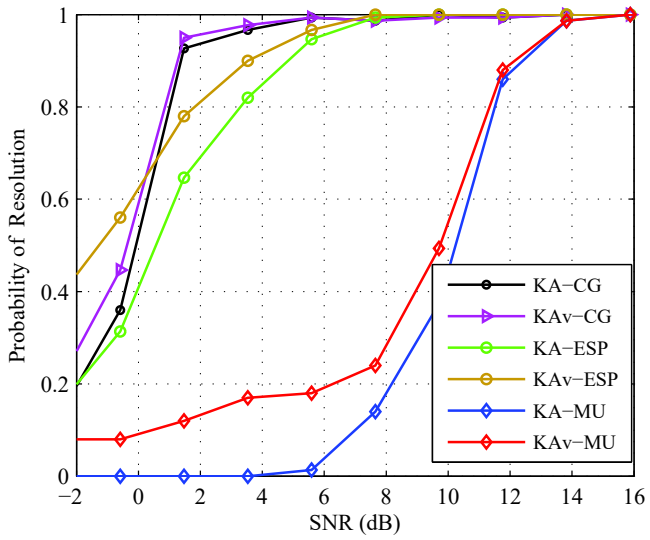

Figure 2.15: Probability of resolution of the KA versions and KAv versions of CG, ESPRIT and MUSIC versus SNR with $M = 12$, $N = 180$, $P = 2$, $L = 200$ runs, unknown uncorrelated sources at $(89.05, 90.95)^\circ$

3

Multi-Step Knowledge-aided Iterative ESPRIT algorithm

3.1

Introduction

Traditional high-resolution methods for DOA estimation such as the multiple signal classification (MUSIC) method [8], the root-MUSIC algorithm [9], the estimation of signal parameters via rotational invariance techniques (ESPRIT) [10] and other recent subspace techniques [20,31,32] exploit the eigenstructure of the input data matrix. These techniques may fail for reduced data sets or low signal-to-noise ratio (SNR) levels where the expected estimation error is not asymptotic to the Cramér-Rao bound (CRB) [33]. The accuracy of the estimates of the covariance matrix is of fundamental importance in parameter estimation. Low levels of SNR or short data records can result in significant divergences between the true and the sample data covariance matrices. In practice, only a modest number of data snapshots is available and when the number of snapshots is similar to the number of sensor array elements, the estimated and the true subspaces can differ significantly. Several approaches have been developed with the aim of enhancing the computation of the covariance matrix [34–43].

Diagonal loading [34] and shrinkage [35–37] techniques can enhance the estimate of the data covariance matrix by weighing and individually increasing its diagonal by a real constant. Nevertheless, the eigenvectors remain the same, which leads to unaltered estimates of the signal and noise projection matrices obtained from the enhanced covariance matrix. Additionally, an improvement of the estimates of the covariance matrix can be achieved by employing forward/backward averaging and spatial smoothing approaches [38, 39]. The former leads to twice the number of the original samples and its corresponding enhancement. The latter extracts the array covariance matrix as the average of all covariance matrices from its sub-arrays, resulting in a greater number of samples. Both techniques are employed in signal decorrelation. An approach to

improve MUSIC dealing with the condition in which the number of snapshots and the sensor elements approach infinity was presented in [40]. Nevertheless, this technique is not that effective for reduced number of snapshots. Other approaches to deal with reduced data sets or low SNR levels [41, 43] consist of reiterating the procedure of adding pseudo-noise to the observations which results in new estimates of the covariance matrix. Then, the set of solutions is computed from previously stored DOA estimates. In [44], two aspects resulting from the computation of DOAs for reduced data sets or low SNR levels have been studied using the root-MUSIC technique. The first aspect dealt with the probability of estimated signal roots taking a smaller magnitude than the estimated noise roots, which is an anomaly that leads to wrong choices of the closest roots to the unit circle. To mitigate this problem, different groups of roots are considered as potential solutions for the signal sources and the most likely one is selected [45]. The second aspect previously mentioned, shown in [46], refers to the fact that a reduced part of the true signal eigenvectors exists in the sample noise subspace (and vice-versa). Such coexistence has been expressed by a Frobenius norm of the related irregularity matrix and introduced its mathematical foundation. An iterative technique to enhance the efficacy of root-MUSIC by reducing this anomaly making use of the gradual reshaping of the sample data covariance matrix has been reported. Inspired by the work in [44], we have developed an ESPRIT-based method known as Two-Step KAI-ESPRIT (TS-ESPRIT) [47], which combines that modifications of the sample data covariance matrix with the use of prior knowledge [12, 21, 48–50] about the covariance matrix of a set of impinging signals to enhance the estimation accuracy in the finite sample size region. In practice, this prior knowledge could be from the signals coming from known base stations or from static users in a system. TS-ESPRIT determines the value of a correction factor that reduces the undesirable terms in the estimation of the signal and noise subspaces in an iterative process, resulting in better estimates.

In [51], we presented preliminary results of the Multi-Step KAI ESPRIT (MS-KAI-ESPRIT) approach that refines the covariance matrix of the input data via multiple steps of reduction of its undesirable terms. This chapter presents the MS-KAI-ESPRIT in further detail, an analysis of the mean squared error (MSE) of the data covariance matrix free of undesired terms (side effects), a more accurate study of the computational complexity and a comprehensive study of MS-KAI-ESPRIT and other competing techniques for scenarios with both uncorrelated and correlated signals. Unlike TS-ESPRIT, which makes use of only one iteration and available known DOAs, MS-KAI-

ESPRIT employs multiple iterations and obtains prior knowledge on line. At each iteration of MS-KAI-ESPRIT, the initial Vandermonde matrix is updated by replacing an increasing number of steering vectors of initial estimates with their corresponding refined versions. In other words, at each iteration, the knowledge obtained on line is updated, allowing the direction finding algorithm to correct the sample covariance matrix estimate, which yields more accurate estimates.

3.2

Proposed MS-KAI-ESPRIT Algorithm

In this section, we present the proposed MS-KAI-ESPRIT algorithm [52] and detail its main features. For this purpose we make use of the same system model described in 2.1.7.1. We start by expanding (2-52) using (2-49), as derived in [44]:

$$\begin{aligned}\hat{\boldsymbol{R}} &= \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{A}\,s(i)+\boldsymbol{n}(i))\,(\boldsymbol{A}\,s(i)+\boldsymbol{n}(i))^H \\ &= \boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{n}^H(i) \\ &+ \underbrace{\boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{n}^H(i)\right\} + \left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H}_{\text{"undesirable terms"}}\end{aligned} \tag{3-1}$$

The first two terms of $\hat{\boldsymbol{R}}$ in (4-20) can be considered as estimates of the two summands of $\boldsymbol{R}$ given in (2-51), which represent the signal and the noise components, respectively. The last two terms in (4-20) are undesirable side effects, which can be seen as estimates for the correlation between the signal and the noise vectors. The system model under study is based on noise vectors which are zero-mean and also independent of the signal vectors. Thus, the signal and noise components are uncorrelated to each other. As a consequence, for a large enough number of samples N, the last two terms of (4-20) tend to zero. Nevertheless, in practice the number of available samples can be limited. In such situations, the last two terms in (4-20) may have significant values, which causes the deviation of the estimates of the signal and the noise subspaces from the true signal and noise subspaces.

The key point of the proposed MS-KAI-ESPRIT algorithm is to modify the sample data covariance matrix estimate at each iteration by gradually incorporating the knowledge provided by the newer Vandermonde matrices

which progressively embody the refined estimates from the preceding iteration. Based on these updated Vandermonde matrices, refined estimates of the projection matrices of the signal and noise subspaces are calculated. These estimates of projection matrices associated with the initial sample covariance matrix estimate and the reliability factor are employed to reduce its side effects and allow the algorithm to choose the set of estimates that has the highest likelihood of being the set of the true DOAs. The modified covariance matrix is computed by computing a scaled version of the undesirable terms of $\hat{\boldsymbol{R}}$, as pointed out in (4-20).

The steps of the proposed algorithm are listed in Table 3.1. The algorithm starts by computing the sample data covariance matrix (2-59). Next, the DOAs are estimated using the ESPRIT algorithm. The superscript $(\cdot)^{(1)}$ refers to the estimation task performed in the first step. Now, a procedure consisting of $n = 1 : P$ iterations starts by forming the Vandermonde matrix using the DOA estimates. Then, the amplitudes of the sources are estimated such that the square norm of the differences between the observation vector and the vector containing estimates and the available known DOAs is minimized. This problem can be formulated [44] as:

$$\hat{\boldsymbol{s}}(i) = \arg\min_{\boldsymbol{s}} \parallel \boldsymbol{x}(i) - \hat{\boldsymbol{A}}\mathbf{s} \parallel_2^2 . \tag{3-2}$$

The minimization of (5-15) is achieved using the least squares technique and the solution is described by

$$\hat{\boldsymbol{s}}(i) = (\hat{\mathbf{A}}^H \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}} \boldsymbol{x}(i) \tag{3-3}$$

The noise component is then estimated as the difference between the estimated signal and the observations made by the array, as given by

$$\hat{\boldsymbol{n}}(i) = \boldsymbol{x}(i) - \hat{\boldsymbol{A}} \hat{\boldsymbol{s}}(i). \tag{3-4}$$

After estimating the signal and noise vectors, the third term in (4-20) can be

computed as:

$$\begin{aligned}\boldsymbol{V} &\triangleq \hat{\boldsymbol{A}}\left\{\frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{s}}(i)\hat{\mathbf{n}}^{H}(i)\right\}\\ &= \hat{\boldsymbol{A}}\left\{\frac{1}{N}\sum_{i=1}^{N}(\hat{\mathbf{A}}^{H}\,\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}^{H}\boldsymbol{x}(i)\right.\\ &\left.\times(\boldsymbol{x}^{H}(i)-\boldsymbol{x}^{H}(i)\hat{\mathbf{A}}(\hat{\mathbf{A}}^{H}\hat{\mathbf{A}})^{-1}\,\hat{\mathbf{A}}^{H})\right\}\\ &= \hat{\mathbf{Q}}_{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}(i)\boldsymbol{x}^{H}(i)\,\left(\mathbf{I}_{M}\,-\,\hat{\mathbf{Q}}_{A}\right)\right\}\\ &= \hat{\mathbf{Q}}_{A}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_{A}^{\perp},\end{aligned} \tag{3-5}$$

where

$$\hat{\mathbf{Q}}_{A} \triangleq \hat{\mathbf{A}}\,(\hat{\mathbf{A}}^{H}\,\hat{\mathbf{A}})^{-1}\,\hat{\mathbf{A}}^{H} \tag{3-6}$$

is an estimate of the projection matrix of the signal subspace, and

$$\hat{\mathbf{Q}}_{A}^{\perp} \triangleq \mathbf{I}_{M}\,-\,\hat{\mathbf{Q}}_{A} \tag{3-7}$$

is an estimate of the projection matrix of the noise subspace.

Next, as part of the procedure consisting of $n = 1 : P$ iterations, the modified data covariance matrix $\hat{\mathbf{R}}^{(n+1)}$ is obtained by computing a scaled version of the estimated terms from the initial sample data covariance matrix as given by

$$\hat{\mathbf{R}}^{(n+1)} = \hat{\mathbf{R}}\,-\,\mu\,(\mathbf{V}^{(n)}\,+\,\mathbf{V}^{(n)H}), \tag{3-8}$$

where the superscript $(\cdot)^{(n)}$ refers to the n^{th} iteration performed. The scaling or reliability factor μ increases from 0 to 1 incrementally, resulting in modified data covariance matrices. Each of them gives origin to new estimated DOAs also denoted by the superscript $(\cdot)^{(n+1)}$ by using the ESPRIT algorithm, as briefly described ahead.

The rank P is assumed to be known, which is an assumption frequently found in the literature. Alternatively, the rank P could be estimated by model-order selection schemes [53] such as Akaike's Information Theoretic Criterion (AIC) [54] and the Minimum Descriptive Length (MDL) Criterion [55].

In order to estimate the signal and the orthogonal subspaces from the data records, we may consider two approaches [56,57]: the direct data approach and the covariance approach. The direct data approach makes use of singular value decomposition(SVD) of the data matrix $\mathbf{X}$, composed of the ith data snapshot

(2-49) of the M-dimensional array data vector:

$$\begin{aligned}\boldsymbol{X} =&[\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(N)] \\ =&\boldsymbol{A}[\boldsymbol{s}(1), \boldsymbol{s}(2), \ldots, \boldsymbol{s}(N)] + [\boldsymbol{n}(1), \boldsymbol{n}(2), \ldots, \boldsymbol{n}(N)] \\ =&\boldsymbol{A}(\boldsymbol{\Theta})\ \boldsymbol{S} + \ \boldsymbol{N}\ \in \mathbb{C}^{M\times N}\end{aligned} \tag{3-9}$$

Since the number of the sources is assumed known or can be estimated by AIC [54] or MDL [55] , as previously mentioned, we can write $\mathbf{X}$ as:

$$\boldsymbol{X}\ = \begin{bmatrix}\ \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n\ \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Gamma}}_s & 0 \\ 0 & \hat{\mathbf{\Gamma}}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{bmatrix}, \tag{3-10}$$

where the diagonal matrices $\hat{\mathbf{\Gamma}}_s$ and $\hat{\mathbf{\Gamma}}_n$ contain the P largest singular values and the $M-P$ smallest singular values, respectively. The estimated signal subspace $\hat{\mathbf{U}}_s \in \mathbb{C}^{M\times P}$ consists of the singular vectors corresponding to $\hat{\mathbf{\Gamma}}_s$ and the orthogonal subspace $\hat{\mathbf{U}}_n \in \mathbb{C}^{M\times(M-P)}$ is related to $\hat{\mathbf{\Gamma}}_n$. If the signal subspace is estimated, a rank-P approximation of the SVD can be applied.

The covariance approach applies the eigenvalue decomposition (EVD) of the sample covariance matrix (2-59), which is related to the data matrix (3-9):

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}(i)\boldsymbol{x}^H(i) = \frac{1}{N}\boldsymbol{X}\boldsymbol{X}^H\ \in \mathbb{C}^{M\times M} \tag{3-11}$$

Then, the EVD of (3-11) can be carried out as follows:

$$\hat{\mathbf{R}}\ = \begin{bmatrix}\ \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n\ \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Lambda}}_s & 0 \\ 0 & \hat{\mathbf{\Lambda}}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{bmatrix}, \tag{3-12}$$

where the diagonal matrices $\hat{\mathbf{\Lambda}}_\mathbf{s}$ and $\hat{\mathbf{\Lambda}}_\mathbf{n}$ contain the P largest and the M-P smallest eigenvalues, respectively. The estimated signal subspace $\hat{\mathbf{U}}_s \in \mathbb{C}^{M\times P}$ corresponding to $\hat{\mathbf{\Gamma}}_\mathbf{s}$ and the orthogonal subspace $\hat{\mathbf{U}}_n \in \mathbb{C}^{M\times(M-P)}$ complies with $\hat{\mathbf{\Gamma}}_\mathbf{n}$. If the signal subspace is estimated, a rank-P approximation of the EVD can be applied. With infinite precision arithmetic, both SVD and EVD can be considered equivalent. However, as in practice, finite precision arithmetic is employed, 'squaring' the data to obtain the Gramian $\boldsymbol{X}\boldsymbol{X}^H$ (3-11) can result in round-off errors and overflow. These are potential problems to be aware when using the covariance approach.

Now, we can briefly review ESPRIT. We start by forming a twofold subarray configuration, as each row of the array steering matrix $\boldsymbol{A}(\boldsymbol{\Theta})$ corresponds to

one sensor element of the antenna array. The subarrays are specified by two $(s \times M)$-dimensional selection matrices $\mathbf{J_1}$ and $\mathbf{J_2}$ which choose s elements of the M existing sensors, respectively, where s is in the range $P \leq s < M$. For maximum overlap, the matrix $\mathbf{J_1}$ selects the first $s = M - 1$ elements and the matrix $\mathbf{J_2}$ selects the last $s = M - 1$ rows of $\boldsymbol{A}(\boldsymbol{\Theta})$.

Since the matrices $\mathbf{J_1}$ and $\mathbf{J_2}$ have now been computed, we can estimate the operator $\boldsymbol{\Psi}$ by solving the approximation of the shift invariance equation (3-13) given by

$$\mathbf{J}_1 \, \hat{\mathbf{U}}_s \, \boldsymbol{\Psi} \; \approx \; \mathbf{J}_2 \, \hat{\mathbf{U}}_s, \tag{3-13}$$

where $\hat{U}_s$ is obtained in (3-12).

Using the least squares (LS) method, which yields

$$\hat{\boldsymbol{\Psi}} = \arg\min_{\boldsymbol{\Psi}} \| \, \mathbf{J}_2 \, \hat{\mathbf{U}}_s \; - \; \mathbf{J}_1 \, \hat{\mathbf{U}}_s \, \boldsymbol{\Psi} \, \|_F = \left(\mathbf{J}_1 \, \hat{\mathbf{U}}_s \right)^{\dagger} \; \mathbf{J}_2 \, \hat{\mathbf{U}}_s, \tag{3-14}$$

where $\| \cdot \|_F$ denotes the Frobenius norm and $(\cdot)^{\dagger}$ stands for the pseudo-inverse.

Lastly, the eigenvalues λ_i of $\hat{\boldsymbol{\Psi}}$ contain the estimates of the spatial frequencies γ_i computed as:

$$\gamma_i \; = \; \arg\left(\lambda_i\right), \tag{3-15}$$

so that the DOAs can be calculated as:

$$\hat{\theta}_i \; = \; \arcsin\left(\frac{\gamma_i \, \lambda_c}{2\pi \; d} \right) \tag{3-16}$$

where for (3-15) and (3-16) $\mathrm{i} = 1, \cdots, \mathrm{P}$.

Then, a new Vandermonde matrix $\hat{\mathbf{B}}^{(n+1)}$ is formed by the steering vectors of those refined estimates of the DOAs. By using this updated matrix, it is possible to compute the refined estimates of the projection matrices of the signal $\hat{\mathbf{Q}}_B^{(n+1)}$ and the noise $\hat{\mathbf{Q}}_B^{(n+1)\perp}$ subspaces.

Next, employing the refined estimates of the projection matrices, the initial sample data matrix, $\hat{\boldsymbol{R}}$, and the number of sensors and sources, the stochastic maximum likelihood objective function $U^{(n+1)}(\mu)$ [45] is computed for each value of μ at the n^{th} iteration, as follows:

$$U^{(n+1)}(\mu) = \ln \det\Big(\hat{\mathbf{Q}}_B^{(n+1)} \, \hat{\mathbf{R}} \, \hat{\mathbf{Q}}_B^{(n+1)} \tag{3-17}$$

$$+ \; \frac{\mathrm{Trace}\{ \hat{\mathbf{Q}}_B^{\perp \, (n+1)} \, \hat{\mathbf{R}} \}}{\mathrm{M} - \mathrm{P}} \; \hat{\mathbf{Q}}_B^{(n+1)\perp} \Big). \tag{3-18}$$

The previous computation selects the set of unavailable DOA estimates that have a higher likelihood at each iteration. Then, the set of estimated DOAs corresponding to the optimum value of μ that minimizes (3-18) also at each n^{th}

iteration is determined. Finally, the output of the proposed MS-KAI-ESPRIT algorithm is formed by the set of the estimates obtained at the P^{th} iteration, as described in Table 3.1.

The proposed approach will be investigated further and extended for an arbitrary number of iterations in the next chapters.

3.3 Analysis

In this section, we carry out an analysis of the MSE of the data covariance matrix free of side effects along with a study of the computational complexity of the proposed MS-KAI-ESPRIT and existing direction finding algorithms.

3.3.1 MSE Analysis

In this subsection we show that at the first of the P iterations, the MSE of the data covariance matrix free of side effects $\hat{\mathbf{R}}^{(n+1)}$ is less than or equal to the MSE of the original one $\hat{\mathbf{R}}$. This can be formulated as:

$$\text{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\Big|_{n=1} \leq \text{MSE}\left(\hat{\mathbf{R}}\right) \tag{3-19}$$

or, alternatively, as

$$\text{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\Big|_{n=1} - \text{MSE}\left(\hat{\mathbf{R}}\right) \leq 0 \tag{3-20}$$

In what follows, we provide the proof of this inequality.

