Direction of Arrival Estimation and Tracking with Sparse Arrays

Jian-Feng Gu

A Thesis
in
The Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montréal, Québec, Canada

November 2013

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Jian-Feng Gu

Entitled: Direction of Arrival Estimation and Tracking with Sparse Arrays

and submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Electrical & Computer Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Maria Elektorowicz _____ Chair

_____ External Examiner
Dr. Henry Leung

_____ Examiner
Dr. C.-Y. Su

_____ Examiner
Dr. S. H. Zad

_____ Examiner
Dr. M.O. Ahmad

_____ Supervisor
Dr. W.-P. Zhu

_____ Supervisor
Dr. M.N.S. Swamy

Approved by _____
Dr. Abdel R. Sebak, Graduate Program Director

November 13 2013

Dr. Robin A.L. Drew, Dean
Faculty of Engineering & Computer Science

ABSTRACT

Direction of Arrival Estimation and Tracking with Sparse Arrays

Jian-Feng Gu, Ph.D.
Concordia University, 2013

Direction of Arrival (DOA) estimation and tracking of a plane wave or multiple plane waves impinging on an array of sensors from noisy data are two of the most important tasks in array signal processing, which have attracted tremendous research interest over the past several decades. It is well-known that the estimation accuracy, angular resolution, tracking capacity, computational complexity, and hardware implementation cost of a DOA estimation and/or tracking technique depend largely on the array geometry. Large arrays with many sensors provide accurate DOA estimation and perfect target tracking, but they usually suffer from a high cost for hardware implementation. Sparse arrays can yield similar DOA estimates and tracking performance with fewer elements for the same-size array aperture as compared to the traditional uniform arrays. In addition, the signals of interest may have rich temporal information that can be exploited to effectively eliminate background noise and significantly improve the performance and capacity of DOA estimation and tracking, and/or even dramatically reduce the computational burden of estimation and tracking algorithms. Therefore, this thesis aims to provide some solutions to improving the DOA estimation and tracking performance by designing sparse arrays and exploiting prior knowledge of the incident signals such as AR modeled sources and known waveforms.

First, we design two sparse linear arrays to efficiently extend the array aperture

and improve the DOA estimation performance. One scheme is called minimum redundancy sparse subarrays (MRSSA), where the subarrays are used to obtain an extended correlation matrix according to the principle of minimum redundancy linear array (MRLA). The other linear array is constructed using two sparse ULAs, where the inter-sensor spacing within the same ULA is much larger than half wavelength. Moreover, we propose a 2-D DOA estimation method based on sparse L-shaped arrays, where the signal subspace is selected from the noise-free correlation matrix without requiring the eigen-decomposition to estimate the elevation angle, while the azimuth angles are estimated based on the modified total least squares (TLS) technique.

Second, we develop two DOA estimation and tracking methods for autoregressive (AR) modeled signal source using sparse linear arrays together with Kalman filter and LS-based techniques. The proposed methods consist of two common stages: in the first stage, the sources modeled by AR processes are estimated by the celebrated Kalman filter and in the second stage, the efficient LS or TLS techniques are employed to estimate the DOAs and AR coefficients simultaneously. The AR-modeled sources can provide useful temporal information to handle cases such as the ones, where the number of sources is larger than the number of antennas. In the first method, we exploit the symmetric array to transfer a complex-valued nonlinear problem to a real-valued linear one, which can reduce the computational complexity, while in the second method, we use the ordinary sparse arrays to provide a more accurate DOA estimation.

Finally, we study the problem of estimating and tracking the direction of arrivals (DOAs) of multiple moving targets with known signal source waveforms and unknown

gains in the presence of Gaussian noise using a sparse sensor array. The core idea is to consider the output of each sensor as a linear regression model, each of whose coefficients contains a pair of DOAs and gain information corresponding to one target. These coefficients are determined by solving a linear least squares problem and then updating recursively, based on a block QR decomposition recursive least squares (QRD-RLS) technique or a block regularized LS technique. It is shown that the coefficients from different sensors have the same amplitude, but variable phase information for the same signal. Then, simple algebraic manipulations and the well-known generalized least squares (GLS) are used to obtain an asymptotically-optimal DOA estimate without requiring a search over a large region of the parameter space.