We start by expressing the MSE of the original data covariance matrix (2-59) as:

$$\text{MSE}\left(\hat{\mathbf{R}}\right) = \mathbb{E}\left[\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2\right]. \tag{3-21}$$

where $\mathbf{R}$ is the true covariance matrix. Similarly, the MSE of the data covariance matrix free of side effects $\hat{\mathbf{R}}^{(n+1)}$ can be expressed for the first

Table 3.1: Proposed MS-KAI-ESPRIT Algorithm

Inputs:

$M,\ d,\ \lambda,\ N,\ P$

Received vectors $\boldsymbol{x}(1),\ \boldsymbol{x}(2),\cdots,\ \boldsymbol{x}(N)$

Outputs:

Estimates $\hat{\theta}_1^{(n+1)}(\mu\ opt),\ \hat{\theta}_2^{(n+1)}(\mu\ opt),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu\ opt)$

First step:

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}(i)\boldsymbol{x}^H(i)$$

$$\{\hat{\theta}_1^{(1)},\ \hat{\theta}_2^{(1)},\cdots,\hat{\theta}_P^{(1)}\}\ \underleftarrow{ESPRIT}\ (\hat{\mathbf{R}}, P, d, \lambda)$$

$$\hat{\mathbf{A}}^{(1)} = \left[\mathbf{a}(\hat{\theta}_1^{(1)}), \mathbf{a}(\hat{\theta}_2^{(1)}), \cdots, \mathbf{a}(\hat{\theta}_P^{(1)})\right]$$

Second step:

for $n = 1 : P$

$$\hat{\mathbf{Q}}_A^{(n)} = \hat{\mathbf{A}}^{(n)}\left(\hat{\mathbf{A}}^{(n)H}\ \hat{\mathbf{A}}^{(n)}\right)^{-1}\hat{\mathbf{A}}^{(n)H}$$

$$\hat{\mathbf{Q}}_A^{(n)\perp} = \mathbf{I}_M\ -\ \hat{\mathbf{Q}}_A^{(n)}$$

$$\mathbf{V}^{(n)} = \hat{\mathbf{Q}}_A^{(n)}\ \hat{\mathbf{R}}\ \hat{\mathbf{Q}}_A^{(n)\perp}$$

for $\mu = 0{:}\ \iota\ : 1$

$$\hat{\mathbf{R}}^{(n+1)} = \hat{\mathbf{R}}\ -\ \mu\left(\mathbf{V}^{(n)}\ +\ \mathbf{V}^{(n)H}\right)$$

$$\{\hat{\theta}_1^{(n+1)},\ \hat{\theta}_2^{(n+1)},\cdots,\hat{\theta}_P^{(n+1)}\}\ \underleftarrow{ESPRIT}\ (\hat{\mathbf{R}}^{(n+1)},\ P, d, \lambda)$$

$$\hat{\mathbf{B}}^{(n+1)} = \left[\mathbf{a}(\hat{\theta}_1^{(n+1)}), \mathbf{a}(\hat{\theta}_2^{(n+1)}), \cdots, \mathbf{a}(\hat{\theta}_P^{(n+1)})\right]$$

$$\hat{\mathbf{Q}}_B^{(n+1)} = \hat{\mathbf{B}}^{(n+1)}\left(\hat{\mathbf{B}}^{(n+1)H}\ \hat{\mathbf{B}}^{(n+1)}\right)^{-1}\hat{\mathbf{B}}^{(n+1)H}$$

$$\hat{\mathbf{Q}}_B^{(n+1)\perp} = \mathbf{I}_M\ -\ \hat{\mathbf{Q}}_B^{(n+1)}$$

$$U^{(n+1)}(\mu) = \ln\det\left(\hat{\mathbf{Q}}_B^{(n+1)}\ \hat{\mathbf{R}}\ \hat{\mathbf{Q}}_B^{(n+1)} + \frac{\mathrm{Trace}\{\hat{\mathbf{Q}}_B^{\perp\,(n+1)}\ \hat{\mathbf{R}}\}}{\mathrm{M}-\mathrm{P}}\ \hat{\mathbf{Q}}_B^{(n+1)\perp}\right)$$

$$\mu_{\mathrm{opt}}^{(n+1)} = \arg\min\ U^{(n+1)}(\mu)$$

$$\mathrm{DOAs}^{(n+1)} = \{\hat{\theta}_1^{(n+1)}(\mu\ opt),\ \hat{\theta}_2^{(n+1)}(\mu\ opt),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu\ opt)\}$$

$$\hat{\mathbf{A}}^{(n+1)} = \left\{\mathbf{a}(\hat{\theta}_{\{1,\cdots,n\}}^{(n+1)})\right\}\bigcup\left\{\mathbf{a}(\hat{\theta}_{\{1,\cdots,P\}-\{1,\cdots,n\}}^{(1)})\right\}$$

end for

end for

iteration $n = 1$ by making use of (5-21), as follows

$$\begin{aligned}
\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\Big|_{n=1} &= \mathrm{MSE}\left(\hat{\mathbf{R}}^{(2)}\right) \\
&= \mathbb{E}\left[\|\hat{\mathbf{R}}^{(2)} - \mathbf{R}\|_F^2\right] \\
&= \mathbb{E}\left[\|\hat{\mathbf{R}} \; - \; \mu\left(\mathbf{V}^{(1)} \; + \; \mathbf{V}^{(1)H}\right) - \mathbf{R}\|_F^2\right] \\
&= \mathbb{E}\left[\|\left(\hat{\mathbf{R}} - \mathbf{R}\right) \; - \; \mu\left(\mathbf{V}^{(1)} \; + \; \mathbf{V}^{(1)H}\right)\|_F^2\right] \qquad (3\text{-}22)
\end{aligned}$$

where for the sake of simplicity, from now on we omit the superscript $^{(1)}$, which refers to the first iteration. In order to expand the result in (3-22), we make use of the following proposition:

Lemma 1: The squared Frobenius norm of the difference between any two matrices $\mathbf{A} \in \mathbb{C}^{m\times m}$ and $\mathbf{B} \in \mathbb{C}^{m\times m}$ is given by

$$\|\mathbf{A} - \mathbf{B}\|_F^2 = \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 - \left(\mathrm{Tr}\,\mathbf{A}^H\mathbf{B} + \mathrm{Tr}\,\mathbf{A}\mathbf{B}^H\right) \qquad (3\text{-}23)$$

Proof of Lemma 1:
The Frobenius norm of any $\mathbf{D} \in \mathbb{C}^{m\times m}$ matrix is defined [1] as

$$\|\mathbf{D}\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{m}|d_{ij}|^2\right)^{\frac{1}{2}} = \left[\mathrm{Tr}\left(\mathbf{D}^H\mathbf{D}\right)\right]^{\frac{1}{2}} \qquad (3\text{-}24)$$

We express $\mathbf{D}$ as a difference between two matrices $\mathbf{A}$ and $\mathbf{B}$, both also $\in \mathbb{C}^{m\times m}$. Making use of Lemma1 and the properties of the trace, we obtain

$$\begin{aligned}
\|\mathbf{A} - \mathbf{B}\|_F^2 &= \mathrm{Tr}\left[\left(\mathbf{A} - \mathbf{B}\right)^H\left(\mathbf{A} - \mathbf{B}\right)\right] \\
&= \mathrm{Tr}\left[\left(\mathbf{A}^H - \mathbf{B}^H\right)\left(\mathbf{A} - \mathbf{B}\right)\right] \\
&= \mathrm{Tr}\left[\left(\mathbf{A}^H\mathbf{A}\right) - \mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\right) - \mathrm{Tr}\left(\mathbf{B}^H\mathbf{A}\right) + \mathrm{Tr}\left(\mathbf{B}^H\mathbf{B}\right)\right] \\
&= \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 - \left(\mathrm{Tr}\,\mathbf{A}^H\mathbf{B} + \mathrm{Tr}\,\mathbf{A}\mathbf{B}^H\right), \qquad (3\text{-}25)
\end{aligned}$$

which is the desired result.

Now, assuming that the true $\mathbf{R}$ [5] and the data covariance matrices $\hat{\mathbf{R}}$ [5] are Hermitian and using (3-22) combined with Lemma1, the cyclic [58] property of the trace and the linearity [59] property of the expected value, we get

$$\begin{aligned}
\mathrm{MSE}\left(\hat{\mathbf{R}}^{(2)}\right) = \mathbb{E}\Big\{&\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2 + \mu^2\,\|\mathbf{V} \; + \; \mathbf{V}^H\|_F^2 \\
&- \mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)^H\mu\left(\mathbf{V} \; + \; \mathbf{V}^H\right)\right] \\
&- \mathrm{Tr}\left[\mu\left(\mathbf{V} + \mathbf{V}^H\right)^H\left(\hat{\mathbf{R}} \; + \; \mathbf{R}\right)\right]\Big\}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left\{\|\hat{\mathbf{R}}-\mathbf{R}\|_F^2+\mu^2\,\|\mathbf{V}\,+\,\mathbf{V}^H\|_F^2\right.\\
&-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)^H\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\\
&\left.-\mu\,\mathrm{Tr}\left[\left(\mathbf{V}+\mathbf{V}^H\right)^H\left(\hat{\mathbf{R}}\,+\,\mathbf{R}\right)\right]\right\}\\
&= \mathbb{E}\left\{\|\hat{\mathbf{R}}-\mathbf{R}\|_F^2+\mu^2\,\|\mathbf{V}\,+\,\mathbf{V}^H\|_F^2\right.\\
&-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\\
&\left.-\mu\,\mathrm{Tr}\left[\left(\mathbf{V}^H+\mathbf{V}\right)\left(\hat{\mathbf{R}}\,+\,\mathbf{R}\right)\right]\right\}\\
&= \mathbb{E}\left\{\|\hat{\mathbf{R}}-\mathbf{R}\|_F^2+\mu^2\,\|\mathbf{V}\,+\,\mathbf{V}^H\|_F^2\right.\\
&-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\\
&\left.-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}}\,+\,\mathbf{R}\right)\left(\mathbf{V}+\mathbf{V}^H\right)\right]\right\}\\
&= \mathbb{E}\left\{\|\hat{\mathbf{R}}-\mathbf{R}\|_F^2\right\}+\mu^2\mathbb{E}\left\{\|\mathbf{V}\,+\,\mathbf{V}^H\|_F^2\right\}\\
&-2\mu\mathbb{E}\left\{\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\right\}\\
&= \mathrm{MSE}\left(\hat{\mathbf{R}}\right)+\mu^2\mathbb{E}\left\{\|\mathbf{V}\,+\,\mathbf{V}^H\|_F^2\right\}\\
&-2\mu\mathbb{E}\left\{\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\right\} \qquad (3\text{-}26)
\end{aligned}
$$

By moving the first summand of (3-26) to its first element, we obtain the intended expression for the difference between the $MSEs$ of the data covariance matrix free of perturbations and the original one, i.e.:

$$
\begin{aligned}
\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\Big|_{n=1}-\mathrm{MSE}\left(\hat{\mathbf{R}}\right) &= \mu^2\mathbb{E}\left\{\|\mathbf{V}+\mathbf{V}^H\|_F^2\right\}\\
&-2\mu\mathbb{E}\left\{\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\right\}. \qquad (3\text{-}27)
\end{aligned}
$$

Now, we expand the expressions inside braces of the second member of (3-27) individually. We start with the first summand

$$
\begin{aligned}
\|\mathbf{V}+\mathbf{V}^H\|_F^2 &= \|\mathbf{V}\|_F^2+\|\mathbf{V}^H\|_F^2+\mathrm{Tr}\left(\mathbf{V}^H\mathbf{V}^H\right)\\
&+\mathrm{Tr}\left(\left(\mathbf{V}^H\right)^H\mathbf{V}\right)\\
&= \|\mathbf{V}\|_F^2+\|\mathbf{V}^H\|_F^2+\mathrm{Tr}\left(\mathbf{V}^H\mathbf{V}^H\right)+\mathrm{Tr}\left(\mathbf{V}\mathbf{V}\right). \qquad (3\text{-}28)
\end{aligned}
$$

The equation (3-28) can be computed by using the projection matrices of the signal and the noise subspaces and the data covariance matrix by using (5-18), (3-7), the idempotence [1] [58] of $\hat{\mathbf{Q}}_A$ and the cyclic property [58] of the trace.

Starting with the computation of its fourth summand, we have

$$\begin{aligned}
\mathrm{Tr}\left(\mathbf{VV}\right) &= \mathrm{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)\right] \\
&= \mathrm{Tr}\left[\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\right] \\
&= \mathrm{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)\right. \\
&\left.\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)\right] \\
&= \mathrm{Tr}\left[\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right. \\
&\left.-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}+\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right] \\
&= \mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\
&- \mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right) \\
&+ \mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\
&= \mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right) \\
&- \mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)+\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)=0. \qquad (3\text{-}29)
\end{aligned}$$

Taking into account that the data covariance matrix $\hat{\mathbf{R}}$ and the estimate of the projection matrix of the noise subspace $\hat{\mathbf{Q}}_A^{\perp}$ are Hermitian, we can evaluate the third summand of (3-28) as follows:

$$\begin{aligned}
&\mathrm{Tr}\left(\mathbf{V}^H\mathbf{V}^H\right) = \mathrm{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)^H\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)^H\right] \\
&= \mathrm{Tr}\left\{\left[\left(\hat{\mathbf{Q}}_A^{\perp}\right)^H\,\hat{\mathbf{R}}^H\,\hat{\mathbf{Q}}_A^H\right]\left[\left(\hat{\mathbf{Q}}_A^{\perp}\right)^H\,\hat{\mathbf{R}}^H\,\hat{\mathbf{Q}}_A^H\right]\right\} \\
&= \mathrm{Tr}\left\{\left[\hat{\mathbf{Q}}_A^{\perp}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\left[\hat{\mathbf{Q}}_A^{\perp}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\right\} \\
&= \mathrm{Tr}\left\{\left[\left(\mathbf{I_M}-\hat{\mathbf{Q}}_{\mathbf{A}}\right)\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\left[\left(\mathbf{I_M}-\hat{\mathbf{Q}}_{\mathbf{A}}\right)\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\right\} \\
&= \mathrm{Tr}\left\{\left[\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\left[\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\right\} \\
&= \mathrm{Tr}\left\{\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A-\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right. \\
&\left.-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A+\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right\} \\
&= \mathrm{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)-\mathrm{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\
&- \mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)+\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\
&= \mathrm{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)-\mathrm{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right) \\
&- \mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)+\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)=0. \qquad (3\text{-}30)
\end{aligned}$$

By using (3-24), we can expand the first and the second summands of (3-28)

as follows:

$$\begin{aligned}\|\mathbf{V}\|_F^2 + \|\mathbf{V}^H\|_F^2 &= \mathrm{Tr}\left(\mathbf{V}^H\mathbf{V}\right) + \mathrm{Tr}\left(\left(\mathbf{V}^H\right)^H\mathbf{V}^H\right)\\ &= \mathrm{Tr}\left(\mathbf{V}^H\mathbf{V}\right) + \mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right)\\ &= \mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right) + \mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right)\\ &= 2\,\mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right). \end{aligned} \tag{3-31}$$

Equation (3-31) can be expressed in terms of the projection matrices of the signal and the noise subspaces and the data covariance, in a similar way as for the third and fourth summands of (3-28), as follows:

$$\begin{aligned}2\,\mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right) &= 2\,\mathrm{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)^H\right]\\ &= 2\,\mathrm{Tr}\left\{\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\left[\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\right]^H\right\}\\ &= 2\,\mathrm{Tr}\left\{\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)^H\right\}\\ &= 2\,\mathrm{Tr}\left\{\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_A-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right.\\ &\quad\left.-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A+\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\ &= 2\left\{\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right.\\ &\quad-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)+\left.\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\}\\ &= 2\left\{\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\right)-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right.\\ &\quad-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)+\left.\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\}\\ &= 2\left\{\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\right)-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\} \end{aligned} \tag{3-32}$$

From (3-28), (3-29), (3-30), (3-31) and (3-32), we obtain the first summand of (3-27), as follows:

$$\mu^2\mathbb{E}\left\{\|\mathbf{V}+\mathbf{V}^H\|_F^2\right\} = 2\mu^2\mathbb{E}\left\{\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\right)-\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\} \tag{3-33}$$

In order to finish the expansion of the expressions inside braces of the second member of (3-27), now we deal with its second summand, in which we make use of the cyclic property [58] of the trace and the idempotence [1] [58] of $\hat{\mathbf{Q}}_A$.

$$
\begin{aligned}
&\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}+\mathbf{V}^{H}\right)\right]\\
&=\left\{\mathrm{Tr}\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left[\hat{\mathbf{Q}}_{A}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_{A}^{\perp}+\left(\hat{\mathbf{Q}}_{A}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_{A}^{\perp}\right)^{H}\right]\right\}\\
&=\mathrm{Tr}\left\{\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left[\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_{A}\right)\right.\right.\\
&\left.\left.+\left(\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_{A}\right)\right)^{H}\right]\right\}\\
&=\mathrm{Tr}\left\{\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left[\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\right.\right.\\
&\left.\left.+\left(\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\right)^{H}\right]\right\}\\
&=\mathrm{Tr}\left\{\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left[\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}+\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}-\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\right]\right\}\\
&=\mathrm{Tr}\left\{\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}+\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}-2\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\right.\\
&\left.-\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\mathbf{R}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}+2\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\right\}\\
&=\mathrm{Tr}\,\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}+\mathrm{Tr}\,\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}-2\,\mathrm{Tr}\,\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\\
&-\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}+2\,\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\\
&=\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{R}}+\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{R}}-2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\\
&-\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{R}}+2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\\
&=2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{R}}-2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\\
&-\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{R}}+2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\\
&=2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{R}}-2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\\
&-\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{R}}+2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}
\end{aligned}
\tag{3-34}
$$

By using (3-34), we can straightforwardly write the second summand of the second member of (3-27) in terms of the projection matrices of the signal and the noise subspaces and the data covariance matrix as follows:

$$
\begin{aligned}
&-2\mu\mathbb{E}\left\{\mathrm{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\;+\;\mathbf{V}^{H}\right)\right]\right\}\\
&=-2\mu\mathbb{E}\left\{2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{R}}-2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}-\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\right.\\
&\left.-\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{R}}+2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\right\}\\
&=-4\mu\mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{R}}-\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\right\}\\
&-2\mu\left\{-\mathrm{Tr}\,\mathbb{E}\left[\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\right]-\mathrm{Tr}\,\mathbb{E}\left[\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{R}}\right]\right.\\
&\left.+2\,\mathrm{Tr}\,\mathbb{E}\left[\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\right]\right\}\\
&=-4\mu\mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{R}}-\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{A}\hat{\mathbf{R}}\right\}\\
&-2\mu\left\{-\mathrm{Tr}\,R\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\mathbb{E}\left[\hat{\mathbf{R}}\right]-\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\hat{\mathbf{Q}}_{A}\mathbf{R}\mathbb{E}\left[\hat{\mathbf{R}}\right]\right.\\
&\left.+2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_{A}\mathbf{R}\hat{\mathbf{Q}}_{A}\mathbb{E}\left[\hat{\mathbf{R}}\right]\right\}
\end{aligned}
\tag{3-35}
$$

Now, by using (3-33) and (3-35), and assuming that $\mathbb{E}\left[\hat{\mathbf{R}}\right]$ is an unbiased estimate of $\hat{\mathbf{R}}$, i.e., $\mathbb{E}\left[\hat{\mathbf{R}}\right] = \mathbf{R}$, we can rewrite (3-27) as follows:

$$
\begin{aligned}
&\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\Big|_{n=1} - \mathrm{MSE}\left(\hat{\mathbf{R}}\right) \\
&= \mu^2 \mathbb{E}\left\{\|\mathbf{V} + \mathbf{V}^H\|_F^2\right\} \\
&- 2\mu \mathbb{E}\left\{\mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)\left(\mathbf{V} + \mathbf{V}^H\right)\right]\right\} \\
&= 2\mu^2 \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\
&- 4\mu \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\
&- 2\mu\left\{-\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R}\right. \\
&\left.+2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\} \\
&= 2\mu^2 \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\
&- 4\mu \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\
&- 2\mu\left\{-2\,\mathrm{Tr}\,\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R} + 2\,\mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\} \\
&= 2\mu^2 \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\
&- 4\mu \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\
&- 4\mu\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\} \\
&= \left(2\mu^2 - 4\mu\right)\mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\
&- 4\mu\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\}
\end{aligned}
\tag{3-36}
$$

Next, we will discuss equation (3-36). For this purpose, we assume that the estimate of the projection matrix of the signal subspace $\hat{\mathbf{Q}}_A$ [1], the true $\mathbf{R}$ [5] and the data covariance matrices $\hat{\mathbf{R}}$ [5] are Hermitian. For the next steps we will make use of the following Theorem which is proved in [60]:

Theorem 1: For two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order,

$$
\mathrm{Tr}\left(\mathbf{AB}\right)^{2^k} \leq \mathrm{Tr}\left(\mathbf{A}^{2^k}\mathbf{B}^{2^k}\right), \tag{3-37}
$$

where k is in integer.

By replacing $\mathbf{A}$ with $\hat{\mathbf{Q}}_A$ and $\mathbf{B}$ with $\hat{\mathbf{R}}$ in (3-37) and also considering $k = 1$,

we have

$$\begin{aligned} &\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)^2 \leq \mathrm{Tr}\left(\hat{\mathbf{Q}}_A^2\hat{\mathbf{R}}^2\right) \\ &\therefore \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}} \leq \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} \\ &\Rightarrow \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}} \geq 0 \end{aligned} \tag{3-38}$$

Similarly, making $\mathbf{A} = \hat{\mathbf{Q}}_A$ and $\mathbf{B} = \mathbf{R}$ for $k = 1$, we obtain

$$\begin{aligned} &\mathrm{Tr}\left(\hat{\mathbf{Q}}_A\mathbf{R}\right)^2 \leq \mathrm{Tr}\left(\hat{\mathbf{Q}}_A^2\mathbf{R}^2\right) \\ &\therefore \mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R} \leq \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R} \\ &\Rightarrow \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R} \geq 0 \end{aligned} \tag{3-39}$$

Next, we analyze the behavior of the expressions -4μ and $(2\mu^2 - 4\mu)$ based on the reliability factor $\mu \in [0\ 1]$, as defined in (5-21). In order to illustrate the case being studied, we assume that both expressions are continuous functions as depicted in Fig. 3.1. It can be seen in it that in the range $[0\ 1]$ both expressions

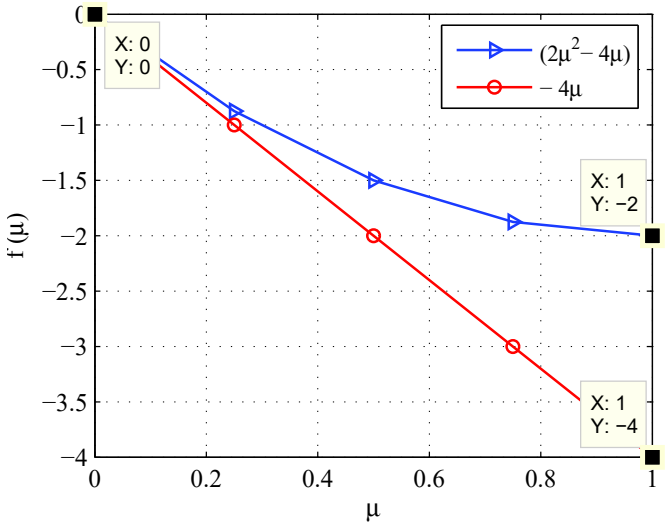

Figure 3.1: Behavior of $(2\mu^2 - 4\mu)$ and -4μ for $\mu \in [0\ 1]$.

assume values $\mathrm{f}(\mu) \leq 0$, i.e.:

$$\text{For } \mu \in [0\ 1]: \begin{cases} \left(2\mu^2 - 4\mu\right) \leq 0 \\ -4\mu \leq 0 \end{cases} \tag{3-40}$$

Now, we can consider the traces which form the subtraction in (3-38) as

different random variables $y(\omega)$ and $x(\omega)$, i.e.:

$$\left.\begin{array}{l} \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} = y(\omega) \\ \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}} = x(\omega) \end{array}\right\},\ \forall\,\omega\ \in \Omega. \tag{3-41}$$

In addition, we can suppose that there is a random variable $z(\omega)$ always greater than zero, i.e., $z(\omega) \geq 0$, so that

$$z(\omega) = y(\omega) - x(\omega) \geq 0,\ \forall\,\omega\ \in \Omega \tag{3-42}$$

Taking the expectation of (3-42) and applying its properties of linearity and monotonicity [59, 61], we obtain

$$\mathbb{E}\left[z(\omega)\right] = \mathbb{E}\left[y(\omega) - x(\omega)\right] \geq 0, \tag{3-43}$$

which, by making use of (3-41), results in

$$\begin{aligned} \mathbb{E}\left[z(\omega)\right] &= \mathbb{E}\left[y(\omega) - x(\omega)\right] \\ &= \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \geq 0 \end{aligned} \tag{3-44}$$

Next, we can combine the inequalities (3-40) with (3-44) to compute the second member of (3-36), for $\mu \in [0\ 1]$.

For its first summand, we combine (3-40) and (3-44), as follows:

$$\begin{cases} \mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \geq 0 \\ \left(2\mu^2 - 4\mu\right) \leq 0,\ \mu \in [0\ 1], \end{cases} \tag{3-45}$$

to obtain in a straightforward way

$$\left(2\mu^2 - 4\mu\right)\mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \leq 0 \tag{3-46}$$

Similarly, we can compute its second member, by combining (3-40) and (3-39), as described by

$$\begin{cases} \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}} \geq 0 \\ -4\mu \leq 0,\ \mu \in [0\ 1], \end{cases} \tag{3-47}$$

to obtain also straightforwardly the expression given by

$$-4\mu\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\} \leq 0 \tag{3-48}$$

By combining the inequalities (3-46) and (3-48) with (3-36), we have

$$\begin{aligned}
&\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\Big|_{n=1} - \mathrm{MSE}\left(\hat{\mathbf{R}}\right) \\
&= \underbrace{\left(2\mu^2 - 4\mu\right)\mathbb{E}\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}}_{\leq 0} \\
&\underbrace{-4\mu\left\{\mathrm{Tr}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R} - \mathrm{Tr}\,\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\}}_{\leq 0} \\
&\therefore\ \mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\Big|_{n=1} - \mathrm{MSE}\left(\hat{\mathbf{R}}\right) \leq\ 0
\end{aligned} \tag{3-49}$$

which is the desired result.

3.3.2

Computational Complexity Analysis

In this section, we evaluate the computational cost of the proposed MS-KAI-ESPRIT algorithm which is compared to the following classical subspace methods: ESPRIT [10], MUSIC [8], Root-MUSIC [9], Conjugate Gradient (CG) [15,16], Auxiliary Vector Filtering (AVF) [13] and TS-ESPRIT [47]. The ESPRIT and MUSIC-based methods use the Singular Value Decomposition (SVD) of the sample covariance matrix (2-59). The computational complexity of MS-KAI-ESPRIT in terms of number of multiplications and additions is depicted in Table 3.2, where $\tau = \frac{1}{\iota} + 1$. The increment ι is defined in Table 3.1.

Table 3.2: Computational complexity - MS-KAI-ESPRIT [52]

Multiplications
$P\,\tau[\frac{10}{3}\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P}+2) + \mathrm{M}(\frac{5}{2}\mathrm{P}^2 + \frac{1}{2}\mathrm{P} + 8\mathrm{N}^2)\ + \mathrm{P}^2(\frac{17}{2}\mathrm{P} + \frac{1}{2})]$
$+\mathrm{P}\,[2\mathrm{M}^3 + \mathrm{M}^2(\mathrm{P}) + \mathrm{M}(\frac{3}{2}\mathrm{P}^2 + \frac{1}{2}\mathrm{P}) + \mathrm{P}^2(\frac{\mathrm{P}}{2} + \frac{3}{2})]$
$+2\mathrm{M}^2(\mathrm{P}) + \mathrm{M}(\mathrm{P}^2 - \mathrm{P} + 8\mathrm{N}^2) + \mathrm{P}^2(8\mathrm{P} - 1)$
Additions
$P\,\tau[\frac{10}{3}\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P}-1) + \mathrm{M}(\frac{5}{2}\mathrm{P}^2 - \frac{9}{2}\mathrm{P} + 8\mathrm{N}^2)\ + \mathrm{P}(8\mathrm{P}^2 - 2\mathrm{P} - \frac{5}{2})]$
$+\mathrm{P}\,[2\mathrm{M}^3 + \mathrm{M}^2(\mathrm{P}-2) + \mathrm{M}(\frac{3}{2}\mathrm{P}^2 - \frac{1}{2}\mathrm{P}) - \mathrm{P}(\mathrm{P} + \frac{1}{2})]$
$+2\mathrm{M}^2(\mathrm{P}) + \mathrm{M}(\mathrm{P}^2 - 4\mathrm{P} + 8\mathrm{N}^2) + \mathrm{P}(8\mathrm{P}^2 - \mathrm{P} - 2)$

As can be seen, for this specific configuration used in the simulations 3.4 MS-KAI-ESPRIT shows a relatively high computational burden with

$\mathcal{O}(P\tau(3M^3 + 8MN^2))$, where τ is typically an integer that ranges from 2 to 20. It can be noticed that for the configuration used in the simulations $(P = 4, M = 40, N = 25)$ $3M^3$ and $8MN^2$ are comparable, resulting in two dominant terms. It can also be seen that the number of multiplications required by the proposed algorithm is more significant than the number of additions. For this reason, in Table 3.3, we computed only the computational burden of the previously mentioned algorithms in terms of multiplications for the purpose of comparisons. In that table, Δ stands for the search step.

Table 3.3: Computational complexity - other algorithms

Algorithm	Multiplications
MUSIC [8]	$\frac{180}{\Delta}[\mathrm{M}^2 + \mathrm{M}(2 - \mathrm{P}) - \mathrm{P}] + 8\mathrm{MN}^2$
root-MUSIC [9]	$2\mathrm{M}^3 - \mathrm{M}^2\mathrm{P} + 8\mathrm{MN}^2$
AVF [13]	$\frac{180}{\Delta}[\mathrm{M}^2(3\mathrm{P} + 1) + \mathrm{M}(4\mathrm{P} - 2) + \mathrm{P} + 2]$ $+\mathrm{M}^2\mathrm{N}$
CG [15]	$\frac{180}{\Delta}[\mathrm{M}^2(\mathrm{P} + 1) + \mathrm{M}(6\mathrm{P} + 2) + \mathrm{P} + 1] + \mathrm{M}^2\mathrm{N}$
ESPRIT [10]	$2\mathrm{M}^2\mathrm{P} + \mathrm{M}(\mathrm{P}^2 - 2\mathrm{P} + 8\mathrm{N}^2) + 8\mathrm{P}^3 - \mathrm{P}^2$
TS-ESPRIT [47]*	$\tau[3\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P} + 2) + \mathrm{M}(\frac{5}{2}\mathrm{P}^2 - \frac{3}{2}\mathrm{P} + 8\mathrm{N}^2)$ $+\mathrm{P}^2(\frac{17}{2}\mathrm{P} + \frac{1}{2}) + 1]$ $+[2\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P}) + \mathrm{M}(\frac{5}{2}\mathrm{P}^2 - \frac{3}{2}\mathrm{P} + 8\mathrm{N}^2)$ $+\mathrm{P}^2(\frac{17}{2}\mathrm{P} + \frac{1}{2})]$

Next, we will evaluate the influence of the number of sensor elements on the number of multiplications based on Tables 3.2 and 3.3, respectively. Supposing $\mathrm{P} = 4$ narrowband signals impinging on a ULA of M sensor elements and $\mathrm{N} = 25$ available snapshots, we obtain Fig. 3.2. We can see the main trends in terms of computational cost measured in multiplications of the proposed and analyzed algorithms. By examining Fig. 3.2, it can be noticed that in the range $M = [20\ 70]$ sensors, the curves describing the exact number of multiplications in MS-KAI-ESPRIT and AVF tend to merge. For $M = 40$, this ratio tends to 1, i.e. their numbers of multiplications are almost equivalent.

3.4

Simulations

In this section, we examine the performance of the proposed MS-KAI-ESPRIT in terms of probability of resolution and RMSE and compare them to the standard ESPRIT [10], the Iterative ESPRIT (IESPRIT), which is also developed here by combining the approach in [44] that exploits knowledge of the structure of the covariance matrix and its perturbation terms, the Conjugate Gradient

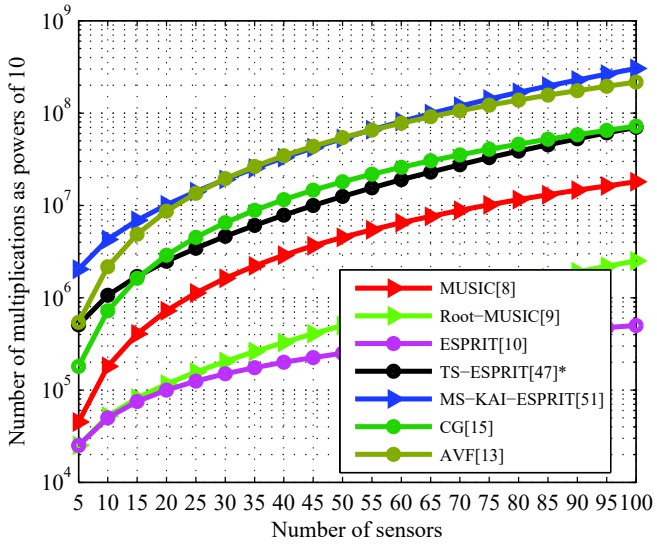

Figure 3.2: Number of multiplications as powers of 10 versus number of sensors for $P = 4$, $N = 25$.

(CG) [15], the Root-MUSIC [9], and the MUSIC [8] algorithms. Despite TS-ESPRIT is based on the knowledge of available known DOAS and the proposed MS-KAI-ESPRIT does not have access to prior knowledge, TS-ESPRIT is plotted with the aim of illustrating the comparisons. For a fair comparison in terms of RMSE and probability of resolution of all studied algorithms, we suppose that we do not have prior knowledge, that is to say that although we have available known DOAs, we compute TS-ESPRIT as they were unavailable. We employ a ULA with M=40 sensors, inter-element spacing $\Delta = \frac{\lambda_c}{2}$ and assume there are four uncorrelated complex Gaussian signals with equal power impinging on the array. The closely-spaced sources are separated by 2.4^o, at $(10.2^o, 12.6^o, 15^o, 17.4^o)$, and the number of available snapshots is N=25. For TS-ESPRIT, as previously mentioned, we presume a priori knowledge of the last true DOAS $(15^o, 17.4^o)$.

In Fig. 3.3, we show the probability of resolution versus SNR. We take into account the criterion [29], in which two sources with DOA θ_1 and θ_2 are said to be resolved if their respective estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are such that both $\left|\hat{\theta}_1 - \theta_1\right|$ and $\left|\hat{\theta}_2 - \theta_2\right|$ are less than $|\theta_1 - \theta_2|/2$. The proposed MS-KAI-ESPRIT algorithm outperforms IESPRIT developed here, based on [44, 62], and the standard ESPRIT [10] in the range between -6 and $5dB$ and MUSIC [8] from -6 to $8.5dB$. MS-KAI-ESPRIT also outperforms CG [15, 16] and Root-Music [9] throughout the whole range of values. The poor performance of the latter could be expected from the results for two closed signals obtained in [44, 62]. When compared to TS-ESPRIT, which as previously discussed,

was supposed to have the best performance, the proposed MS-KAI-ESPRIT algorithm is outperformed by the former only in the range between -6 and $-2dB$. From this last point to $20dB$ its performance is superior or equal to the other algorithms.

In Fig. 3.4, it is shown the RMSE in dB versus SNR, where the term CRB refers to the square root of the deterministic Cramér-Rao bound [18]. The RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{L\,P}\sum_{l=1}^{L}\sum_{p=1}^{P}(\theta_p - \hat{\theta}_p(l))^2}, \tag{3-50}$$

where L is the number of trials.

The results show the superior performance of MS-KAI-ESPRIT in the range between -2.5 and 5 dB. From this last point to 20 dB, MS-KAI-ESPRIT, IESPRIT, ESPRIT and TS-ESPRIT have similar performance. The only range in which MS-KAI-ESPRIT is outperformed lies in the range between -6 and -2.5 dB. From this last point to 20 dB its performance is better or similar to the others.

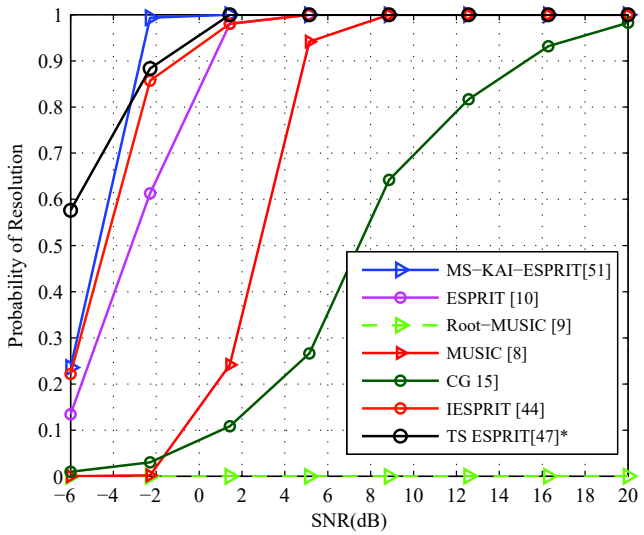

Figure 3.3: Probability of resolution versus SNR with $P = 4$ uncorrelated sources, $M = 40$, $N = 25$, $L = 100$ runs.

Now, we focus on the performance of MS-KAI-ESPRIT under more severe conditions, i.e., we analyze it in terms of RMSE when at least two of the four equal-powered Gaussian signals are strongly correlated, as shown in the following signal correlation matrix $\boldsymbol{R}_{ss}$ (3-51):

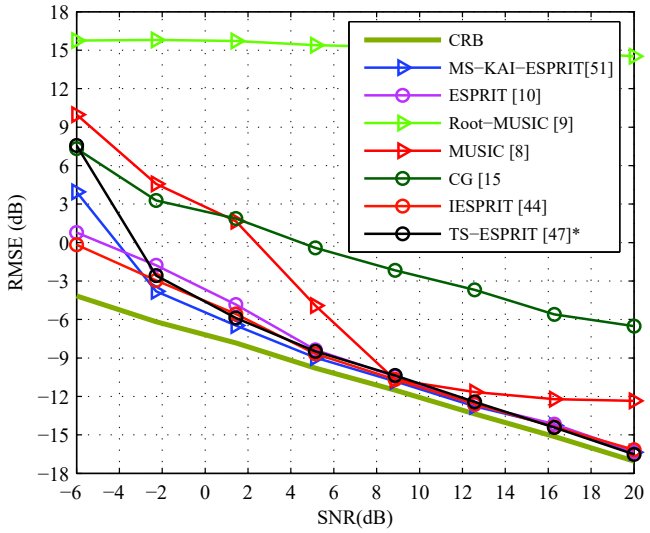

Figure 3.4: RMSE and the square root of CRB versus SNR with $P = 4$ uncorrelated sources, $M = 40$, $N = 25$, $L = 100$ runs.

$$\boldsymbol{R}_{ss} = \sigma_s^2 \begin{bmatrix} 1 & 0.9 & 0.6 & 0 \\ 0.9 & 1 & 0.4 & 0.5 \\ 0.6 & 0.4 & 1 & 0 \\ 0 & 0.5 & 0 & 1 \end{bmatrix}. \tag{3-51}$$

The signal-to-noise ratio (SNR) is defined as $SNR \triangleq 10\log_{10}\left(\frac{\sigma_s^2}{\sigma_n^2}\right)$. In Fig.

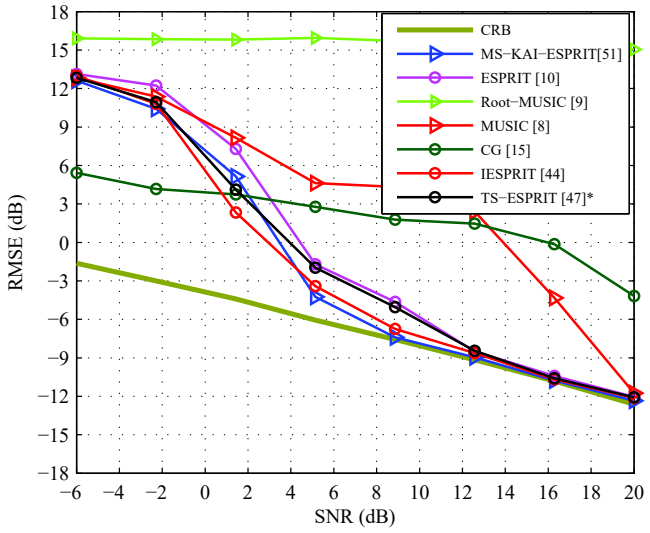

Figure 3.5: RMSE and the square root of CRB versus SNR with $P = 4$ correlated sources, $M = 40$, $N = 25$, $L = 250$ runs.

3.5, we can see the performance of the same algorithms plotted in Fig. 3.4 in terms of RMSE(dB) versus SNR computed after 250 runs, when the signal correlation matrix is given by (3-51). As can be seen, the superior performance

of MS-KAI-ESPRIT occurs in the whole range between 4.0 and 12 dB, which can be considered a small but consistent gain. From 12dB to 20dB MS-KAI-ESPRIT, TS-ESPRIT, IESPRIT and ESPRIT have similar performance. The values for which MS-KAI-ESPRIT is outperformed are in the range between -6.0 and 4.0dB.

In Fig. 3.6, we have provided further simulations to illustrate the performance of each iteration of MS-KAI ESPRIT in terms of RMSE. The resulting iterations can be compared to each other and to the the original ESPRIT, which corresponds to the first step of MS-KAI ESPRIT. For this purpose, we have considered the same scenario employed before, except for the number of the trials, which is $L = 200$ runs for all simulations. In particular, we have considered the case of correlated sources. From Fig. 3.7, which is a magnified detail of Fig. 3.6, it can be seen that the estimates become more accurate with the increase of iterations.

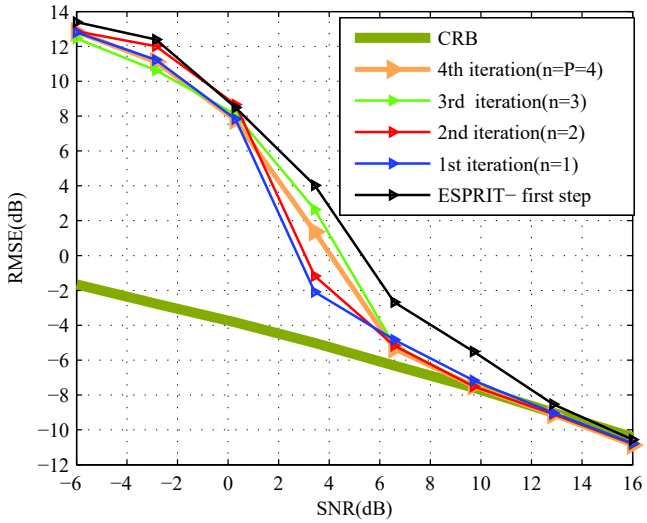

Figure 3.6: RMSE for each iteration of MS-KAI ESPRIT,original ESPRIT and CRB versus SNR with $P = 4$ correlated sources, $M = 40$, $N = 25$, $L = 200$ runs.

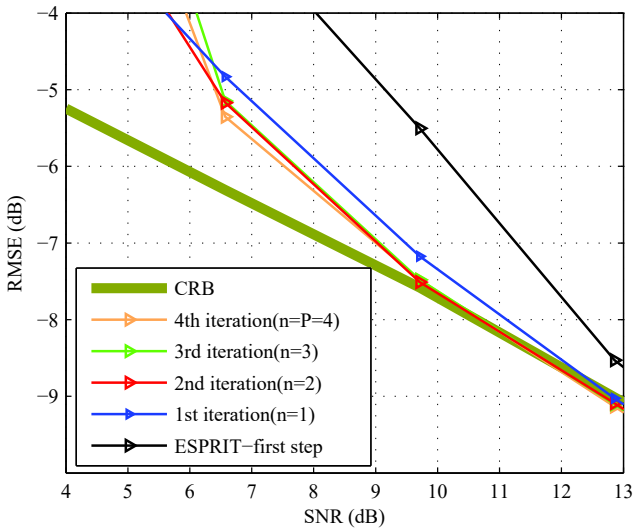

Figure 3.7: RMSE for each iteration of MS-KAI ESPRIT,original ESPRIT and CRB versus SNR with $P = 4$ correlated sources, $M = 40$, $N = 25$, $L = 200$ runs -magnification.

4

Multi-Step Knowledge-aided Iterative Conjugate Gradient algorithm

4.1

Introduction

In this chapter, we present the Multi-Step Knowledge-Aided Iterative Conjugate Gradient (MS-KAI-CG) [66] algorithm, whose preliminary results have been shown in [65]. Both works combine distinct but complementary approaches developed in [30] and [47]. The former approach, termed KA-CG [30], introduced a method that replaces the available known DOAs employed to compute the a priori knowledge matrix, which is employed in the traditional a priori knowledge [49] approach, with previous estimates obtained via the CG algorithm. This a priori knowledge matrix combined with the data covariance matrix results in an enhanced covariance matrix, which, after being processed by the CG algorithm, results in more accurate estimates. This approach is not restricted to CG, the CG algorithm, can be combined with other types of algorithms and can also be viewed as obtaining knowledge on line. The latter and more modern approach [47], which is termed TS-KAI, makes use of ESPRIT and was the first step toward the MS-KAI techniques. Similarly to the MS-KAI technique, TS-KAI (ESPRIT) employs refinements of the covariance matrix combined with incorporation of knowledge to improve the accuracy of DOA estimation. However, these approaches differ in two key aspects: instead of using multiple steps, TS-KAI makes use of only two of them; and instead of acquiring knowledge on line like MS-KAI approach, which can be viewed as its evolution, TS-KAI employs available knowledge of DOAS, like those from base stations or static users. MS-KAI-CG follows the MS-KAI approach applied to ESPRIT, which was described in Chapter 3, and is complemented by its version equipped with forward-backward spatial smoothing, denoted as MS-KAI-CG-FB, which can deal with correlated signals. Unlike prior KAI approaches, MS-KAI-CG and MS-KAI-CG-FB are no longer limited to P iterations, where P is the number of source signals. Moreover, the CG-based

algorithms are particularly effective for scenarios with very few source signals and closely-spaced angles of arrival. For these situations, the MS-KAI-CG approach is particularly appealing and addresses some of the weaknesses of existing CG-based approaches by improving the quality of the covariance matrix estimates. We then carry out an analysis of the computational complexity of the proposed and existing direction finding algorithms along with a simulation study for scenarios with closely-spaced source signals.

4.2

Proposed MS-KAI-CG Algorithm

In this section, we present the proposed MS-KAI-CG algorithm [66] applied to uncorrelated sources and detail its main features. For this purpose we make use of the same system model described in Subsection 2.1.6. We start by expanding the estimate of the data covariance matrix (2-52), which here is denoted by $\hat{\boldsymbol{R}}_o$, using (2-49), as derived in 3.2:

$$\begin{aligned}\hat{\boldsymbol{R}}_o &= \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{A}\,s(i)+\boldsymbol{n}(i))\,(\boldsymbol{A}\,s(i)+\boldsymbol{n}(i))^H \\ &= \boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{n}^H(i) \\ &+ \underbrace{\boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{n}^H(i)\right\} + \left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H}_{\text{"undesirable by-products"}}\end{aligned} \tag{4-1}$$

Similarly to Section 3.2, the first two terms of $\hat{\boldsymbol{R}}_o$ in (4-1) can be considered as estimates of the two summands of $\boldsymbol{R}$ given in (2-51), which represent the signal and the noise components, respectively. The last two terms in (4-1) are undesirable by-products, which can be seen as estimates for the correlation between the signal and the noise vectors. The system model under study is based on noise vectors which are zero-mean and also independent of the signal vectors. Therefore, the signal and noise components are uncorrelated to each other. As a consequence, for a large enough number of samples N, the last two terms pointed out in (4-1) tend to zero. However, in practice the number of available samples can be limited. In such situations, the last two terms in (4-1) may have non negligible values, which causes the deviation of the estimates of the signal and the noise subspaces from the true signal and noise ones.