To my dear wife Lang Luo for her emotional support and boundless love, and to our beloved son Steven

ACKNOWLEDGEMENTS

My sincerest thanks go to my supervisors Drs. Wei-Ping Zhu and M.N.S. Swamy, for their guidance, patience, and encouragement throughout my research at Concordia University. I am thankful to Dr. S. C. Chan of the Department of Electrical Engineering at the University of Hong Kong for his guidance during my internship research in his lab. I want to acknowledge the members of my Ph.D. supervisory committee at Concordia University: Dr. M. Omair Ahmad, Dr. Shahin Hashtrudi Zad, and Dr. Chun-Yi Su. Their time and efforts on reading my doctoral proposal, seminal report and thesis and providing comments on my research are greatly appreciated. I would also like to thank the external examiner Dr. Henry Leung from University of Calgary for his careful review of my thesis and his participation in my defense.

I am grateful to all my friends from Concordia University and University of Electronic Science and Technology of China (UESTC). They gave me great help and encouragement.

Finally, I would like to say thanks to my parents, my parents-in-law, my sister and her family for always believing in me and encouraging me to achieve my goals. This thesis would not have been possible without my wife, Lang Luo, who provided emotional support and boundless love. Our beloved son Steven born during my research thesis has always cheered me up with his lovely behavior.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$\boldsymbol{X}^T$	Transpose of $\boldsymbol{X}$
$\boldsymbol{X}^H$	Complex conjugate transpose of $\boldsymbol{X}$
$\boldsymbol{X}^*$	Conjugate transpose of $\boldsymbol{X}$
$\boldsymbol{X}^{-1}$	Inverse of of $\boldsymbol{X}$
$\boldsymbol{X}^\dagger$	Pseudo-inverse (Moore-Penrose inverse) of of of $\boldsymbol{X}$
$\mathrm{Tr}(\boldsymbol{X})$	Trace of $\boldsymbol{X}$
$\mathrm{diag}(\boldsymbol{x})$	Diagonal matrix with its diagonal elements denoted by $\boldsymbol{x}$
$\|\boldsymbol{X}\|_F$	Frobenius norm of $\boldsymbol{X}$
$\otimes$	Kronecker product
$*$	Khatri-Rao matrix product
$\circ$	Hadamard product
δ	Dirac delta function
$\lceil x \rceil$	Largest integer less than x
$\lfloor x \rfloor$	Smallest integer greater than x
$E[x]$	Expectation of x or $E[x] \triangleq \lim_{N\to\infty} \frac{\sum_{i=1}^{N}(x)}{N}$
$\boldsymbol{I}_n$	Identity matrix or unit matrix of size n
$\boldsymbol{X}(a:b,k)$	From a^{th} to b^{th} elements of the k^{th} column of $\boldsymbol{X}$
$\boldsymbol{x}(a:b)$	From a^{th} to b^{th} elements of $\boldsymbol{x}$
$\arg(x)$	Argument or phase of x
mod	Modulo operation
$\det(\boldsymbol{X})$	Determinant of $\boldsymbol{X}$
$P_{\boldsymbol{X}}$	Projection matrix of $\boldsymbol{X}$
$P_{\boldsymbol{X}}^{\perp}$	Orthogonal projection matrix of $\boldsymbol{X}$
$\|x\|$	Absolute value of x

Rank($\boldsymbol{X}$)	Rank of $\boldsymbol{X}$
Re($\boldsymbol{x}$) or Im($\boldsymbol{x}$)	Real or imaginary part of $\boldsymbol{x}$
vec($\boldsymbol{X}$)	Vectorization of $\boldsymbol{X}$
$[\boldsymbol{X}]_{a,b}$	a^{th} row and b^{th} column element of $\boldsymbol{X}$

LIST OF ACRONYMS

ANST	Alternating Null-Steering Technique
AP	Alternating Projection
ASP	Array Signal Processing
BLUE	Best Linear Unbiased Estimator
CDEML	Coherent Decoupled Maximum Likelihood
CCM	Cross Correlation Matrix
CRT	Chinese remainder theorem
DF	Direction Finding
DOA	Direction of Arrival
EM	Expectation Maximization
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Techniques
EVD	Eigenvalue Decomposition
KF	Kalman Filter
GLS	Generalized Least Squares
i.i.d	independent and identically distributed
IQML	Iterative Quadratic Maximum Likelihood
LS	Least Squares
MODE	Method of Direction Estimation
MRLA	Minimum Redundancy Linear Arrays
MRSSA	Minimum Redundancy Sparse Subarrays

MUSIC	MUltiple SIgnal Classification
MVUE	Minimum Variance Unbiased Estimator
PADEL	PArallel DEComposition
QRD-RLS	QR Decomposition based Recursive Least Squares
RMSE	Root Mean Square Error
SLA	Sparse Linear Array
SNR	Signal-to-Noise Ratio
TLS	Total Least Squares
WDEML	White Decoupled Maximum Likelihood