The key approach of the proposed MS-KAI-CG algorithm is to reshape the sample data covariance matrix estimate at each iteration by gradually incorporating the knowledge provided by the updated Vandermonde matrices which progressively incorporate the newer estimates from the preceding iteration. Based on these updated Vandermonde matrices, refined estimates of the projection matrices of the signal and noise subspaces are calculated. These estimates of projection matrices associated with the initial sample covariance matrix estimate and the reliability (scaling) factor employed to reduce its by-products allow to choose the set of estimates that has the minimum value of the stochastic maximum likelihood objective function (SMLOF), i.e., the highest likelihood of being the set of the true DOAs. The modified covariance matrix is computed by deriving a scaled version of the undesirable terms from $\hat{\boldsymbol{R}}_o$, which are pointed out in (4-1).

The steps of the proposed algorithm are listed in Table 4.2. The algorithm starts by computing the sample data covariance matrix (2-52). Next, the DOAs are estimated using the CG direction finding algorithm reported in [15, 16, 27].

In this chapter, the rank P is assumed to be known, which is an assumption frequently found in the literature. Alternatively, the rank P could be estimated by model-order selection schemes [53] such as Akaike's Information Theoretic Criterion (AIC) [54] and the Minimum Descriptive Length (MDL) Criterion [55]. The CG method, from which the first and the last steps of the MS-KAI-CG are based on, is used to minimize a cost function, or analogously, to solve a linear system of equations by approaching the optimal solution step by step via a line search along successive directions, which are sequentially determined at each direction [19]. As a result of the application of the CG algorithm to direction finding, we have a system of equations that is iteratively solved for $\boldsymbol{w}$ at each search angle:

$$\boldsymbol{R}\boldsymbol{w} = \boldsymbol{b}(\theta), \tag{4-2}$$

where $\boldsymbol{R}$ is the covariance matrix and $\boldsymbol{b}(\theta)$ is the initial vector defined as

$$\boldsymbol{b}(\theta) = \frac{\boldsymbol{R}\,\boldsymbol{a}(\theta)}{\|\boldsymbol{R}\,\boldsymbol{a}(\theta)\|}, \tag{4-3}$$

where $\boldsymbol{a}(\theta)$ is the search vector.

The extended signal subspace of rank P is obtained by means of the CG algorithm, which is summarized in the Table 4.1. The set of orthogonal residual vectors

$$\boldsymbol{G}_{cg,P+1}(\theta) = [\boldsymbol{g}_{cg,0}(\theta), \boldsymbol{g}_{cg,1}(\theta), \ldots, \boldsymbol{g}_{cg,P}(\theta)], \tag{4-4}$$

where $\boldsymbol{b}(\theta)= \boldsymbol{g}_0(\theta)$ generates the well-known extended Krylov subspace comprised of the true signal subspace of dimension P and the search vector itself. All the residual vectors are normalized except for the last one. If $\theta \in \{\theta_1, \ldots, \theta_P\}$, the initial vector $\boldsymbol{b}(\theta)$ lies in the true signal subspace space spanned by the $[\boldsymbol{g}_{cg,0}(\theta), \boldsymbol{g}_{cg,1}(\theta), \ldots, \boldsymbol{g}_{cg,P-1}(\theta)]$ basis vectors of the extended Krylov subspace. Thus, the rank of the generated signal subspace drops from $P+1$ to P and we have

$$\boldsymbol{g}_{cg,P}(\theta) = 0, \tag{4-5}$$

where $\boldsymbol{g}_{cg,P}$ is the last unnormalized residual vector. In order to exploit this behavior, the proposed KA-CG algorithm makes use of the spectral function defined in [13]:

$$\mathcal{P}_{\mathcal{K}}(\theta^{(n)}) = \frac{1}{\|\boldsymbol{g}_{cg,P}^H(\theta^{(n)})\boldsymbol{G}_{cg,P+1}(\theta^{(n-1)})\|^2}, \tag{4-6}$$

where $\theta^{(n)}$denotes the search angle in the whole angle range $\{-90^o, \ldots, 90^o\}$ with $\theta^{(n)} = n\Delta^o - 90^o$, where Δ^o is the search step and $n = 0, 1, \ldots, 180^o/\Delta^o$. The matrix $\boldsymbol{G}_{cg,P+1}(\theta^{(n-1)})$ contains all residual vectors at the $(n-1)$th vector calculated at the current search step n. If $\theta^{(n)} \in \{\theta_1, \ldots, \theta_P\}$, $\boldsymbol{g}_{cg,P}(\theta^{(n)}) = 0$ and we can expect a peak in the spectrum. Taking into account that $\hat{\boldsymbol{R}}_o$ in (2-52) and (4-1) is only a sample average estimate, which is unknown in practical applications, $\boldsymbol{g}_{cg,P}(\theta^{(n)})$ and $\boldsymbol{G}_{cg,P+1}(\theta^{(n-1)})$ become approximations. Hence the spectral function in (4-6) can just provide very large values but they do not tend to infinity as for the original covariance matrix.

The superscript $(\cdot)^{(1)}$ refers to the estimation task performed in the first step. Now, a procedure consisting of $n = 1 : I$ iterations starts by forming the Vandermonde matrix using the DOA estimates. Then, the amplitudes of the sources are estimated such that the square norm of the differences between the observation vector and the vector containing estimates and the available known DOAs is minimized. This problem can be formulated [44] as:

$$\hat{\boldsymbol{s}}(i) = \arg\min_{\boldsymbol{s}} \| \boldsymbol{x}(i) - \hat{\boldsymbol{A}}\boldsymbol{s} \|_2^2 . \tag{4-7}$$

The minimization of (4-7) is achieved using the least squares technique and the solution is described by

$$\hat{\boldsymbol{s}}(i) = (\hat{\boldsymbol{A}}^H \hat{\boldsymbol{A}})^{-1} \hat{\boldsymbol{A}} \boldsymbol{x}(i) \tag{4-8}$$

The noise component is then estimated as the difference between the estimated

Table 4.1: Summary of the Conjugate Gradient Algorithm

$\boldsymbol{w}_0 = 0,\ \boldsymbol{d}_1 = \boldsymbol{g}_{cg,0} = b,\ \rho_0 = \boldsymbol{g}_{cg,0}^H \boldsymbol{g}_{cg,0}$
for i=1 to P
$\boldsymbol{v}_i = \boldsymbol{R}\,\boldsymbol{d}_i$
$\alpha_i = \rho_{i-1} \,/\, \boldsymbol{d}_i^H \boldsymbol{v}_i$
$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \boldsymbol{\alpha}_i \boldsymbol{d}_i$
$\boldsymbol{g}_{cg,i} = \boldsymbol{g}_{cg,i-1}$ - $\boldsymbol{\alpha}_i \boldsymbol{v}_i$
$\rho_i = \boldsymbol{g}_{cg,i}^H \boldsymbol{g}_{cg,i}$
$\beta_i = \rho_i \,/\, \rho_{i-1} = \|\boldsymbol{g}_{cg,i}\|^2 \,/\, \|\boldsymbol{g}_{cg,i-1}\|^2$
$\boldsymbol{d}_{i+1} = \boldsymbol{g}_{cg,i} + \boldsymbol{\beta}_i \boldsymbol{d}_i$
end for
form $\boldsymbol{G}_{cg,P+1}(\theta)$ (4-4)
compute $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ (4-6)
find $\hat{P}$ largest peaks of $\mathcal{P}_{\mathcal{K}}(\theta^{(n)})$ to obtain estimates $\hat{\theta}_l$ of the DOA

signal and the observations made by the array, as given by

$$\hat{\boldsymbol{n}}(i) = \boldsymbol{x}(i) \ - \ \hat{\boldsymbol{A}}\,\hat{\boldsymbol{s}}(i). \tag{4-9}$$

After estimating the signal and noise vectors, the third term in (4-1) can be computed as

$$\begin{aligned}
\boldsymbol{V} &\triangleq \hat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{s}}(i) \hat{\boldsymbol{n}}^H(i) \right\} \\
&= \hat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{i=1}^{N} (\hat{\boldsymbol{A}}^H\ \hat{\boldsymbol{A}})^{-1} \hat{\boldsymbol{A}}^H \boldsymbol{x}(i) \right. \\
&\left. \times (\boldsymbol{x}^H(i) - \boldsymbol{x}^H(i) \hat{\boldsymbol{A}} (\hat{\boldsymbol{A}}^H \hat{\boldsymbol{A}})^{-1}\ \hat{\boldsymbol{A}}^H) \right\} \\
&= \hat{\boldsymbol{Q}}_A \left\{ \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}(i) \boldsymbol{x}^H(i)\ \left(\boldsymbol{I}_M \ - \ \hat{\boldsymbol{Q}}_A \right) \right\} \\
&= \hat{\boldsymbol{Q}}_A\, \hat{\boldsymbol{R}}\, \hat{\boldsymbol{Q}}_A^{\perp},
\end{aligned} \tag{4-10}$$

where

$$\hat{\boldsymbol{Q}}_A \triangleq \hat{\boldsymbol{A}}\ (\hat{\boldsymbol{A}}^H\ \hat{\boldsymbol{A}})^{-1}\ \hat{\boldsymbol{A}}^H \tag{4-11}$$

is an estimate of the projection matrix of the signal subspace, and

$$\hat{\boldsymbol{Q}}_A^{\perp} \triangleq \boldsymbol{I}_M \ - \ \hat{\boldsymbol{Q}}_A \tag{4-12}$$

is an estimate of the projection matrix of the noise subspace.

Subsequently, as part of the procedure with $n = 1 : I$ iterations, the modified data covariance matrix $\hat{\boldsymbol{R}}^{(n+1)}$ is calculated by computing a scaled version of the estimated terms from the initial sample data covariance matrix as given

$$\hat{\boldsymbol{R}}^{(n+1)} = \hat{\boldsymbol{R}}_o \; - \; \mu \, (\boldsymbol{V}^{(n)} \; + \; \boldsymbol{V}^{(n)H}), \tag{4-13}$$

where the superscript $(\cdot)^{(n)}$ refers to the n^{th} iteration performed. The scaling or reliability factor μ increases from 0 to 1 incrementally, resulting in modified data covariance matrices. Each of them gives origin to new estimated DOAs also denoted by the superscript $(\cdot)^{(n+1)}$ by using the CG algorithm, which was previously described. Then, a new Vandermonde matrix $\hat{\boldsymbol{B}}^{(n+1)}$ is formed by the steering vectors of those newer estimated DOAs. By using this new matrix, it is possible to compute the newer estimates of the projection matrices of the signal $\hat{\boldsymbol{Q}}_B^{(n+1)}$ and the noise $\hat{\boldsymbol{Q}}_B^{(n+1)\perp}$ subspaces.

Next, employing the refined estimates of the projection matrices, the initial sample data matrix, $\hat{\boldsymbol{R}}_o$, and the number of sensors and sources, the stochastic maximum likelihood objective function (SMLOF) $U^{(n+1)}(\mu)$ [45] is computed for each value of μ at the n^{th} iteration, as follows:

$$U^{(n+1)}(\mu) = \ln \det (\cdot) \, , \tag{4-14}$$

where

$$(\cdot) = \left(\hat{\boldsymbol{Q}}_B^{(n+1)} \, \hat{\boldsymbol{R}}_o \, \hat{\boldsymbol{Q}}_B^{(n+1)} + \frac{\text{Trace}\{\hat{\boldsymbol{Q}}_B^{\perp \, (n+1)} \, \hat{\boldsymbol{R}}_o\}}{M - P} \, \hat{\boldsymbol{Q}}_B^{(n+1)\perp} \right)$$

The previous computation selects the set of unavailable DOA estimates that have a higher likelihood at each iteration. Then, the set of estimated DOAs corresponding to the optimum value of μ that minimizes (4-14) also at each n^{th} iteration is determined. Finally, the output of the proposed MS-KAI-CG algorithm is formed by the set of the estimates obtained at the I^{th} iteration, as described in Table 4.2.

4.3

Proposed MS-KAI CG-FB Algorithm

Most direction finding algorithms experience performance degradation in the presence of correlated signals. This is also verified for the proposed MS-KAI-CG algorithm, as will be shown later via simulations. In this section, we present an approach the combines the proposed MS-KAI-CG algorithm and the well-known forward-backward spatial smoothing (FBSS) [38,63,64] technique, denoted as MS-KAI-CG-FB algorithm, for dealing with correlated signals. In

Table 4.2: Proposed MS-KAI-CG Algorithm

Inputs:
$M,\ d,\ \lambda,\ N,\ I$
Received vectors $\boldsymbol{x}(1),\ \boldsymbol{x}(2),\cdots,\boldsymbol{x}(N)$
Outputs:
Estimates $\hat{\theta}_1^{(n+1)}(\mu\ opt),\ \hat{\theta}_2^{(n+1)}(\mu\ opt),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu\ opt)$
First step:
$\hat{\boldsymbol{R}}_o = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}(i)\boldsymbol{x}^H(i)$
$\{\hat{\theta}_1^{(1)},\ \hat{\theta}_2^{(1)},\cdots,\hat{\theta}_P^{(1)}\}\ \underleftarrow{CG}\ (\hat{\boldsymbol{R}}_o, P, d, \lambda)$
$\hat{\boldsymbol{A}}^{(1)} = \left[\boldsymbol{a}(\hat{\theta}_1^{(1)}), \boldsymbol{a}(\hat{\theta}_2^{(1)}),\cdots,\boldsymbol{a}(\hat{\theta}_P^{(1)})\right]$
Second step:
for $n = 1 : I$
$\hat{\boldsymbol{Q}}_A^{(n)} = \hat{\boldsymbol{A}}^{(n)}\ (\hat{\boldsymbol{A}}^{(n)H}\ \hat{\boldsymbol{A}}^{(n)})^{-1}\ \hat{\boldsymbol{A}}^{(n)H}$
$\hat{\boldsymbol{Q}}_A^{(n)\perp} = \boldsymbol{I}_M\ -\ \hat{\boldsymbol{Q}}_A^{(n)}$
$\boldsymbol{V}^{(n)} = \hat{\boldsymbol{Q}}_A^{(n)}\ \hat{\boldsymbol{R}}_o\ \hat{\boldsymbol{Q}}_A^{(n)\perp}$
for $\boldsymbol{\mu} =$ 0 : increment : 1
$\hat{\boldsymbol{R}}^{(n+1)} = \hat{\boldsymbol{R}}_o\ -\ \mu\ (\boldsymbol{V}^{(n)}\ +\ \boldsymbol{V}^{(n)H})$
$\{\hat{\theta}_1^{(n+1)},\ \hat{\theta}_2^{(n+1)},\cdots,\hat{\theta}_P^{(n+1)}\}\ \underleftarrow{CG}\ (\hat{\boldsymbol{R}}^{(n+1)},\ P, d, \lambda)$
$\hat{\boldsymbol{B}}^{(n+1)} = \left[\boldsymbol{a}(\hat{\theta}_1^{(n+1)}), \boldsymbol{a}(\hat{\theta}_2^{(n+1)}),\cdots,\boldsymbol{a}(\hat{\theta}_P^{(n+1)})\right]$
$\hat{\boldsymbol{Q}}_B^{(n+1)} = \hat{\boldsymbol{B}}^{(n+1)}\ (\hat{\boldsymbol{B}}^{(n+1)H}\ \hat{\boldsymbol{B}}^{(n+1)})^{-1}\ \hat{\boldsymbol{B}}^{(n+1)H}$
$\hat{\boldsymbol{Q}}_B^{(n+1)\perp} = \boldsymbol{I}_M\ -\ \hat{\boldsymbol{Q}}_B^{(n+1)}$
$$U^{(n+1)}(\mu) = \ln\det\left(\hat{\boldsymbol{Q}}_B^{(n+1)}\ \hat{\boldsymbol{R}}_o\ \hat{\boldsymbol{Q}}_B^{(n+1)} + \frac{\mathrm{Trace}\{\hat{\boldsymbol{Q}}_B^{\perp\,(n+1)}\ \hat{\boldsymbol{R}}_o\}}{M-P}\ \hat{\boldsymbol{Q}}_B^{(n+1)\perp}\right)$$
$\mu_o^{(n+1)} = \arg\min\ U^{(n+1)}(\mu)$
$\mathrm{DOAs}^{(n+1)} = \{\hat{\theta}_1^{(n+1)}(\mu_o),\ \hat{\theta}_2^{(n+1)}(\mu_o),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu_o)\}$
if $n <=\ P$
$\hat{\boldsymbol{A}}^{(n+1)} = \left\{\boldsymbol{a}(\hat{\theta}_{\{1,\cdots,n\}}^{(n+1)}(\mu_o))\right\} \bigcup \left\{\boldsymbol{a}(\hat{\theta}_{\{1,\cdots,P\}-\{1,\cdots,n\}}^{(1)})\right\}$
else
$\hat{\boldsymbol{A}}^{(n+1)} = \left[\boldsymbol{a}(\hat{\theta}_1^{(n+1)}(\mu_o)), \boldsymbol{a}(\hat{\theta}_2^{(n+1)}(\mu_o)),\cdots,\boldsymbol{a}(\hat{\theta}_P^{(n+1)}(\mu_o))\right]$
end if
end for
end for

---

the proposed MS-KAI-CG-FB algorithm, the FBSS covariance matrix (4-15) is obtained from the initial sample covariance matrix $\hat{\boldsymbol{R}}_o$ (2-52), as follows:

$$\hat{\boldsymbol{R}} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{Z}_K\ \tilde{\boldsymbol{R}}\ \boldsymbol{Z}_k^T, \tag{4-15}$$

where the number of its subarrays is obtained by

$$K = M - L + 1, \tag{4-16}$$

In (4-16), L means the number of sensors of the subarrays and M the number of sensors of the original ULA. The matrix $\boldsymbol{Z}_k$ is given by

$$\boldsymbol{Z}_k = \left[\boldsymbol{0}_{L\times(k-1)} \mid \boldsymbol{I}_{L\times L} \mid \boldsymbol{0}_{L\times M-(L+k-1)} \right] \tag{4-17}$$

The forward-backward modified matrix $\tilde{\boldsymbol{R}}$ is defined as:

$$\tilde{\boldsymbol{R}} = \frac{1}{2}\left(\hat{\boldsymbol{R}}_o \, + \, \boldsymbol{J}\, \hat{\boldsymbol{R}}_o^{*}\, \boldsymbol{J}\right), \tag{4-18}$$

where $\boldsymbol{J}$ is is an off-diagonal exchange matrix

$$\boldsymbol{J} = \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}, \tag{4-19}$$

and $(*)$ means the complex conjugate.

Next, we expand (4-15) using (2-49) as follows:

$$\begin{aligned} \hat{\boldsymbol{R}} &= \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{A}\, s(i) + \boldsymbol{n}(i))\,(\boldsymbol{A}\, s(i) + \boldsymbol{n}(i))^H \\ &= \boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{n}^H(i) \\ &+ \underbrace{\boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{n}^H(i)\right\} + \left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H}_{\text{"undesirable by-products"}} \end{aligned} \tag{4-20}$$

The first two terms of $\hat{\boldsymbol{R}}$ in (4-20) can be considered as estimates of the two summands of $\boldsymbol{R}$ given in (2-51), which represent the signal and the noise components, respectively. The last two terms in (4-20) are undesirable by-products, which can be seen as estimates for the correlation between the signal and the noise vectors. Moreover, the last two terms in (4-20) may have large values, which results in estimates of the signal and the noise subspaces different from the actual subspaces.

The key aspect of the proposed MS-KAI-CG-FB algorithm is to modify the FBSS covariance matrix estimate $\hat{\boldsymbol{R}}$ (4-15) at each iteration by gradually incorporating the knowledge provided by the newer Vandermonde matrices which progressively embody the newer estimates from the preceding iteration. Based on these updated Vandermonde matrices, refined estimates of the projection matrices of the signal and noise subspaces are calculated. These estimates of projection matrices associated with the FBSS covariance matrix estimate estimate $\hat{\boldsymbol{R}}$ and the reliability factor employed to reduce its by-products allow to choose the set of estimates that has the minimum value

of the SMLOF, i.e., the highest likelihood of being the set of the true DOAs. The modified covariance matrix is computed by deriving a scaled version of the undesirable terms from $\hat{\boldsymbol{R}}$, which are pointed out in (4-20).

The steps of the proposed MS-KAI-CG-FB algorithm are listed in Table 4.3. The algorithm starts by computing the initial sample data covariance matrix (2-52). Then, the FBSS covariance matrix estimate (4-15) is calculated. Subsequently, the DOAs are estimated using the original CG algorithm described in Subsection 2.1.6 and summarized in section 4.2. The superscript $(\cdot)^{(1)}$ refers to the estimation task performed in the first step. Next, a procedure consisting of $n = 1 : I$ iterations starts by forming the Vandermonde matrix using the DOA estimates. Then, the amplitudes of the sources are estimated such that the square norm of the differences between the observation vector and the vector containing estimates and the available known DOAs is minimized. This problem can be formulated as

$$\hat{\boldsymbol{s}}(i) = \arg\min_{\boldsymbol{s}} \parallel \boldsymbol{x}(i) - \hat{\boldsymbol{A}}\boldsymbol{s} \parallel_2^2 . \tag{4-21}$$

The minimization of (4-21) is achieved using the least squares technique and the solution is described by

$$\hat{\boldsymbol{s}}(i) = (\hat{\boldsymbol{A}}^H \hat{\boldsymbol{A}})^{-1} \hat{\boldsymbol{A}} \boldsymbol{x}(i) \tag{4-22}$$

The noise component is then estimated as the difference between the estimated signal and the observations made by the array, as given by

$$\hat{\boldsymbol{n}}(i) = \boldsymbol{x}(i) - \hat{\boldsymbol{A}} \hat{\boldsymbol{s}}(i). \tag{4-23}$$

After estimating the signal and noise vectors, the third term in (4-20) can be computed as

$$\begin{aligned}
\boldsymbol{V} &\triangleq \hat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{s}}(i) \hat{\boldsymbol{n}}^H(i) \right\} \\
&= \hat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{i=1}^{N} (\hat{\boldsymbol{A}}^H \hat{\boldsymbol{A}})^{-1} \hat{\boldsymbol{A}}^H \boldsymbol{x}(i) \right. \\
&\left. \times (\boldsymbol{x}^H(i) - \boldsymbol{x}^H(i) \hat{\boldsymbol{A}} (\hat{\boldsymbol{A}}^H \hat{\boldsymbol{A}})^{-1} \hat{\boldsymbol{A}}^H) \right\} \\
&= \hat{\boldsymbol{Q}}_A \left\{ \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}(i) \boldsymbol{x}^H(i) \left(\boldsymbol{I}_M - \hat{\boldsymbol{Q}}_A \right) \right\} \\
&= \hat{\boldsymbol{Q}}_A \hat{\boldsymbol{R}} \hat{\boldsymbol{Q}}_A^{\perp},
\end{aligned} \tag{4-24}$$

where

$$\hat{\boldsymbol{Q}}_A \triangleq \hat{\boldsymbol{A}} (\hat{\boldsymbol{A}}^H \hat{\boldsymbol{A}})^{-1} \hat{\boldsymbol{A}}^H \tag{4-25}$$

is an estimate of the projection matrix of the signal subspace, and

$$\hat{\boldsymbol{Q}}_A^{\perp} \triangleq \boldsymbol{I}_M \; - \; \hat{\boldsymbol{Q}}_A \tag{4-26}$$

is an estimate of the projection matrix of the noise subspace.

Next, also as part of the procedure of $n = 1 : I$ iterations, the modified data covariance matrix $\hat{\boldsymbol{R}}^{(n+1)}$ is calculated by computing a scaled version of the estimated terms from the FBSS covariance matrix estimate (4-15), as given

$$\hat{\boldsymbol{R}}^{(n+1)} = \hat{\boldsymbol{R}} \; - \; \mu \, (\boldsymbol{V}^{(n)} \; + \; \boldsymbol{V}^{(n)H}), \tag{4-27}$$

where the superscript $(\cdot)^{(n)}$ refers to the n^{th} iteration performed. The scaling or reliability factor μ increases from 0 to 1 incrementally, resulting in modified data covariance matrices. Each of them gives origin to new estimated DOAs also denoted by the superscript $(\cdot)^{(n+1)}$ by using the standard CG algorithm, which was described in Subsection 2.1.6. Then, a new Vandermonde matrix $\hat{\boldsymbol{B}}^{(n+1)}$ is formed by the steering vectors of those newer estimated DOAs. By using this new matrix, it is possible to compute the newer estimates of the projection matrices of the signal $\hat{\boldsymbol{Q}}_B^{(n+1)}$ and the noise $\hat{\boldsymbol{Q}}_B^{(n+1)\perp}$ subspaces.

Afterwards, employing the newer estimates of the projection matrices, the FBSS covariance matrix estimate $\hat{\boldsymbol{R}}$ (4-15), the number of sensors of the subarrays obtained in the FBSS technique L (4-16) and the number of the sources P, the stochastic maximum likelihood objective function $U^{(n+1)}(\mu)$ [45] is computed for each value of μ at the n^{th} iteration, as follows:

$$U^{(n+1)}(\mu) = \ln \det (\cdot) \, , \tag{4-28}$$

where

$$(\cdot) = \left(\hat{\boldsymbol{Q}}_B^{(n+1)} \, \hat{\boldsymbol{R}} \, \hat{\boldsymbol{Q}}_B^{(n+1)} + \frac{\mathrm{Trace}\{\hat{\boldsymbol{Q}}_B^{\perp \, (n+1)} \, \hat{\boldsymbol{R}}\}}{L - P} \, \hat{\boldsymbol{Q}}_B^{(n+1)\perp} \right)$$

The preceding computation selects the set of unavailable DOA estimates that have a higher likelihood at each iteration. Then, the set of estimated DOAs corresponding to the optimum value of μ that minimizes (4-28) also at each n^{th} iteration is determined. Finally, the output of the proposed MS-KAI-CG-FB algorithm is formed by the set of the estimates obtained at the I^{th} iteration, as described in Table 4.3.

Table 4.3: Proposed MS-KAI-CG-FB Algorithm

Inputs:
$M,\ d,\ \lambda, N,\ P$
Received vectors $\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(N)$
Outputs:
Estimates $\hat{\theta_1}^{(n+1)}(\mu\ opt), \hat{\theta_2}^{(n+1)}(\mu\ opt), \cdots, \hat{\theta_P}^{(n+1)}(\mu\ opt)$
First step:
$\hat{\boldsymbol{R}}_{\boldsymbol{o}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}(i)\boldsymbol{x}^H(i)$
$\hat{\boldsymbol{R}} \ \underleftarrow{FBSS}\ \hat{\boldsymbol{R}}_{\boldsymbol{o}}$
$\{\hat{\theta_1}^{(1)},\ \hat{\theta_2}^{(1)}, \cdots, \hat{\theta_P}^{(1)}\} \ \underleftarrow{CG}\ (\hat{\boldsymbol{R}}, P, d, \lambda)$
$\hat{\boldsymbol{A}}^{(1)} = \left[\boldsymbol{a}(\hat{\theta_1}^{(1)}), \boldsymbol{a}(\hat{\theta_2}^{(1)}), \cdots, \boldsymbol{a}(\hat{\theta_P}^{(1)})\right]$
Second step:
for $n = 1 : I$
$\hat{\boldsymbol{Q}}_A^{(n)} = \hat{\boldsymbol{A}}^{(n)} (\hat{\boldsymbol{A}}^{(n)H} \hat{\boldsymbol{A}}^{(n)})^{-1} \hat{\boldsymbol{A}}^{(n)H}$
$\hat{\boldsymbol{Q}}_A^{(n)\perp} = \boldsymbol{I}_M - \hat{\boldsymbol{Q}}_A^{(n)}$
$\boldsymbol{V}^{(n)} = \hat{\boldsymbol{Q}}_A^{(n)} \hat{\boldsymbol{R}} \hat{\boldsymbol{Q}}_A^{(n)\perp}$
for $\boldsymbol{\mu} = 0 : \iota : 1$
$\hat{\boldsymbol{R}}^{(n+1)} = \hat{\boldsymbol{R}} - \mu (\boldsymbol{V}^{(n)} + \boldsymbol{V}^{(n)H})$
$\{\hat{\theta_1}^{(n+1)},\ \hat{\theta_2}^{(n+1)}, \cdots, \hat{\theta_P}^{(n+1)}\} \ \underleftarrow{CG}\ (\hat{\boldsymbol{R}}^{(n+1)},\ P, d, \lambda)$
$\hat{\boldsymbol{B}}^{(n+1)} = \left[\boldsymbol{a}(\hat{\theta_1}^{(n+1)}), \boldsymbol{a}(\hat{\theta_2}^{(n+1)}), \cdots, \boldsymbol{a}(\hat{\theta_P}^{(n+1)})\right]$
$\hat{\boldsymbol{Q}}_B^{(n+1)} = \hat{\boldsymbol{B}}^{(n+1)} (\hat{\boldsymbol{B}}^{(n+1)H} \hat{\boldsymbol{B}}^{(n+1)})^{-1} \hat{\boldsymbol{B}}^{(n+1)H}$
$\hat{\boldsymbol{Q}}_B^{(n+1)\perp} = \boldsymbol{I}_M - \hat{\boldsymbol{Q}}_B^{(n+1)}$
$U^{(n+1)}(\mu) = \ln \det \left(\hat{\boldsymbol{Q}}_B^{(n+1)} \hat{\boldsymbol{R}} \hat{\boldsymbol{Q}}_B^{(n+1)} + \frac{\text{Trace}\{\hat{\boldsymbol{Q}}_B^{\perp\,(n+1)} \hat{\boldsymbol{R}}\}}{L - P} \hat{\boldsymbol{Q}}_B^{(n+1)\perp} \right)$
$\mu_{\text{o}}^{(n+1)} = \arg\min\ U^{(n+1)}(\mu)$
$\text{DOAs}^{(n+1)} = \{\hat{\theta_1}^{(n+1)}(\mu_o),\ \hat{\theta_2}^{(n+1)}(\mu_o), \cdots,\ \hat{\theta_P}^{(n+1)}(\mu_o)\}$
if $n <= P$
$\hat{\boldsymbol{A}}^{(n+1)} = \left\{\boldsymbol{a}(\hat{\theta}^{(n+1)}_{\{1,\cdots,n\}}(\mu_o))\right\} \bigcup \left\{\boldsymbol{a}(\hat{\theta}^{(1)}_{\{1,\cdots,P\}-\{1,\cdots,n\}})\right\}$
else
$\hat{\boldsymbol{A}}^{(n+1)} = \left[\boldsymbol{a}(\hat{\theta_1}^{(n+1)}(\mu_o)), \boldsymbol{a}(\hat{\theta_2}^{(n+1)}(\mu_o)), \cdots, \boldsymbol{a}(\hat{\theta_P}^{(n+1)}(\mu_o))\right]$
end if
end for
end for

4.4

Computational Complexity Analysis

In this section, we evaluate the computational cost of the proposed MS-KAI-CG and MS-KAI-CG-FB [66] algorithms which are compared to the

following classical subspace methods: ESPRIT [10], MUSIC [8], Root-MUSIC [9], Conjugate Gradient (CG) [15], Auxiliary Vector Filtering (AVF) [13] and TS-ESPRIT [47]. The ESPRIT and MUSIC-based methods use the Singular Value Decomposition (SVD) of the sample covariance matrix (2-59). The computational complexity of MS-KAI-CG/MS-KAI-CG-FB in terms of number of multiplications is depicted in Table 4.4, where $\tau = \frac{1}{\iota} + 1$. The increment ι is defined in Table 4.3.

Considering the number of multiplications, it can be seen that for this specific configuration used in the simulations 3.4 MS-KAI-CG/MS-KAI-CG-FB show a relatively high computational burden with $\mathcal{O}(P\tau\left[\frac{180}{\Delta}\left(M^2\left(P+1\right)+M\left(6P+2\right)\right)\right])$, where τ is typically an integer that ranges from 2 to 20. Similarly, the order of additions reaches $\mathcal{O}(P\tau\left[\frac{180}{\Delta}\left(M^2\left(P+1\right)+M\left(5P+1\right)\right)\right])$. By examining the expressions for multiplications and additions for the proposed algorithms, that the number of multiplications required by the proposed algorithms is more significant than the number of additions and serves as an appropriate indicator of the computational complexity of the proposed and existing algorithms. For this reason, in Table 4.4, we consider the computational burden of the proposed and previously reported algorithms in terms of multiplications for the purpose of comparisons. In that table, Δ stands for the search step.

Next, based on Table 4.4, we have evaluated the influence of the number of sensor elements on the number of multiplications based on the specific configuration composed of $P = 4$ narrowband signals impinging on a ULA of M sensor elements and $N = 100$ available snapshots. In Fig. 4.1, we can see the main trends in terms of computational cost measured in multiplications of the proposed and analyzed algorithms. By examining Fig. 4.1, it can be noticed that in the whole range $M = [0\ 100]$ sensors, the curves describing the exact number of multiplications computed in MS-KAI-CG-FB and MS-KAI-CG has been merged, which means equivalent burden in terms of this kind of operation. It can also be noticed that in the range $M = [5\ 20]$, MS-KAI-CG-FB, MS-KAI-CG and MS-KAI-ESPRIT require a similar cost.

4.5

Simulations

In this section, we evaluate the performance of the proposed MS-KAI CG-FB and MS-KAI-CG algorithms, the standard CG [15, 16, 27] and the forward-

Table 4.4: Computational complexity in multiplications of algorithms

MS-KAI-CG [65] ∼ MS-KAI-CG -FB [65]	$P\,\tau\{\frac{180}{\Delta}[M^2(P+1)+M(6P+2)+P+1]$ $+\frac{10}{3}M^3+M^2(N+P+3)+M(\frac{3}{2}P^2+\frac{1}{2}P)$ $+P^2(\frac{1}{2}P+\frac{3}{2})\}$ $+P\,[2M^3+M^2(P)+M(\frac{1}{2}P)+P^2(\frac{P}{2}+\frac{3}{2})]$ $+\frac{180}{\Delta}[M^2(P+1)+M(6P+2)+P+1]$ $+M^2(N+2)+MP$
MUSIC [8]	$\frac{180}{\Delta}[M^2+M(2-P)-P]+8MN^2$
Root-MUSIC [9]	$2M^3-M^2P+8MN^2$
AVF [13]	$\frac{180}{\Delta}[M^2(3P+1)+M(4P-2)+P+2]$ $+M^2N$
CG [15]	$\frac{180}{\Delta}[M^2(P+1)+M(6P+2)+P+1]+M^2N$
ESPRIT [10]	$2M^2P+M(P^2-2P+8N^2)+8P^3-P^2$
TS-ESPRIT [47]	$\tau[3M^3+M^2(3P+2)+M(\frac{5}{2}P^2-\frac{3}{2}P+8N^2)$ $+P^2(\frac{17}{2}P+\frac{1}{2})+1]$ $+[2M^3+M^2(3P)+M(\frac{5}{2}P^2-\frac{3}{2}P+8N^2)$ $+P^2(\frac{17}{2}P+\frac{1}{2})]$

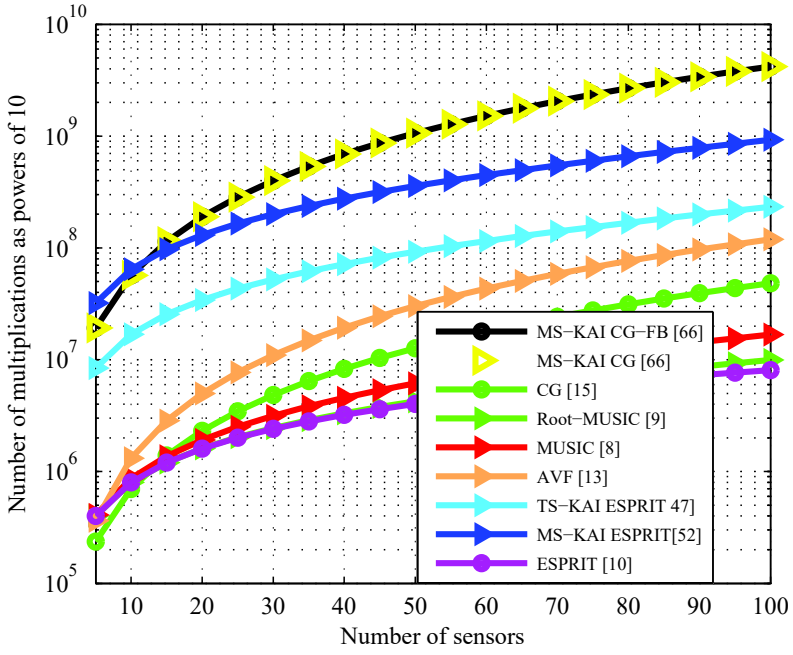

Figure 4.1: *Number of multiplications as powers of 10 versus number of sensors for* $P = 4$, $N = 100$.

backward spatially smoothed CG (CG-FB) [15,38], the ESPRIT [10], and the MUSIC [8] algorithms in terms of RMSE and probability of resolution (PR).

The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{S\,P}\sum_{s=1}^{S}\sum_{p=1}^{P}(\theta_p - \hat{\theta}_p(s))^2}, \tag{4-29}$$

where S is the number of trials. The signal-to-noise ratio (SNR) is defined as $SNR \triangleq 10\log_{10}\left(\frac{\sigma_s^2}{\sigma_n^2}\right)$. For comparisons in terms of RMSE (dB), we have computed and plotted the square root of the deterministic CRB [28]. To assess the performance in terms of PR, we take into account the criterion of [29], in which two sources with DOA θ_1 and θ_2 are said to be resolved if their respective estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are such that both $\left|\hat{\theta}_1 - \theta_1\right|$ and $\left|\hat{\theta}_2 - \theta_2\right|$ are less than $|\theta_1 - \theta_2|/2$. We have set the search step to $\Delta = 0.2^o$ in all algorithms that make use of peak search. We first consider a scenario with $P = 2$ uncorrelated complex Gaussian signals with equal power impinging on a ULA with $N = 12$ sensors. The sources have been separated by $\xi(\theta) = 2.0^o$, at $(15^o, 17^o)$, and the number of available snapshots was set to $N = 100$. The computations of RMSE have used 150 independent trials. In Fig. 4.2, we show the PR against the SNR, whereas in Fig. 4.3 the RMSE performance against the SNR is depicted. From the curves it can be noticed the improvement of the performance of MS-KAI-CG in terms of both PR and RMSE as a result of the improved covariance matrix estimates.

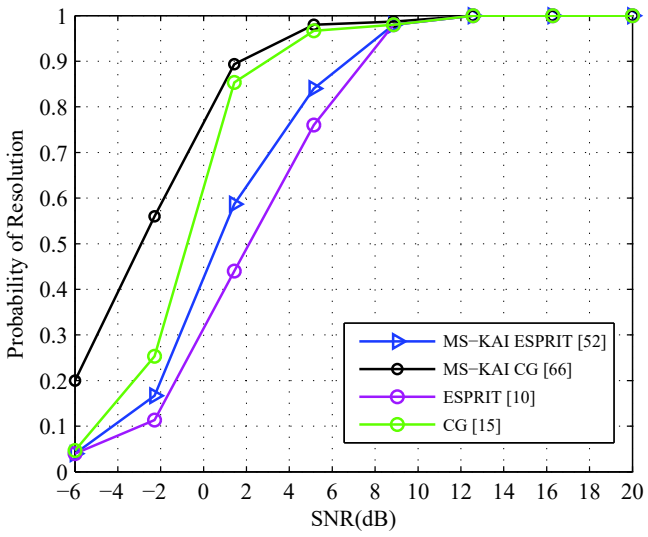

Figure 4.2: *Probability of resolution versus SNR with* $P = 2$, $M = 12$, $N = 100$, $L = 150$ *runs*, $\xi(\theta) = 2.0^o$.

In Fig. 4.4, we show the influence of the iterations carried out at the second step. It can be noticed the gradual and consistent improvement of the performance of MS-KAI-CG in terms of RMSE as a result of the increasing number of iterations.

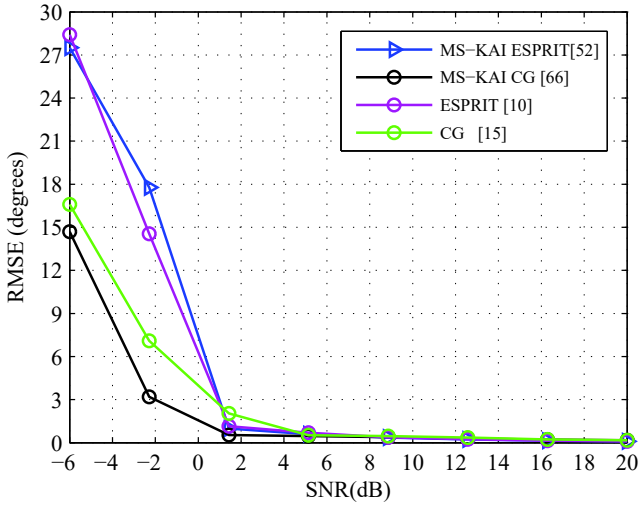

Figure 4.3: *RMSE in degrees versus SNR with* $P = 2$, $M = 12$, $N = 100$, $L = 150$ *runs*, $\xi(\theta) = 2.0^o$.

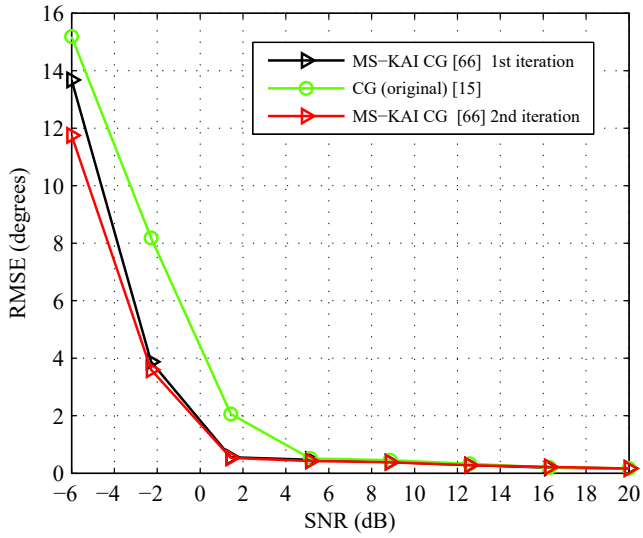

Figure 4.4: *Influence of the iterations in terms of RMSE in degrees versus SNR with* $P = 2$, $M = 12$, $N = 100$, $L = 150$ *runs*, $\xi(\theta) = 2.0^o$.

In the next examples, we have examined the performance of the proposed MS-KAI-CG-FB when employed to estimate strongly correlated closely spaced sources. To this end, we consider a scenario composed of Gaussian signals with equal power impinging on a ULA. In particular, we have $P = 2$ sources separated by $\xi(\theta) = 2.0^o$, at $(15^o, 17^o)$, $M = 12$ sensors and $N = 70$ snapshots. We have employed $L = 150$ trials for these simulations. The source signals have been correlated according to the following correlation matrix:

$$\boldsymbol{R}_{ss} = \sigma_s^2 \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}. \tag{4-30}$$

In Fig. 4.5, we can notice that in terms of PR the proposed MS-KAI-CG-FB outperforms the standard CG algorithm equipped with forward-backward spatial smoothing, denoted as CG-FB, the standard CG algorithm, MUSIC and ESPRIT in most of the considered range of SNR values. In Fig. 4.6, we

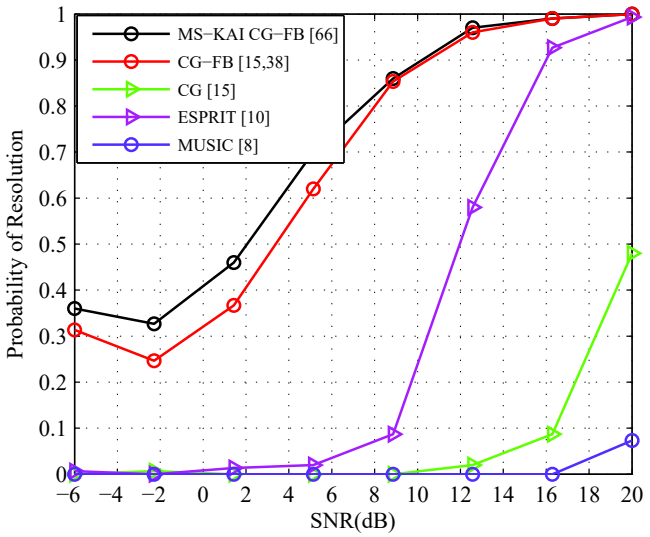

Figure 4.5: *Probability of resolution versus SNR with* $P = 2$, $M = 12$, $N = 70$, $L = 150$ *runs*, $\xi(\theta) = 2.0^o$.

can see that in terms of RMSE the proposed MS-KAI-CG-FB provides the best performance in the range [1.8 16] dB. It can also be seen that in the ranges [−6 1.8) dB and (16 20] dB its performance is similar to the best. This performance can be better noticed in Fig. 4.7, which shows the RMSE in terms of dB.