Chapter 1

Introduction

1.1 Background and Motivation

Array signal processing (ASP), as an important sub-area of signal processing, has been widely used in diverse fields of science and engineering such as radar, sonar, seismic exploration, speech enhancement, deep space communications, navigation and wireless communications. [1–4]. In radar and sonar systems, antenna arrays or hydrophone arrays are often exploited to estimate the source location, range, and velocity of objects such as aircraft, missile, and submarine [5,6]. Seismic arrays are widely used for oil exploration and detection of underground nuclear tests [7]. In acoustic and speech signal processing, microphone arrays are often used to extract some signals of interest by enhancing the reception in one or multiple directions [8]. Very large antenna arrays are employed in deep space network (DSN) to compensate for signal-to-noise ratio (SNR) of the signal from a receding deep-space spacecraft [9]. Array antenna technique, also called smart antenna [10, 11], has emerged as one of the key features in the third

generation and beyond wireless communication systems such as TD-SCDMA, which can significantly improve system operating parameters, such as capacity, quality, and coverage, and reduce the cost for green communications. The most important objective of ASP is to estimate and/or track the parameters of the source signal or capture the waveform of the signal itself by fusing temporal and spatial information of the signal sources impinging onto a set of judiciously placed antenna sensors. Though the fundamental theories and basic methods of ASP have been extensively studied over the past four decades, many issues concerning its practical applications remain to be solved. For example, in many practical applications for estimating and tracking the direction of arrivals (DOA), the array systems are limited to a light load that requires limited sensors and accessories due to the cost of hardware and computational complexity. In addition, it is also necessary for these systems to provide higher accuracy with limited sensors as well as better tracking performance compared with the traditional array configurations and estimators. Therefore, the aims of this thesis are to achieve high angular resolution and accurate estimates at a low cost by considering two strategies, one to design sparse arrays and the other to employ the temporal information. The performance of an array for both DOA estimation and tracking is closely related to its array aperture. The larger the array aperture, the more accurate the angle measurements are. Sparse arrays have fewer elements for the same size aperture as compared to fully populated arrays. The temporal information is usually utilized in wireless communications and active sonar/radar to identify the different users/targets or for other purposes. The temporal information can be used to improve the estimation accuracy and reject the noise and interference. The subsequent section

reviews the state of the art on sparse array design and waveform-based techniques for DOA estimation and tracking as well as two dimensional DOA estimators.

1.2 Literature Review

1.2.1 Sparse Arrays for DOA Estimation

DOA estimation, also called spatial spectrum estimation, refers to the estimation of direction finding signals impinging on antenna arrays. Uniformly spaced linear array (ULA) is one of the most important arrays due to the natural Fourier relationship between the beam pattern and the excitation at the array, which allows the DOA estimation problem to be treated equivalently as spectral estimation. Therefore, most of the work for DOA estimation with ULA has addressed the issue of disambiguity according to the spatial sampling theory, i.e., the inter-element spacing must be less than or equal to half the wavelength of the impinging sources. It is known that the performance of DOA estimation depends on the size of the arrays aperture [1, 12]. A large array aperture can produce more accurate DOA estimation and higher resolution for closely spaced sources, inspiring us to improve the performance of DOA estimation by extending the array aperture. Generally speaking, there are two key aspects to be considered. One is the special feature or type of the source signal such as non-Gaussian signal [13], temporally correlated sources [14], noncircular sources [15], cyclostationary signals [16], AR modeled sources [17, 18] and quasi-stationary signals [19]. The other is the array configuration [20–27]. Recently, researchers have paid more attention to

array pattern design considering that in some practical circumstances there is no a priori knowledge for the characteristics of the received signal, especially for noncooperative signals. For example, in some instances, there are only a few sensors available for system implementation, the classical regular arrays such as ULA cannot provide accurate DOA estimate. This is because the aperture size of ULA is very small due to the constraint of the spatial sampling theory. This limitation has triggered the development of arrays with inter-element space greater than half the signal wavelength. Further, to mitigate the ambiguous problem in DOA estimation, it is desirable to constitute arrays such as minimum redundancy linear arrays (MRLA) [20]. The MRLA is designed so that the number of sensor pairs that have the same spatial correlation lag is as small as possible. The authors of [21] exploited the covariance augmentation technique to extend the principle of MRLA to planar geometries. However, it is very difficult to construct a MRLA when the number of sensors is relatively large because of the involvement of the heuristic search procedure and a NP hard problem in obtaining a perfect array. In order to combat this weakness, Pal and Vaidyanathan recently presented two simple and closed-form design schemes [22, 23] to extend the effective aperture of arrays. One is named nested array, which is constructed by two or multiple uniform linear subarrays with different inter-element spacing. The other is formed by two ULAs which satisfy the so-called co-prime relationship in the inter-element spacing as well as the number of sensors. Unfortunately, these design techniques for the above mentioned arrays are still constrained by the customary half-wavelength. Therefore, in order to further extend the array aperture, some researchers have suggested configurations of sparse subarrays [24–27] each of which is constructed by a