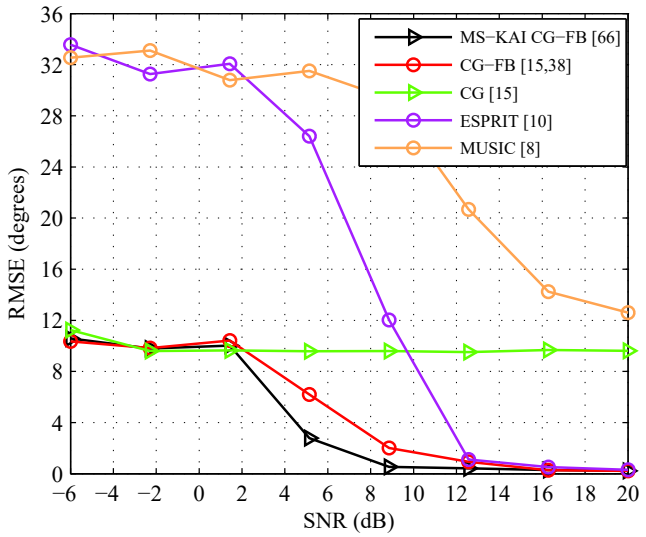

Figure 4.6: *RMSE in degrees versus SNR with* $P = 2$, $M = 12$, $N = 70$, $L = 150$ *runs*, $\xi(\theta) = 2.0^o$.

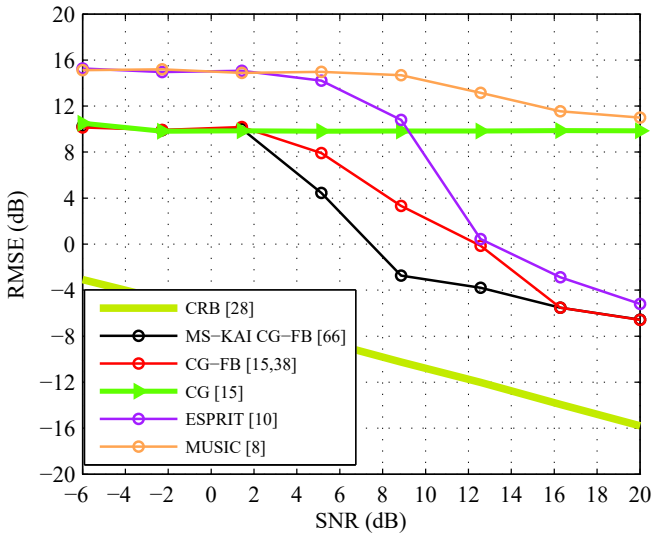

Figure 4.7: *RMSE and the square root of CRB in dB versus SNR with* $P = 2$, $M = 12$, $N = 70$, $L = 150$ *runs*, $\xi(\theta) = 2.0^o$.

5

Multi-Step Knowledge-Aided Iterative MUSIC for Nested Sensor Arrays

5.1

Introduction

In the previous chapters of this thesis, we have presented algorithms based on a ULA-based system model, whose major applications such as DOA estimation and beamforming suffer from a key limitation: the number of sources that can be resolved with an N element ULA using conventional subspace based methods like MUSIC [1], [8] is *N-1*. Over the years, the question of detecting more sources than sensors has been dealt with by different approaches. In [67,68], the use of minimum redundancy arrays (MRA) [69] and the construction of an enlarged covariance matrix for achieving higher degrees of freedom (DOF) has not been successful. In [70, 71], an approach to convert the enlarged matrix into an appropriate positive definite Toeplitz matrix has been proposed and relies on MRA. Despite those efforts to achieve more DOF to process more sources than sensors, there is no closed form expression for the array geometry. Moreover, such arrays demand hard designs which are limited to computer simulations or complex algorithms for locating the sensors [72–76]. In [77–79], it was shown that an approach [80] using fourth-order cumulants succeeded in increasing the DOF, however it is limited to non-Gaussian sources. In [81,82], by using the Khatri-Rao (KR) product and the hypothesis of quasi-stationary sources, which finds applications in microphone array processing of speech [83], one can recognize *2N-1* sources through a N element ULA without the need for high-order statistics. In [84], the increase of the DOF results from building a virtual array making use of a MIMO radar. Since the creation of that array relies on active sensing, the method is not suitable for passive sensing. In [85, 86], by exploring the class of non-uniform arrays, a structure called nested has been introduced, which is formed by combining two or more ULAs, to obtain a difference co-array. This structure can provide an increase of DOF and, therefore, can resolve more sources than the number of sensors. In a

subsequent work [87], linear nested arrays were employed to estimate DOAs of distributed sources. Moreover, in [88] robust beamforming for these arrays based on interference-plus-noise reconstruction and steering vector estimation has been developed. The studies in [85–88] focus on scenarios with multiple but not necessarily closely spaced sources in order to assess their performance. To this end, their signal models assume that the sources are uncorrelated. However, the required vectorization of the initial covariance matrix resulting from uncorrelated sources already leads to an equivalent source signal vector whose powers of their sources behave like fully coherent ones. For this reason, these methods require spatial smoothing.

In Chapter 3, we have presented Multi-Step KAI-ESPRIT [51, 52], and in Chapter 4, the Krylov subspace based Multi-Step KAI-Conjugate Gradient [65, 66]. Both perform refinements of the covariance matrix estimates via cancellation [44, 62] of their undesirable terms. However, neither MUSIC-type algorithms nor non-uniform arrays have yet been considered with the MS-KAI approach.

In this chapter, in order to satisfy such needs, we present a MUSIC-type algorithm for DOA estimation using nested arrays, denoted multi-step knowledge-aided iterative MUSIC method [89] (MS-KAI-MUSIC). The basic idea of MS-KAI-MUSIC is to exploit prior knowledge about the signals and the mathematical structure of the spatially smoothed [38,90,91] covariance matrix of the data of nested arrays, which are obtained online. An iterative procedure to perform cancellation of undesirable terms of the estimate of the spatially smoothed covariance matrix is then developed. Unlike existing knowledge-aided methods applied to ULAs, which exploit available known DOAs to improve the estimation of the covariance matrix of the input data, MS-KAI-MUSIC exploits knowledge of the structure of the spatially smoothed covariance matrix obtained from part of a difference co-array of a two-level nested array and the gradual incorporation of prior knowledge, which is obtained on line.

In this Chapter, after describing the nested arrays system model and method, we present the MS-KAI-MUSIC technique. We also discuss its computational complexity and present simulations, which show its performance.

Notation: the superscript H denotes the Hermitian transposition, $\mathbb{E}[\cdot]$ expresses the expectation operator. $\boldsymbol{I}$ stands for the identity matrix and $\odot$ means the Khatri-Rao product.

5.2

System Model

Let us consider a two-level nested sensor array composed of M sensors, which is a concatenation of two ULAs. The inner ULA has M_1 sensors with intersensor spacing d_1 and the outer has M_2 sensors with intersensor spacing $d_2 = (M_1 + 1)\, d_1$. Specifically, it consists of a linear array with sensors positions obtained by the union of the sets $I_{nner} = \{md_1 \mid \ m = 1, 2, \ldots, M_1\}$ and $O_{uter} = \{n\,\{M_1 + 1\}\, d_1 \mid \ n = 1, 2, \ldots, M_2\}$. Fig. 5.1 illustrates a two-level nested array composed of 6 sensors. Assuming P uncorrelated narrowband

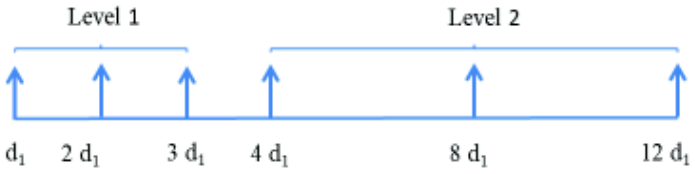

Figure 5.1: A two level nested array with 3 sensors at each level.

signals from far-field sources at directions $\{\theta_p, p = 1, 2, \ldots, P\}$ impinging on this array, the ith data snapshot of the M-dimensional array output vector can be modeled as

$$\boldsymbol{y}(i) = \boldsymbol{F}\, \boldsymbol{s}(i) + \boldsymbol{n}(i), \qquad i = 1, 2, \ldots, N, \tag{5-1}$$

where $\boldsymbol{y}(i) = [y_1(i), y_2(i), \ldots, y_M(i)]^T$ is the received signal vector at the snapshot i, $\boldsymbol{s}(i) = [s_1(i), s_2(i), \ldots, s_P(i)]^T$ is the source signal vector and $s_p(i) \sim N^C\left(0, \sigma_p^2\right)$. Additionally, we assume that $\boldsymbol{n}(i) = [n_1(i), n_2(i), \ldots, n_M(i)]^T$ is the white Gaussian noise vector with power σ_n^2 and that its components and the source vector ones are uncorrelated to each other. We also consider that $\boldsymbol{f}(\theta_p) = \left\{e^{-j2\pi \frac{d_1}{\lambda_c} r_n \sin\theta_p} \mid n = 1, 2, \ldots, M\right\}$ denotes the steering vector of the pth signal, where λ_c stands for the carrier wavelength and

$$\begin{aligned}\{r_n \mid n = 1, 2, \ldots, M\} = \{&0, 1, \ldots, M_1 - 1, M_1,\\ &2\,(M_1 + 1) - 1, \ldots,\\ &M_2\,(M_1 + 1) - 1\}\end{aligned} \tag{5-2}$$

is a vector that contains the location of the sensors. Next, the array manifold containing the steering vectors of the signals can be formed as

$$\boldsymbol{F}(\Theta) = [\boldsymbol{f}(\theta_1), \boldsymbol{f}(\theta_2), \ldots, \boldsymbol{f}(\theta_P)] \tag{5-3}$$

By averaging the N collected snapshots over time, we can express the sample covariance matrix as

$$\begin{aligned}\hat{\boldsymbol{R}}_{\mathbf{1}} &= \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}\left(i\right)\boldsymbol{y}^{H}\left(i\right)\\ &\approx \mathbb{E}\left[\boldsymbol{y}\left(i\right)\boldsymbol{y}^{H}\left(i\right)\right]\\ &= \boldsymbol{F}\,\boldsymbol{R}_{s}\boldsymbol{F}^{H}+\sigma_{n}^{2}\boldsymbol{I}\\ &= \boldsymbol{F}\begin{bmatrix}\sigma_{1}^{2} & & & \\ & \sigma_{2}^{2} & & \\ & & \ddots & \\ & & & \sigma_{P}^{2}\end{bmatrix}\boldsymbol{F}^{H}+\sigma_{n}^{2}\boldsymbol{I}\end{aligned} \tag{5-4}$$

Next, following [81], by the vectorization of $\hat{\boldsymbol{R}}_{\mathbf{1}}$ (5-4), one can obtain a long vector $\boldsymbol{z}_1$, as shown bellow:

$$\begin{aligned}\boldsymbol{z}_1 &= \mathrm{vec}\left(\hat{\boldsymbol{R}}_{\mathbf{1}}\right)\\ &= \mathrm{vec}\left[\sum_{i=1}^{N}\sigma_{i}^{2}\left(\boldsymbol{f}(\theta_1)\boldsymbol{f}(\theta_1^{H})\right)\right]+\sigma_{n}^{2}+\overrightarrow{\mathbf{1}}_{n}\\ &= \left(\boldsymbol{F}^{*}\odot\boldsymbol{F}\right)\,\boldsymbol{p}+\overrightarrow{\mathbf{1}}_{n},\end{aligned} \tag{5-5}$$

where

$$\boldsymbol{p}=\left[\sigma_{1}^{2},\sigma_{2}^{2},\ldots,\sigma_{P}^{2}\right]^{T}, \tag{5-6}$$

$$\overrightarrow{\mathbf{1}}_{n}=\left[\boldsymbol{e}_{1}^{T},\boldsymbol{e}_{2}^{T}\cdots\boldsymbol{e}_{M}^{T}\right]^{T} \tag{5-7}$$

and $\odot$ means the Khatri-Rao product.

Since in the long vector (5-5) some elements appear more than once, one can remove duplicated rows and sort them so that the *ith* row corresponds to the sensor located at $\left(-\bar{M}+i\right)d_1$, where $\bar{M}=(M^2/4+M/2)$. Then, we can obtain the following new vector:

$$\boldsymbol{z}=\boldsymbol{G}\,\boldsymbol{p}+\sigma_{n}^{2}\boldsymbol{e}, \tag{5-8}$$

where

$$\boldsymbol{G}(\Theta)=[\boldsymbol{g}(\theta_1),\boldsymbol{g}(\theta_2),\ldots,\boldsymbol{g}(\theta_P)], \tag{5-9}$$

in which

$$\boldsymbol{g}(\theta_p) = \Big[e^{-j2\pi\frac{d_1}{\lambda_c}\left(-\bar{M}+1\right)\sin\theta_p}, e^{-j2\pi\frac{d_1}{\lambda_c}\left(-\bar{M}+2\right)\sin\theta_p}, \\ \ldots, e^{-j2\pi\frac{d_1}{\lambda_c}\left(\bar{M}-2\right)\sin\theta_p}, e^{-j2\pi\frac{d_1}{\lambda_c}\left(\bar{M}-1\right)\sin\theta_p}\Big]^T, \tag{5-10}$$

$$\boldsymbol{e} \in \Re^{\left(2\bar{M}-1\right)\times 1} \tag{5-11}$$

is a vector of all zeros, except for a *1* at the center position and $\boldsymbol{p}$ has already been given by (5-6).

By comparing (5-8) with (5-1), we can notice that $\boldsymbol{z}$ in (5-8) behaves like the signal received by a longer difference coarray [92], whose sensors locations can be determined by the distinct values in the set $\{r_i - r_j \mid 1 \leq i, j \leq M\}$. The equivalent source signal vector $\boldsymbol{p}$ (5-6) consists of powers σ_p^2 of the actual sources and thus they behave like fully coherent sources [85,86]. This, combined with the fact that the difference coarray is a filled ULA, motivates to apply spatial smoothing to $\boldsymbol{z}$ (5-8) to obtain a full rank covariance matrix $\widetilde{R}$ as follows:

$$\begin{aligned} \widetilde{\boldsymbol{R}} = & \quad \frac{1}{M^2/4+M/2} \sum_{i=1}^{M^2/4+M/2} \boldsymbol{z}_i \boldsymbol{z}_i^H \\ = & \quad \frac{1}{M^2/4+M/2} \left(\boldsymbol{G}_1 \boldsymbol{R}_s \boldsymbol{G}_1^H + \sigma_n^2 \boldsymbol{I}\right)^2, \end{aligned} \tag{5-12}$$

where $\boldsymbol{z}_i$ corresponds to the

$$\left(M^2/4 + M/2 - i + 1\right) th \text{ to } \left(\left(M^2 - 2\right)/2 + M - i + 1\right) th \tag{5-13}$$

rows of $\boldsymbol{z}$ and $\boldsymbol{G}_1$ is a manifold array composed of the last $\bar{M}$ rows of $\boldsymbol{G}$. It can be shown [86] that the smoothed covariance matrix $\widetilde{\boldsymbol{R}}$ (5-12) can be expressed as $\widetilde{\boldsymbol{R}} = \hat{\boldsymbol{R}}^2$, where $\hat{\boldsymbol{R}}$ has the same form as the covariance received by a longer ULA composed of $M^2/4 + M/2$ sensors. Since $\hat{\boldsymbol{R}}$ and $\widetilde{\boldsymbol{R}}$ share the same set of eigenvectors and the eigenvalues of $\hat{\boldsymbol{R}}$ are the square roots of those of $\widetilde{\boldsymbol{R}}$, by eigendecomposition of $\widetilde{\boldsymbol{R}}$, we can find the eigenvectors corresponding to the smallest $(M^2/4 + M/2) - P$ eigenvalues of $\hat{\boldsymbol{R}}$. Due to the previously mentioned reasons and also for being PSD by construction, which results from the sum of vector outer products, the spatially smoothed matrix $\widetilde{\boldsymbol{R}}$ can be used as the basis for the proposed MS-KAI-MUSIC algorithm.

5.3

Proposed MS-KAI-MUSIC algorithm

The idea behind the MS-KAI-Nested-MUSIC algorithm is to expand the estimated spatially smoothed covariance matrix $\widetilde{\boldsymbol{R}}$ (5-12) as if it were generated by i data snapshots of $L = (M^2/4 + M/2)$ -dimensional array output vectors, where, as explained in Section 5.2, M is the number of physical sensors of the nested array. That is to say that we can employ the estimated spatially smoothed covariance matrix $\widetilde{\boldsymbol{R}}$ as if it were the estimate provided by the sample average formula. Since, as mentioned before, the resulting smoothed covariance matrix is generated by part of a coarray, which is a filled ULA, from now on, our method will make use of the ULA model.

Therefore, after setting $\widetilde{\boldsymbol{R}}$ in (5-12) equal to $\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}(i)\boldsymbol{x}^H(i)$ (2-52), we can expand (5-12) as follows:

$$\begin{aligned}\widetilde{\boldsymbol{R}} &= \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{A}\, s(i) + \boldsymbol{n}(i))\,(\boldsymbol{A}\, s(i) + \boldsymbol{n}(i))^H \\ &= \boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{n}^H(i) \\ &+ \underbrace{\boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{n}^H(i)\right\} + \left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H}_{\text{"undesirable by-products"}}\end{aligned} \tag{5-14}$$

In a similar way to MS-KAI-ESPRIT and MS-KAI-CG, in Chapters 3 and 4, the first two terms of $\widetilde{\boldsymbol{R}}$ in (5-14) can be considered as estimates of the two summands of $\boldsymbol{R} = \boldsymbol{A}\,\boldsymbol{R}_{ss}\boldsymbol{A}^H + \sigma_n^2\boldsymbol{I}_L$, which represent the signal and the orthogonal subspaces, respectively. The last two terms in (5-14) are undesirable by-products, which can be seen as estimates for the correlation between the signal and the noise vectors. The system model under study is based on noise vectors which are zero-mean and statistically independent of the signal vectors. As a consequence, for a large enough number of samples N, the last two terms expressed in (5-14) tend to zero. Nevertheless, in practice the number of available samples can be limited. In such situations, the last two terms in (5-14) may have significant values, which causes the deviation of the estimates of the signal and the noise subspaces from the true signal and noise ones. The key point of the proposed MS-KAI-MUSIC algorithm is to modify the smoothed covariance matrix estimate at each iteration by gradually incorporating the knowledge provided by the updated Vandermonde matrices which progressively incorporate the newer estimates from the preceding iteration. Based on these

updated Vandermonde matrices, refined estimates of the projection matrices of the signal and noise subspaces are calculated. These estimates of projection matrices associated with the initial smoothed covariance matrix estimate and the reliability factor employed to reduce its by-products allow to choose the set of estimates that has the minimum value of the stochastic maximum likelihood objective function (SMLOF), i.e., the highest likelihood of being the set of the true DOAs. The modified smoothed covariance matrix estimate is computed by deriving a scaled version of the undesirable terms from $\widetilde{\boldsymbol{R}}$, which are pointed out in (5-14).

MS-KAI-MUSIC starts by computing the spatially smoothed covariance matrix estimate (5-12). Next, the DOAs are estimated using the original MUSIC [8] algorithm. The superscript $(\cdot)^{(1)}$ refers to the estimation task performed in the 1^{st} step. Now, a procedure composed of $n = 1 : I$ iterations starts by forming the Vandermonde matrix using the DOA estimates. Then, the amplitudes of the sources are estimated such that the squared norm of the differences between the observation vector and the vector containing estimates and the available known DOAs is minimized. This problem can be formulated as

$$\hat{\boldsymbol{s}}(i) = \arg\min_{\boldsymbol{s}} \parallel \boldsymbol{x}(i) - \hat{\boldsymbol{A}}\mathbf{s} \parallel_2^2 . \tag{5-15}$$

The minimization of (5-15) is achieved using the least squares technique and the solution is described by

$$\hat{\boldsymbol{s}}(i) = (\hat{\mathbf{A}}^H \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}} \boldsymbol{x}(i). \tag{5-16}$$

The noise component is then estimated as the difference between the estimated signal and the observations made by the array, as given by

$$\hat{\boldsymbol{n}}(i) = \boldsymbol{x}(i) - \hat{\boldsymbol{A}} \hat{\boldsymbol{s}}(i). \tag{5-17}$$

After estimating the signal and the noise vectors, the third term in (4-20) can

be computed as

$$\begin{aligned}\boldsymbol{V} &\triangleq \hat{\boldsymbol{A}}\left\{\frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{s}}(i)\hat{\mathbf{n}}^{H}(i)\right\}\\ &= \hat{\boldsymbol{A}}\left\{\frac{1}{N}\sum_{i=1}^{N}(\hat{\mathbf{A}}^{H}\,\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}^{H}\boldsymbol{x}(i)\right.\\ &\left.\times(\boldsymbol{x}^{H}(i)-\boldsymbol{x}^{H}(i)\hat{\mathbf{A}}(\hat{\mathbf{A}}^{H}\hat{\mathbf{A}})^{-1}\,\hat{\mathbf{A}}^{H})\right\}\\ &= \hat{\mathbf{Q}}_{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}(i)\boldsymbol{x}^{H}(i)\,\left(\mathbf{I}_{M}\,-\,\hat{\mathbf{Q}}_{A}\right)\right\}\\ &= \hat{\mathbf{Q}}_{A}\,\widetilde{\mathbf{R}}\,\hat{\mathbf{Q}}_{A}^{\perp},\end{aligned} \tag{5-18}$$

where

$$\hat{\mathbf{Q}}_{A} \triangleq \hat{\mathbf{A}}\,(\hat{\mathbf{A}}^{H}\,\hat{\mathbf{A}})^{-1}\,\hat{\mathbf{A}}^{H} \tag{5-19}$$

is an estimate of the projection matrix of the signal subspace, and

$$\hat{\mathbf{Q}}_{A}^{\perp} \triangleq \mathbf{I}_{L}\,-\,\hat{\mathbf{Q}}_{A} \tag{5-20}$$

is an estimate of the projection matrix of the noise subspace.

Next, as part of the process of $n = 1 : I$ iterations, the modified data covariance matrix $\widetilde{\mathbf{R}}^{(n+1)}$ is calculated by computing a scaled version of the estimated terms from the initial smoothed covariance matrix as given

$$\widetilde{\mathbf{R}}^{(n+1)} = \widetilde{\mathbf{R}}\,-\,\mu\,(\mathbf{V}^{(n)}\,+\,\mathbf{V}^{(n)H}), \tag{5-21}$$

where the superscript $(\cdot)^{(n)}$ refers to the n^{th} iteration performed. The scaling or reliability factor μ increases from 0 to 1 incrementally, resulting in modified smoothed covariance matrix estimates. Each of them gives origin to new DOAs estimates also denoted by the superscript $(\cdot)^{(n+1)}$ using the MUSIC algorithm. Here, the rank P is assumed to be known, which is an assumption frequently found in the literature. Alternatively, the rank P could be estimated by model-order selection schemes [53] such as Akaike's Information Theoretic Criterion (AIC) [54] and the Minimum Descriptive Length (MDL) Criterion [55].

Then, a new Vandermonde matrix $\hat{\mathbf{B}}^{(n+1)}$ is formed by the steering vectors of those new DOAs estimates. By using $\hat{\mathbf{B}}^{(n+1)}$, it is possible to compute the newer estimates of the projection matrices of the signal $\hat{\mathbf{Q}}_{B}^{(n+1)}$ and the noise $\hat{\mathbf{Q}}_{B}^{(n+1)\perp}$ subspaces.

Afterwards, employing the newer estimates of the projection matrices, the initial smoothed covariance matrix estimate, $\widetilde{\boldsymbol{R}}$, the number of its corresponding sensors and the number of sources, the stochastic maximum likelihood objective function $U^{(n+1)}(\mu)$ [45] is computed for each value of μ at the n^{th} iteration,

as follows:

$$U^{(n+1)}(\mu) = \ln \det(\cdot), \tag{5-22}$$

where

$$(\cdot) = \left(\hat{\mathbf{Q}}_B^{(n+1)} \widetilde{\mathbf{R}} \, \hat{\mathbf{Q}}_B^{(n+1)} + \frac{\mathrm{Trace}\{\hat{\mathbf{Q}}_B^{\perp\,(n+1)} \widetilde{\mathbf{R}}\}}{\mathrm{L} - \mathrm{P}} \, \hat{\mathbf{Q}}_B^{(n+1)\perp} \right)$$

The preceding computation selects the set of unavailable DOA estimates that have a higher likelihood at each iteration. Then, the set of estimated DOAs corresponding to the optimum value of μ that minimizes (5-22) also at each n^{th} iteration is determined. Lastly, the output of the proposed MS-KAI-MUSIC algorithm is formed by the set of the estimates obtained at the I^{th} iteration, as described in Table 5.1.

5.4 Computational Complexity Analysis

In this section, we evaluate the approximate computational cost of the proposed MS-KAI-MUSIC algorithm in terms of multiplications and additions. For this purpose, we make use of Table 5.2, where $\tau = \frac{1}{\iota} + 1$. The increment ι is defined in Table 5.1.

From Table 5.2, it can be seen that assuming the specific configuration used in the simulations 5.5, MS-KAI-Nested-MUSIC shows a roughly similar computational burden in terms of multiplications and also of additions with $\mathcal{O}\left\{ I\tau \left[\frac{180}{\Delta} \left(\frac{M^2}{4} + \frac{M}{2} \right)^2 \right] + \left(\frac{M^2}{4} + \frac{M}{2} \right) 8N^2 \right\}$, where τ is typically an integer that ranges from 2 to 20, Δ stands for the search step and I is the number of iterations at the 2^{nd} step. The relatively high costs come from the two nested loops for computing $I \times \tau$ times two subprocesses at its second step. These nested loops, from which the last is the most significant, concentrate most of the required operations. For this reason it is responsible for most of the cost of MS-KAI-MUSIC.

5.5 Simulations

In this section, we examine the performance of the proposed MS-KAI-MUSIC algorithm in terms of probability of resolution (PR) and compare it to the corresponding performances of MUSIC for nested arrays (Nested-MUSIC) [86] and of the original MUSIC [8] for ULAs. We focus on the specific case of closely-

Table 5.1: MS-KAI-MUSIC algorithm

Inputs:
$M_1,\ M_2,\ d_1,\ \lambda,\ N,\ P$
Received vectors $\boldsymbol{y}(1),\ \boldsymbol{y}(2),\cdots,\ \boldsymbol{y}(N)$
Outputs:
Estimates $\hat{\theta}_1^{(n+1)}(\mu\ opt),\ \hat{\theta}_2^{(n+1)}(\mu\ opt),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu\ opt)$
First step:
$\{\hat{\theta}_1^{(1)},\ \hat{\theta}_2^{(1)},\cdots,\hat{\theta}_P^{(1)}\}\ \underleftarrow{MUSIC}\ (\widetilde{\boldsymbol{R}}, P, d, \lambda)$
$\hat{\boldsymbol{A}}^{(1)} = \left[\boldsymbol{a}(\hat{\theta}_1^{(1)}), \boldsymbol{a}(\hat{\theta}_2^{(1)}), \cdots, \boldsymbol{a}(\hat{\theta}_P^{(1)})\right]$
Second step:
for $n = 1 : I$
$\hat{\boldsymbol{Q}}_A^{(n)} = \hat{\boldsymbol{A}}^{(n)}\ (\hat{\boldsymbol{A}}^{(n)H}\ \hat{\boldsymbol{A}}^{(n)})^{-1}\ \hat{\boldsymbol{A}}^{(n)H}$
$\hat{\boldsymbol{Q}}_A^{(n)\perp} = \boldsymbol{I}_L\ -\ \hat{\boldsymbol{Q}}_A^{(n)}$
$\boldsymbol{V}^{(n)} = \hat{\boldsymbol{Q}}_A^{(n)}\ \widetilde{\boldsymbol{R}}\ \hat{\boldsymbol{Q}}_A^{(n)\perp}$
for $\boldsymbol{\mu} = 0 : \iota\ : 1$
$\widetilde{\boldsymbol{R}}^{(n+1)} = \widetilde{\boldsymbol{R}}\ -\ \mu\ (\boldsymbol{V}^{(n)}\ +\ \boldsymbol{V}^{(n)H})$
$\{\hat{\theta}_1^{(n+1)},\ \hat{\theta}_2^{(n+1)},\cdots,\hat{\theta}_P^{(n+1)}\}\ \underleftarrow{MUSIC}\ (\widetilde{\boldsymbol{R}}^{(n+1)},\ P, d, \lambda)$
$\hat{\boldsymbol{B}}^{(n+1)} = \left[\boldsymbol{a}(\hat{\theta}_1^{(n+1)}), \boldsymbol{a}(\hat{\theta}_2^{(n+1)}), \cdots, \boldsymbol{a}(\hat{\theta}_P^{(n+1)})\right]$
$\hat{\boldsymbol{Q}}_B^{(n+1)} = \hat{\boldsymbol{B}}^{(n+1)}\ (\hat{\boldsymbol{B}}^{(n+1)H}\ \hat{\boldsymbol{B}}^{(n+1)})^{-1}\ \hat{\boldsymbol{B}}^{(n+1)H}$
$\hat{\boldsymbol{Q}}_B^{(n+1)\perp} = \boldsymbol{I}_L\ -\ \hat{\boldsymbol{Q}}_B^{(n+1)}$
$U^{(n+1)}(\mu) = \ln\det\left(\hat{\boldsymbol{Q}}_B^{(n+1)}\ \widetilde{\boldsymbol{R}}\ \hat{\boldsymbol{Q}}_B^{(n+1)} + \dfrac{\mathrm{Trace}\{\hat{\boldsymbol{Q}}_B^{\perp\,(n+1)}\ \widetilde{\boldsymbol{R}}\}}{L-P}\ \hat{\boldsymbol{Q}}_B^{(n+1)\perp}\right)$
$\mu_{\mathrm{o}}^{(n+1)} = \arg\min\ U^{(n+1)}(\mu)$
$\mathrm{DOAs}^{(n+1)} = \{\hat{\theta}_1^{(n+1)}(\mu_o),\ \hat{\theta}_2^{(n+1)}(\mu_o),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu_o)\}$
if $n <= P$
$\hat{\boldsymbol{A}}^{(n+1)} = \left\{\boldsymbol{a}(\hat{\theta}_{\{1,\cdots,n\}}^{(n+1)}(\mu_o))\right\} \cup \left\{\boldsymbol{a}(\hat{\theta}_{\{1,\cdots,P\}-\{1,\cdots,n\}}^{(1)})\right\}$
else
$\hat{\boldsymbol{A}}^{(n+1)} = \left[\boldsymbol{a}(\hat{\theta}_1^{(n+1)}(\mu_o)), \boldsymbol{a}(\hat{\theta}_2^{(n+1)}(\mu_o)), \cdots, \boldsymbol{a}(\hat{\theta}_P^{(n+1)}(\mu_o))\right]$
end if
end for
end for

spaced sources. We employ $M = 8$ sensors in the algorithms based on two-level nested array and. In the original MUSIC, we use a ULA with $M = 20$ sensors, which is also the same number of sensors $(M^2/4 + M/2)$ of the filled ULA obtained from part of the difference coarray, which is the effective number of sensors employed in the MUSIC for nested arrays and MS-KAI-MUSIC algorithms. The choice of the number of the sensors of each algorithm is a strategy to assess the employment of sensor arrays with a reduced number of sensors. We assume the shortest inter-element spacing $d_1 = \frac{\lambda_c}{2}$ and also

Table 5.2: MS-KAI-MUSIC algorithm

Multiplications

$$\approx I\tau\left\{\frac{180}{\Delta}[\left(\frac{M^2}{4}+\frac{M}{2}\right)^2+\left(\frac{M^2}{4}+\frac{M}{2}\right)(2-P)-P]\right.$$

$$+\left(\frac{M^2}{4}+\frac{M}{2}\right)8N^2+\frac{10}{3}\left(\frac{M^2}{4}+\frac{M}{2}\right)^3+\left(\frac{M^2}{4}+\frac{M}{2}\right)^2(P+2)$$

$$\left.+\left(\frac{M^2}{4}+\frac{M}{2}\right)(P^2+2P)+\frac{P^3}{2}+\frac{3P^2}{2}\right\}$$

Additions

$$\approx I\tau\left\{\frac{180}{\Delta}[\left(\frac{M^2}{4}+\frac{M}{2}\right)^2-\left(\frac{M^2}{4}+\frac{M}{2}\right)(P-1)]\right.$$

$$+\left(\frac{M^2}{4}+\frac{M}{2}\right)8N^2+\frac{10}{3}\left(\frac{M^2}{4}+\frac{M}{2}\right)^3+\left(\frac{M^2}{4}+\frac{M}{2}\right)^2(P-1)$$

$$\left.+\left(\frac{M^2}{4}+\frac{M}{2}\right)\left(\frac{3P^2}{2}+\frac{5P}{2}-1\right)-P^2-\frac{P}{2}\right\}$$

---

that there are two uncorrelated complex Gaussian signals with equal power impinging on the arrays. The closely-spaced sources are separated by 2^o, at $(15^o, 17^o)$. The first figure makes use of $N = 150$ snapshots and $L_r = 250$ trials, whereas the latter employs $3.33dB$ and $L_r = 500$ trials. In both cases, we have set the search step to $\Delta = 0.1^o$.

In Fig. 5.2, we show PR versus SNR. We take into account the criterion [29], in which two sources with DOAs θ_1 and θ_2 are said to be resolved if their respective estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are such that both $\left|\hat{\theta}_1-\theta_1\right|$ and $\left|\hat{\theta}_2-\theta_2\right|$ are less than $|\theta_1-\theta_2|/2$. It can be seen the superior performance of the proposed MS-KAI-MUSIC in the range $(-10 \quad 7)\,dB$. From this point on, all considered algorithms provide similar performance. The gap between the proposed MS-KAI-MUSIC and MUSIC with nested arrays [86] shows a significant improvement achieved in terms of PR. It can be noticed a bigger gap between the proposed MS-KAI-MUSIC and the original MUSIC [8], whose number of physical sensors is 2.5× the number of the physical sensors of the other two-level nested based algorithms under comparison, which means an important saving of sensors.

In Fig. 5.3, it is shown the RMSE in degrees versus SNR. The RMSE is defined as

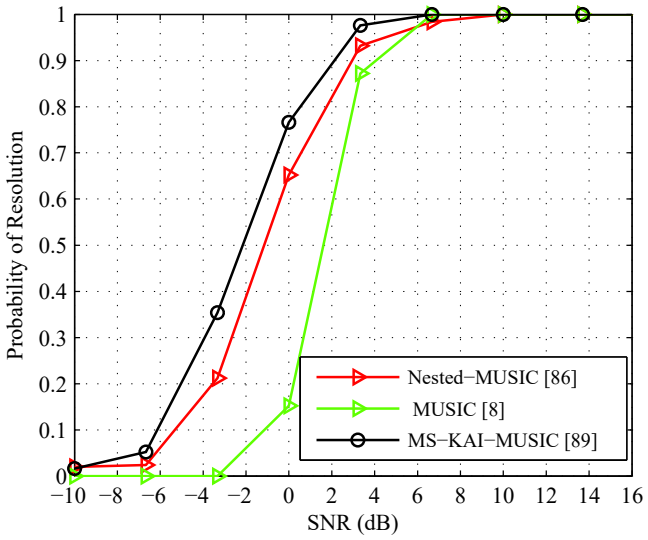

Figure 5.2: Probability of resolution versus SNR with $P = 2$, $M = 8$, $N = 150$, $L_r = 250$ runs

$$\text{RMSE} = \sqrt{\frac{1}{L_r\, P}\sum_{l=1}^{L_r}\sum_{p=1}^{P}(\theta_p - \hat{\theta}_p(l))^2}, \tag{5-23}$$

where L_r is the number of trials.

It can be noticed that the MS-KAI-MUSIC outperforms Nested-MUSIC, in the whole range under consideration. In the range $[-10\ -1.8)\,dB$, it is outperformed by conventional MUSIC, however, it can be noticed that the achieved level shows a clear trend to improvement in accuracy. From -1.8 to $6.7\,dB$ MS-KAI-MUSIC is superior to it. From $10\,dB$ on all algorithms have similar performance. As mentioned before, it must be highlighted that in this specific case MUSIC makes use of a ULA whose number of physical sensors is $2.5\times$ the number of the physical sensors of the other two-level nested based algorithms under comparison.

In Fig. 5.4, it is shown the influence of the number of snapshots on PR. For this purpose we have set the SNR at $3.33\,dB$ and employed 500 trials. From the curves, it can be noticed the superior performance of MS-KAI-MUSIC in the range of 25 to 250 snapshots. From this point on, all algorithms have similar performance.

In Fig.5.5, it is shown the influence of the number of snapshots on RMSE. In this case, we also set the SNR at 3.33 dB and employed 500 trials. It can be seen that the performance of the MS-KAI-MUSIC is superior to Nested-MUSIC. It can also be noticed that except for the range 25 to 50, in which the RMSE has high levels, the performance of MS-KAI-MUSIC is also superior to

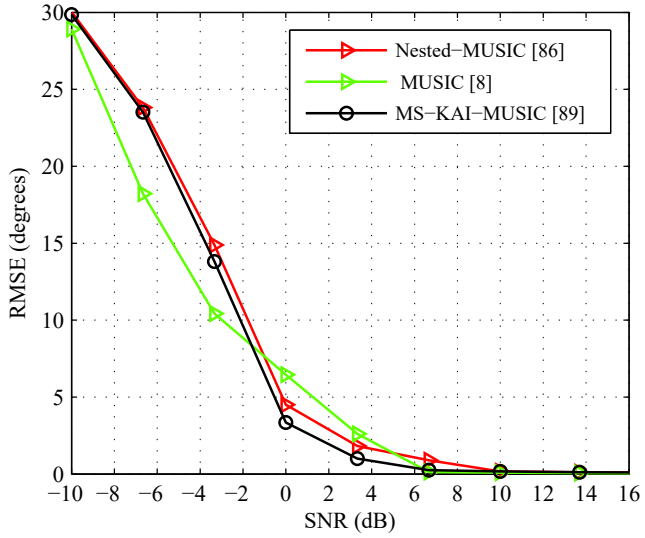

Figure 5.3: RMSE in degrees versus SNR with with $P = 2$, $M = 8$, $N = 150$, $L_r = 250$ runs.

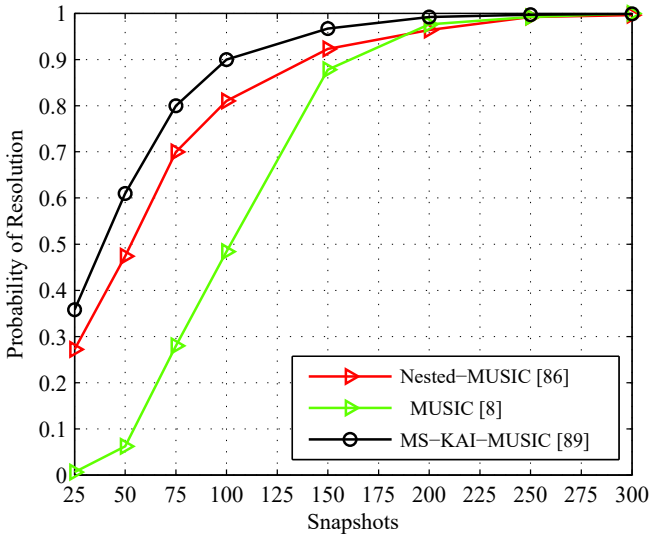

Figure 5.4: Probability of resolution versus SNR with $P = 2$, $M = 8$, $SNR = 3.33\,dB$, $L_r = 500$ runs.

the original MUSIC [8], whose number of physical sensors is 2.5× the number of the physical sensors of the other two-level nested based algorithms under comparison.

Finally, from Figures 5.4 and 5.5, respectively, it can be noticed the significant saving of samples to achieve high probabilities of resolution and reduced RMSEs.

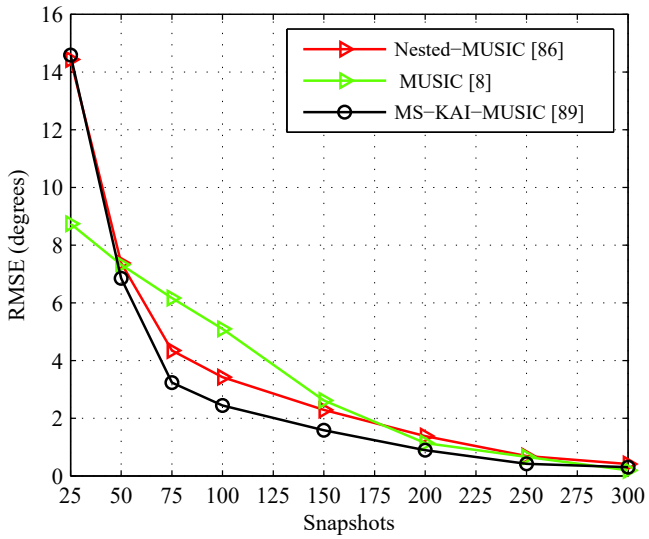

Figure 5.5: RMSE versus SNR with $P = 2$, $M = 8$, $SNR = 3.33\,dB$, $L_r = 500$ runs.

6

Conclusions

In this Chapter, conclusions of this thesis are presented and future directions for this research topic are discussed. It is organized as follows: Section 6.1 summarizes the work and points significant results and Section 6.2 present points to be worked in the future.

6.1

Summary of the Work

This thesis has been devoted to the investigation of high-resolution direction finding techniques exploiting prior knowledge. These methods can be classified into two categories according to the approaches for obtaining this knowledge. The approach previously available in the literature, termed KAv (Subsection 2.1.7.3) makes use of accessible known DOAs to form a rank-deficient known covariance, which can be optimally combined with the sample covariance matrix obtained from snapshots collected by a ULA, resulting in an enhanced covariance matrix estimate. This estimate, which, in this thesis, has been applied to CG (KAv-CG), can also be used as the basis for other several ULA-based algorithms. The KAv approach was the starting point for the approach developed in this thesis, termed KA (Subsection 2.1.7.3), which also employs the CG algorithm. Instead of using available known DOAs to compute the mentioned rank-deficient known covariance to be combined with the sample covariance matrix, the KA approach makes use of initial estimates, which can be considered online knowledge acquisition. The resulting covariance matrix is then processed by the CG algorithm to compute the final estimates. After checking the feasibility of on-line knowledge acquisition for enhancing the data covariance matrix, provided that the initial DOA estimates are sufficiently accurate, this concept has been extended to a new approach, described in Chapter 3 and expanded in Chapter 4: the iterative refinement of the data covariance matrix by progressive incorporation of knowledge on line. This approach, initially applied to algorithms that processed signals impinging on

ULAs, has been further extended to a non-ULA-based algorithm such as MS-KAI-MUSIC (Chapter 5), which makes use of a two-level nested array.

In Subsection 2.1.7.3, in a case study involving the CG algorithm applied to DOA estimation, we introduced the idea of replacing accessible DOAs with those acquired on-line to form a rank-deficient known covariance matrix, which combined with the sample covariance matrix in a mean squared error sense, yields an enhanced covariance matrix. The initial implementation of this approach has made use of a CG-based algorithm (KA-CG) composed of two stages: the former for obtaining the mentioned enhanced covariance matrix and the latter for processing it. The performances of KA-CG and subsequent versions of the KA approach based on MUSIC (KA-MUSIC) and ESPRIT (KA-ESPRIT) have been evaluated in terms of PR under a scenario composed of two signals generated by uncorrelated closely spaced sources impinging on a ULA and sufficient number of snapshots. For the purpose of comparisons, we have also plotted the KAv versions of the mentioned algorithms, i.e., instead of using estimates to form the known covariance, we have used available known DOAs to do that. The gaps between KAv-versions and their correspondent KA-versions have shown that most of the potential to be exploited is situated at middle low signal-to-noise ratios (SNR). Considering that each KAv-version can be considered an upper bound of its KA version the small gap between KA-CG and KAv-CG shows that the former already exploits its potential close to the effective optimal performance. Despite the larger existing gaps between the corresponding versions of MUSIC and ESPRIT, the specific result related to CG could be viewed as a preliminary indication of feasibility of replacing available known DOAs with estimates to form the known covariance, as suggested in the KA-approach.

In Chapter 3, extending the research about prior knowledge obtained on-line applied to DOA estimation, we have introduced the Multi-Step KAI approach applied to ESPRIT algorithm (MS-KAI-ESPRIT). This approach is based on the gradual incorporation of prior knowledge acquired on line, i.e, an increasing number of the obtained estimates, to iteratively refine the covariance matrix of the input data. It has also been presented an analysis of the mean squared error (MSE) of the data covariance matrix free of undesired terms (side effects) that results from low levels of SNR or modest number of snapshots. The analysis has shown that in the first iteration, the MSE of the data covariance matrix free of side effects is already less than or equal to the MSE of the original one. The heavy computational burden faced by MS-ESPRIT can be considered a cost to

be paid for the accuracy achieved, however this is a limitation to be addressed. The performance of MS-ESPRIT has been assessed in terms of PR and RMSE under scenarios composed of four signals generated by uncorrelated and highly correlated closely spaced sources impinging on a ULA and small data set. The comparisons with existing algorithms have evidenced the superior performance of the estimate-based MS-KAI-ESPRIT, which rivals and sometimes surpasses the supposed best performance of an algorithm that makes use of available known DOAs like TS-KAI-ESPRIT. Lastly, extra simulations have illustrated the influence of the iterations of MS-KAI-ESPRIT on its performance.