regular array, while the inter-subarray spacing is much larger than half wavelength of the signal of interest. In general, this DOA estimation method contains two steps. The first step is to implement the traditional DOA methods to obtain the rough DOA estimate without ambiguity and the cyclically ambiguous values of the fine DOA. The second step is to resolve the cyclic ambiguity by some disambiguation procedure such as beamforming, MUSIC or MODE-based method.

1.2.2 2-D DOA Estimation Methods

The problem of two-dimensional (2-D) DOA (i.e., azimuth and elevation angles) has been receiving increasing attention in the recent past. 2-D DOA problem may be closer to some practical environment than 1-D, for instance, using an airborne or a spaceborne array to observe ground-based sources. Additionally, in the last three decades a number of the high-resolution direction finding methods have been studied in the context of 1-D estimation (e.g., the azimuth angle) of multiple plane waves. Among them, MUSIC [28] and ESPRIT [29] are two of the most popular algorithms. Many 2-D DOA methods are based on the two algorithms. Specially, the latter method has two main advantages over the former. First, the ESPRIT algorithm requires less computational burden and storage space due to the fact that it does not require to search over these whole parameter space. Second, independent of the array response, the ESPRIT algorithm is more robust to array calibration errors. Therefore, the ESPRIT algorithm and its variations [30,31], which are widely devoted to the problem of 2-D DOA estimation with planar arrays, have received considerable attention in

the array-processing literature.

Majority of the planar arrays required to implement these techniques can be divided into three types: the triangular array [32,33], the rectangular array [34,35],and the two-orthogonal ULA or the L-shaped array [36–46]. Although the L-shaped array has a simpler configuration compared to the rectangular and triangular ones, it enjoys higher accuracy among these configurations [36]. Thus the L-shaped array has received increased attention in dealing with 2-D DOA estimation problems recently. In [37], Tayem and Kwon presented a computationally simple 2-D DOA estimation with the propagator method using one or two L-shaped arrays. They showed that it is possible to decompose the 2-D problem into two independent 1-D problems by using the L-shaped array for reducing the computational burden significantly. But, the two independent sets of angles would have to be properly paired together using some appropriate techniques [38]. Different approaches have been put forward in the literature for this purpose. For example, Kikuchi et al. [39] have suggested a cross-correlation technique to obtain the correct parameter pairs by constructing a Toeplitz matrix. The first column and first row of the Toeplitz matrix are constructed by the diagonal elements and their conjugate transposes of the cross correlation matrix (CCM). Then, the one-to-one relationship between the elevation and azimuth angles is set up. Unfortunately, the Kikuchis approach still suffers from the pair-matching problem when the difference of the corresponding combinations of the 2-D angles is small and the signal-to-noise ratio (SNR) is low [40]. Furthermore, it only employs the CCM to deal with the pair-matching problem such as the pairing algorithm suggested in [41], but does not exploit its characteristics to improve the estimation performance.

Gu and Wei [40] have proposed a joint singular value decomposition(JSVD) technique that constructs the extension signal subspace by selecting two submatrices from the CCM, which is unaffected by the additive noise. By this scheme, we can make use of the property of the eigendecomposition, i.e.,the eigenvalue and its unique eigenvector, to achieve automatic pairing and estimate 2-D DOA. Therefore, the JSVD technique enjoys at least two advantages over the technique suggested in [39]. First, the JSVD technique needs no additional steps to deal with the pair matching problem. Second, the JSVD is superior in estimating the 2-D DOA, especially at low SNR and with a small number of snapshots. The authors of [43] have presented a generalized ESPRIT-based technique to deal with the problem of pair-matching. Unfortunately, the computational burden of this technique is very high due to its requirement for search over the parameter space of interest and implementation for the eigenvalue decomposition (EVD) of the array correlation matrix. Therefore, a computationally efficient method is proposed in [44] based on propagator method, but it still involves considerable cost in computation (the parameter space searching) to estimate 2-D DOA estimation and deal with the pair-matching problem.