In Chapter 4, we have presented an MS-KAI approach combined with the CG algorithm, which is an expanded approach of that which has been applied to ESPRIT in Chapter 3. In this way, unlike the prior KA-ESPRIT approach MS-KAI-CG versions are no longer limited to P iterations like MS-KAI-ESPRIT one, where P is the number of source signals. Furthermore, we have provided a version equipped with forward-backward spatial smoothing, termed MS-KAI-CG-FB, which can deal with correlated signals. The CG-based algorithms are particularly effective for scenarios with very few source signals and closely-spaced angles of arrival. However, under a small number of samples or low levels of SNR, they suffer from lack of resolution, which can result from false intermediate peaks in the spectrum. In this case, simulations have shown that the MS-KAI-CG approach is particularly attractive and addresses this weakness in existing CG-based approaches by improving the quality of the covariance matrix estimates. Moreover, as expected, the MS-KAI approach applied to CG have show effectiveness in the finite sample region, by reducing the needed number of snapshots to achieve the same levels of accuracy. MS-KAI-CG has also shown the heaviest computational burden, resulting not only from the MS-KAI approach but also from its reliance on structure composed of peak searches.

Chapter 5 have focused on the development of the MS-KAI approach in a non-ULA-class-based method from which a two-level nested array is a representative. It is known that one of the main drawbacks of the spatial smoothing-based MUSIC algorithm applied to a two-level nested array a is that increasing degrees of freedom can be achieved at the cost of significant increase of the number of samples. For achieving this purpose, after obtaining an increased data covariance matrix by vectorization, the method removes redundancies and decorrelates intermediate sources which emerges during the process, resulting in a spatially smoothed matrix which is smaller than the

augmented covariance matrix but even so is greater than the initial one. Since this smoothed covariance matrix and its square, which is a longer ULA, share eigenstructure properties, the MS-KAI approach applied to a two-level nested array based on SS-MUSIC involves only ULA-based computations. The performance of MS-KAI-MUSIC has been evaluated in terms of PR and RMSE against SNR under scenarios composed of two signals generated by uncorrelated closely spaced sources impinging on a two-level nested array composed of eight sensors and sufficient small data set. The comparisons between MS-KAI-MUSIC and existing MUSIC-based algorithms like nested-MUSIC, which employs the same number of sensors, and the original MUSIC algorithm, which is based on a a ULA composed of 2.5× the number of sensors of MS-KAI-MUSIC, have evidenced the superiority of the first over the others. In what concerns RMSE, MS-KAI-MUSIC is superior to nested-MUSIC in all considered range, however in specific subranges it rivals the original MUSIC algorithm based on a a ULA composed of 2.5× its number of sensors. This fact can be viewed as a strategy to employ sensor arrays with a reduced number of sensors. Under the same mentioned scenario and a SNR set at 3.33 dB, the evaluation of the performance of the MS-KAI-MUSIC has shown significant saving of snapshots in terms of PR and RMSE. As expected from a MS-KAI approach, the computational complexity analysis of MS-KAI-MUSIC, has revealed a heavy burden, which is intensified by the increase of operations resulting from the greater dimensions of the matrices involved in the computations.

6.2

Future Work

For future work, efforts to enhance the KAI-approach and to extend it to other types of arrays should be considered along with the development of new algorithms to improve DOA estimation. Some suggestions based on this thesis are given below.

In Chapters 3, 4 and 5, the MS-KAI-approach combined with ESPRIT, CG and MUSIC algorithms, respectively, has faced a characteristic heavy computational burden. This huge number of operations results mainly from a factor termed τ related to the iterative reduction of the undesired by-products, which occurs under short data records or low levels of SNR and can result in considerable deviations of the sample covariance matrix from the true one. This factor is usually an integer that ranges from 2 to 20

and is proportional to the reciprocal of the step in which the reduction of the undesired terms is optimized. Therefore, research involving means of accelerating this optimization by reducing τ is a point to be considered.

In Chapter 5, it has been shown that it is possible to apply the MS-KAI approach to a nested-type non-uniform linear array. However, this approach can also be applied to other types of sensor array geometries. Thus, it is suggested to extend the research about the MS-KAI approach to other types of sensor arrays and geometries like co-prime arrays.

Up to this point, we have pointed out a drawback to be addressed and the feasibility of the application of MS-KAI approach to other geometries of sensor arrays as possible extensions of the present work. But can the mentioned approach also be applied to grouped (a set of) sources? This question can be considered in the form of distributed sources. It is indeed another point that deserves attention when deciding the expansion of the current work.

The suggestions for future works here presented are limited, however variations and new extensions are possible.

Bibliography

[1] TREES, H. V.. **Optimum Array Processing. Part IV of Detection, Estimation and Modulation Theory**. J.Wiley and Sons, 2002.

[2] MANOLAKIS, D.; INGLE, V. ; KOGON, S.. **Statistical and Adaptive Signal Processing: Spectral Estimation, Signal modeling, Adaptive Filtering and Array Processing. Part IV of Detection, Estimation and Modulation Theory**. Artech House, 2005.

[3] VEEN, B. V.; BUCKLEY, K. M.. **Beamforming: a versatile approach to spatial filtering**. IEEE ASSP Magazine, 5(2):4–24, Apr, 1988.

[4] FROST, O. L.. **An algorithm for linearly constrained adaptive array processing**. In: PROCEEDINGS OF THE IEEE, volumen 60, p. 926–935, Apr 1972.

[5] HAYKIN, S.. **Adaptive filter Theory**. Prentice Hall, 2003.

[6] KRIM, H.; VIBERG, M.. **Two decades of array signal processing research: parametric approach**. IEEE Signal Processing Magazine, 13(4):67–94, July, 1996.

[7] CAPON, J.. **High-resolution frequency-wavenumber spectrum analysis**. Proceedings of the IEEE, 57:1408–1418, Aug, 1969.

[8] SCHMIDT, R.. **Multiple emitter location and signal parameter estimation**. IEEE Transactions on Antennas and Propagation, 34:276–280, 2011.

[9] BARABELL, A.. **Performance of orthogonal matching pursuit for multiple measurement vectors with noise**. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, p. 336–339, April 1983.

[10] ROY, R.; KAILATH, T.. **Estimation of signal parameter via rotational invariance technique**. IEEE Transactions on Acoustics, Speech and Signal Processing, 37(7):984–995, 1989.

[11] HAARDT, M.; NOSSEK, J. A.. **Unitary ESPRIT: how to obtain increased estimation accuracy with a reduced computational burden**. IEEE Transactions on Signal Processing, 43(5):1232–1242, May, 1995.

[12] STEINWANDT, J.; DE LAMARE, R. C. ; HAARDT, M.. **Knowledge-aided direction finding based on Unitary ESPRIT**. In: 45TH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS, p. 613–617, 2011.

[13] GROVER, R.; PADOS, D. A. ; MEDLEY, M.. **Subspace direction finding with an auxiliary vector basis**. IEEE Transactions on Signal Processing, 55(2):758–763, Feb, 2007.

[14] ZISKIND, I.; WAX, M.. **Maximum likelihood localization of multiple sources by alternating projection**. IEEE Transactions on Acoustics, Speech and Signal Processing, 36(10):1553–1560, Oct, 1988.

[15] SEMIRA, H.; BELKACEMI, H. ; MARCOS, S.. **High-resolution source localization algorithm based on conjugate gradient**. In: EURASIP JOURNAL ADVANCES IN SIGNAL PROCESSING, p. 9 pp, 2007.

[16] SEMIRA, H.; BELKACEMI, H. ; DOGHMANE, N.. **A novel conjugate gradient-based source localization algorithm**. In: 9TH INTERNATIONAL SYMPOSIUM ON SIGNAL PROCESSING AND ITS APPLICATIONS, p. 1–4, 2007.

[17] KAY, S.. **Fundamentals of Statistical Signal Processing, Estimation Theory**. Prentice Hall, 1993.

[18] STOICA, P.; NEHORAI, A.. **Music, maximum likelihood, and cramer-rao bound**. IEEE Transactions on Acoustics, Speech and Signal Processing, 37(5):720–741, May,1989.

[19] GOLUB, G.; LOAN, C.. **Matrix Computations**. Prentice Hall Inc., 2002.

[20] WANG, L.; R.C.LAMARE ; HAARDT, M.. **Direction finding algorithms based on joint-iterative subspace optimization**. IEEE Transactions on Aerospace and Electronic Systems, 50(4):2541–2553, Oct, 2014.

[21] MELVIN, W.; GUERCI, J.. **Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors**. IEEE Transactions on Aerospace and Electronic Systems, 42(3):983–996, 2006.

[22] BERGIN, J.; TEIXEIRA, C.; TECHAU, P. M. ; GUERCI, J. R.. **Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors**. IEEE Transactions on Aerospace and Electronic Systems, 42(3):997–1009, 2006.

[23] ZHU, X.; LI, J. ; STOICA, P.. **Knowledge-aided adaptive beamforming**. IET Signal Processing, 2(4):335–345, 2008.

[24] FA, R.; LAMARE, R. C. ; NASCIMENTO, V. H.. **Knowledge-aided STAP algorithm using convex combination of inverse covariance matrices for heterogeneous clutter**. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, p. 2742–2745, March 2010.

[25] FA, R.; LAMARE, R. C.. **Knowledge-aided reduced-rank STAP for MIMO radar based on joint iterative constrained optimization of adaptive filters with multiple constraints**. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, p. 2762–2765, March 2010.

[27] SEMIRA, H.; BELKACEMI, H. ; MARCOS, S.. **A novel conjugate gradient-based source localization algorithm**. In: 4TH INTERNATIONAL CONFERENCE SCIENCES OF ELECTRONIC TECHNOLOGIES OF INFORMATION AND TELECOMMUNICATIONS, p. 6 pp, March 2007.

[28] STOICA, P.; LI, J.; ZHU, X. ; GUERCI, J.. **MUSIC, maximum Likelihood, and Cramer-Rao Bound**. IEEE Transactions on Signal Processing, 56(6):2598–2602, 2008.

[29] STOICA, P.; GERSHMAN, A. B.. **Maximum-likelihood DOA estimation by data-supported grid search**. IEEE Signal Processing Letters, 6(10):273–275, 1999.

[30] PINTO, S.; LAMARE, R. C.. **Knowledge-aided parameter estimation based on conjugate gradient algorithms**. In: XXXV BRAZILIAN COMMUNICATIONS AND SIGNAL PROCESSING SYMPOSIUM, p. 2762–2765, March 2010.

[31] STEINWANDT, J.; LAMARE, R. ; HAARDT, M.. **Beamspace direction finding based on the conjugate gradient and the auxiliary vector filtering algorithms**. Signal Processing, 93(4):641–651, 2013.

[32] QIU, L.; LAMARE, R. ; ZHAO, M.. **Reduced-rank DOA estimation algorithms based on alternating low-rank decomposition**. IEEE Signal Processing, 23(5):565–569, 2016.

[33] THOMAS, J.; SCHARF, L. ; TUFTS, D.. **The probability of a subspace swap in a SVD**. IEEE Transactions on Signal Processing, 43(3):730–736, 1995.

[34] CARLSON, B. D.. **Covariance matrix estimation errors and diagonal loading in adaptive arrays**. IEEE Transactions on Aerospace and Electronic Systems, 24(4):397–401, 1988.

[35] CHEN, Y.; ELDAR, Y. ; HERO, A. O.. **Shrinkage algorithms for MMSE covariance estimation**. IEEE Transactions on Signal Processing, 58(10):5016–5028, 2010.

[36] RUAN, H.; LAMARE, R. C.. **Robust adaptive beamforming using a low-complexity shrinkage-based mismatch estimation algorithm**. IEEE Signal Processing Letters, 21(1):60–64, 2014.

[37] RUAN, H.; LAMARE, R. C.. **Robust adaptive beamforming based on low-rank and cross-correlation techniques**. IEEE Transactions on Signal Processing, 64(15):3919–3932, 2016.

[38] PILLAI, S. U.; KNOWN, B. H.. **Forward/backward spatial smoothing techniques for coherent signal identification**. IEEE Transactions on Acoustics, Speech and Signal Processing, 37(1):8–15, 1989.

[39] EVANS, J.; JOHNSON, J. ; SUN, D. F.. **Application of advanced signal processing techniques to angle of arrival estimation in ATC navigation and surveillance systems**. MIT Lincoln Lab, 1982.

[40] MESTRE, X.; LAGUNAS, M. A.. **Modified subspace algorithms for DOA estimation with large arrays**. IEEE Transactions on Signal Processing, 56(2):598–614, 2008.

[41] GERSHMAN, A. B.; BÖHME, J.. **Improved DOA estimation via pseudorandom resampling of spatial spectrum**. IEEE Signal Processing Letters, 4(2):54–57, 1997.

[42] VASYLYSHYN, V.. **Removing the outliers in root-MUSIC via pseudo-noise resampling and conventional beamformer**. Signal Processing, 93(12):3423–3429, 2013.

[43] QIAN, C.; HUANG, L. ; SO, H. C.. **Improved unitary root-MUSIC for DOA estimation based on pseudo-noise resampling**. IEEE Signal Processing Letters, 21(2):140–144, 2014.

[44] SHAGHAGHI, M. A.; VOROBYOV, S. A.. **Subspace leakage analysis and improved DOA estimation with small sample size**. IEEE Transactions on Signal Processing, 63(12):3251–3265, 2015.

[45] STOICA, P.; NEHORAI, A.. **Performance study on conditional and unconditional direction of arrival estimation**. IEEE Transactions on Acoustics Speech and Signal Processing, 38(10):1783–1795, 1990.

[46] JOHNSON, B. A.; ABRAMOVICH, Y. I. ; MESTRE, X.. **MUSIC, G-MUSIC and maximum-likelihood performance breakdown**. IEEE Transactions on Signal Processing, 56(8):3944–3958, 2008.

[47] PINTO, S. F. B.; LAMARE, R. C.. **Two-step knowledge-aided iterative ESPRIT algorithm**. In: 21 TH INTERNATIONAL ITG WORKSHOP ON SMART ANTENNAS, p. 1–5, March 2017.

[48] BERGIN, J. S.; TEIXEIRA, C. M.; TECHAU, P. M. ; GUERCI, J. R.. **Improved clutter mitigation performance using knowledge-aided space-time adaptive processing**. IEEE Transactions on Aerospace and Electronic Systems, 42(3):997–1009, 2006.

[49] STOICA, P.; LI, J. ; AD J. GUERCI, X. Z.. **On using a priori knowledge in space-time adaptive processing**. IEEE Transactions on Signal Processing, 56(6):2598–2602, 2008.

[50] BOULEUX, G.; STOICA, P. ; BOYER, R.. **An optimal prior knowledge-based DOA estimation method**. In: 17TH EUROPEAN SIGNAL PROCESSING CONFERENCE, p. 869–873, August 2009.

[51] PINTO, S. F. B.; LAMARE, R. C.. **Multi-step knowledge-aided iterative ESPRIT for direction finding**. In: 22ND INTERNATIONAL CONFERENCE ON DIGITAL SIGNAL PROCESSING, p. 1–5, August 2017.

[52] PINTO, S. F. B.; LAMARE, R. C.. **Multi-step knowledge-aided iterative ESPRIT: analysis and design**. IEEE Transactions on Aerospace and Electronic Systems, 99(early access):8, 2018.

[53] JR, J. L.; RAPPAPORT, T. S.. **Smart antennas for wireless communications: IS-95 and third generation CDMA applications**. Prentice hall, 1999.

[54] SCHELL, S. V.; GARDNER, W. A.. **High resolution direction finding, Chapter 17, Handbook of Statistics**. Elsevier, 1993.

[55] RISSANEN, J.. **Modeling by shortest data description**. Automatica, 14:465–471, 1978.

[56] LI, F.; VACCARO, R. J.. **Analysis of Min-Norm and MUSIC with arbitrary array geometry**. IEEE Transactions on Aerospace and Electronic Systems, 26(6):976–985, 1990.

[57] HAARDT, M.. **Efficient one-, and multidimensional high-resolution array signal processing**. Shaker Verlag, 1996.

[58] GRAYBILL, F. A.. **Matrices with applications in Statistics**. Wadsworth Publishing Company, 1983.

[59] KARR, A. F.. **Probability**. Springer-Verlag, 1993.

[60] CHANG, D. W.. **A matrix trace inequality for products of hermitian matrices**. Journal of Mathematical Analysis and Applications, 237(6):721–725, 1999.

[61] ALMEIDA, J. P. A.; FORTES, J. P. ; FINAMORE, W. A.. **Probability, Random Variables and Stochastic Processes**. PUC-Rio Interciencia, 2008.

[62] SHAGHAGHI, M. A.; VOROBYOV, S. A.. **Iterative root-MUSIC algorithm for DOA estimation**. In: 5TH IEEE INTERNATIONAL WORKSHOP ON COMPUTATIONAL ADVANCES IN MULTI-SENSOR ADAPTIVE PROCESSING, p. 53–56, 2013.

[63] THAKRE, A.; HAARDT, M. ; GIRIDHAR, K.. **Single snapshot spatial smoothing with improved effective array aperture**. IEEE Signal Processing Letters, 16(6), 2009.

[64] TUNCER, E.; FRIEDLANDER, B.. **Classical and Modern Direction of Arrival Estimation. Chapter 4**. Academic Press, Elsevier, 2009.

[65] PINTO, S. F. B.; LAMARE, R. C.. **Multi-step knowledge-aided iterative conjugate gradient for direction finding**. In: 22 TH INTERNATIONAL ITG WORKSHOP ON SMART ANTENNAS, p. 1–5, March 2018.

[66] PINTO, S. F. B.; LAMARE, R. C.. **Multi-step knowledge-aided iterative conjugate gradient algorithms for DOA estimation**. Digital Signal Processing, (submitted):8, 2018.

[67] PILLAI, S. U.; BAR-NESS, Y. ; HARBER, F.. **A new approach to array geometry for improved spatial spectrum estimation**. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, volumen 73, p. 1522–1524, 1985.

[68] PILLAI, S. U.; HARBER, F.. **Statistical analysis of a high resolution spatial spectrum estimator utilizing an augmented covariance matrix**. IEEE Transactions on Acoustics, Speech and Signal Processing, 35(11):1517–1523, 1987.

[69] MOFFET, A.. **Minimum-redundancy linear arrays**. IEEE Transactions on Antenna and Propagation, 16:172–175, 1968.

[70] ABRAMOVICH, Y. I.; GRAY, D. A.; GOROKHOV, A. Y. ; SPENCER, N. K.. **Positive-semidefinite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays.I. fullly augmentable arrays**. IEEE Transactions on Signal Processing, 46:2458–2471, 1998.

[71] ABRAMOVICH, Y. I.; SPENCER, N. K. ; GOROKHOV, A. Y.. **Positive-semidefinite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays.II. partially augmentable arrays**. IEEE Transactions on Signal Processing, 47:1502–1521, 1999.

[72] JOHNSON, D. H.; DUDGEON, D. E.. **Array Signal Processing - Concepts and Techniques**. Prentice Hall, 1993.

[73] LINEBARGER, D. A.; SUDBOROUGH, I. H. ; TOLLIS, I. G.. **Difference bases and sparse sensor arrays**. IEEE Transactions on Information Theory, 39:716–721, 1993.

[74] CHEN, C. Y.; VAIDYANATHAN, P. P.. **Minimum redundancy MIMO radars**. In: IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS, p. 45–48, 2008.

[75] PEARSON, D.; PILLAI, S. U. ; LEE, Y.. **An algorithm for near-optimal placement of sensor elements**. IEEE Transactions on Information Theory, 36:1280–1284, 1990.

[76] RUF, C. S.. **Numerical annealing of low-redundancy linear arrays**. IEEE Transactions on Antennas and Propagation, 41:85–90, 1993.

[77] DOGAN, M. C.; MENDEL, J. M.. **Aplications of cumulants to array processing. aperture extension and array calibration**. IEEE Transactions on Signal Processing, 43:1200–1216, 1995.

[78] CHEVALIER, P.; ALBERA, L.; FERREOL, A. ; COMON, P.. **On the virtual array concept for higher order array processing**. IEEE Transactions on Signal Processing, 53:1254–1271, 2005.

[79] CHEVALIER, P.; FERREOL, A.. **On the virtual array concept for the fourth-order direction finding problem**. IEEE Transactions on Signal Processing, 47:2592–2595, 1999.

[80] PORAT, B.; FRIEDLANDER, B.. **Direction finding algorithms based on high-order statistics**. IEEE Transactions on Signal Processing, 39:2016–2024, 1991.

[81] MA, W. K.; HSIEH, T. H. ; CHI, C. Y.. **DOA estimation of quasi-stationary signals with less sensors than sources and unknown spatial noise covariance: a Khatri-Rao subspace approach**. IEEE Transactions on Signal Processing, 58(4):2168–2179, 2010.

[82] MA, W. K.; HSIEH, T. H. ; CHI, C. Y.. **DOAestimation of quasi-stationary signals via Khatri-Rao subspace**. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, p. 2165–2168, 2009.

[83] ASSANO, F.; HAYAMIZU, S. ; NAKAMURA, S.. **Speech enhancement based on the subspace method**. IEEE Transactions on Speech and Audio Processing, 8(5):497–507, 2000.

[84] BLISS, D. W.; FORSYTHE, K. W.. **Multiple-input multiple-output (mimo) radar and imaging: degrees of freedom and resolution**. In: 37TH IEEE ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS, volumen 1, p. 54–59, 2003.

[85] PAL, P.; VAIDYNATHAN, P. P.. **A novel array structure for directions-of-arrival estimation with increased degrees of freedom**. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, p. 2606–2609, 2010.

[86] PAL, P.; VAIDYNATHAN, P. P.. **Nested arrays: A novel approach to array processing with enhanced degrees of freedom**. IEEE Transactions on Signal Processing, 58(8):4167–4181, 2010.

[87] HAN, K.; NEHORAI, A.. **Nested array processing for distributed sources**. IEEE Signal Processing Letters, 21(9):1111–1114, 2014.

[88] YANG, J.; LIAO, G. ; LI, J.. **Robust adaptive beamforming in nested array**. Signal Processing, 114:143–149, 2015.

[89] PINTO, S. F. B.; LAMARE, R. C.. **Multi-step knowledge-aided iterative MUSIC for direction finding using nested arrays**. In: 26TH EUROPEAN SIGNAL PROCESSING CONFERENCE, número submitted, 2018.

[90] SHAN, T. J.; WAX, M. ; KAILATH, T.. **On spatial smoothing for direction of arrival estimation of coherent signals**. IEEE Transactions on Acoustics, Speech and Signal Processing, 33(4):806–811, 1985.

[91] FRIEDLANDER, B.; WEISS, A.. **Direction finding using spatial smoothing with interpolated arrays**. IEEE Transactions on Aerospace and Electronic Systems, 28:574–587, 1992.

[92] HOCTOR, R. T.; KASSAM, S. A.. **The unifying role of the coarray in aperture synthesis for coherent and incoherent imaging**. In: PROCEEDINGS OF IEEE, volumen 78, p. 735–752, April 1990.

A

Published paper

The following paper was published ...