1.2.3 DOA Estimation with Known Waveform

As is well known, most DOA estimation algorithms, such as beamforming-based techniques [7, 47], subspace-based techniques [48], and sparsity-based techniques [49,

50], are mainly based on a common assumption that the received signals are non-cooperative signals, that is, they are either unknown deterministic signals or Gaussian type of stochastic signal sources with unknown covariance. In some applications, such as active radar, active sonar, and communication systems, the basic waveform of signal of interest is available to its receiver. This a priori information can be exploited to effectively eliminate background noises and significantly improve the estimation accuracy [51, 52]. In addition, the capacity of DOA estimation can be larger than the number of antenna elements [53–59]. Only a few techniques have been developed so far to handle the DOA estimation problem by making use of the waveforms of signal sources. Li and Compton [53] are among the very first researchers to improve the accuracy of DOA estimation with known waveforms. They obtained initial angle estimates using an iterative quadratic maximum likelihood (IQML) algorithm, and then used the alternating projection (AP) or the expectation maximization (EM) algorithm to estimate the DOAs. Later on, a large sample decoupled ML estimator (DEML) was proposed to estimate the DOAs of incoherent signals with known waveforms [54]. The DEML estimator is computationally efficient, since it decouples the multidimensional minimization problem into a set of 1-D minimization problems. However, this estimator encounters difficulty when the signals impinging on the array are coherent. To lift this constraint, Cederval and Moses [55] extended the DEML estimator to decorrelate the coherency of incident signals and developed the coherent decoupled maximum likelihood (CDEML) algorithm. Both DEML and CDEML belong to the family of large sample ML algorithms, which do not work well in difficult scenarios such as when the SNR is low or the number of snapshots is small. To improve

the accuracy and spatial resolution of the DOA estimation for signals with known waveforms, Li et al. [56] proposed a white decoupled maximum likelihood estimator (WDEML) under the assumption that the observed noise is spatially white. Recently, Atallah and Marcos [57] have presented a parallel decomposition (PADEC) algorithm that yields comparable performance, but with a lower complexity, than that of the ML-based algorithms. The idea behind the PADEC algorithm is to obtain spatial signature of the signals using the least-squares (LS) error criterion, and to decorrelate the coherence of the signals by applying spatial smoothing techniques. However, for large size subarrays, the computational burden of PADEC may be unacceptably high, since the eigen-decomposition is required to obtain the orthogonal projector on the noise subspace or the signal subspace. A computationally simpler and more efficient DOA estimation technique has been proposed in [58], where the DOA of known signal waveforms is computed based on the phase shift between two subarrays. This technique requires that signals from different sources be uncorrelated with one another; thus, it does not perform well when the signals are partially or completely correlated. More recently, Gu et al. [59] have suggested a fast linear operator to deal with DOA estimation of uncorrelated or coherent signal sources based on their waveforms. This method does not require the reconstruction of orthogonal projector in the noise subspace or the signal subspace but its performance approaches to that of the ML-based methods.

1.2.4 DOA Tracking

Another focus of this thesis is on DOA tracking, which is closely related to DOA estimation. A number of standard methods exist for such a problem. Eiegnstructure or so-called subspace tracking techniques [60, 61], for example, attempt to track DOA via repeated implementation of subspace-based DOA estimation techniques such as MUSIC, ESPRIT [28, 29, 48] that rely on recursively updating the eigenstructure or subspace information obtained from either the singular value decomposition (SVD) of the array output or the EVD of the covariance matrix estimate of sampled array data. However, there are two major limitations that are inherited by all subspace tracking approaches. One is that each updated set of DOAs suffers from the data association problem [62]. In other words, although subspace tracking techniques can efficiently estimate and track the whole updated DOA values, they cannot set up a one-to-one relationship between the estimated DOAs and the targets automatically [63]. The second main limitation of the subspace tracking techniques lies in that it is difficult to incorporate a prior knowledge of the signal feature and/or array structure into the eigendecomposition. It is well known that temporal information of signals and special array structure can be exploited to effectively eliminate background noise and significantly improve the performance and capacity of DOA estimation and tracking, and/or even dramatically simplify the computational burden of the estimation and tracking algorithms. Therefore, many approaches to track the DOA of multiple targets make use of the array structure and/or the covariance matrix of signals to directly update the DOA or the spatial signature of the respective targets without

changing the order of DOA estimates [64–75]. By using the structure of the array manifold, Sword et al. [64] proposed a closed-form LS solution to update the most recent DOA estimates, thus avoiding the data association problem. Later, Lo and Li [65] modified the algorithm by implementing the error-correction procedure to reduce the effect of the error propagation due to the use of recursive approximations. The authors of [66] have made use of the inherent dynamical property of the DOA of moving targets to improve the capability for the case of crossing tracks, where two DOAs are very close or even overlap. In [67], Sastry et al. used the property of the Frobenius norm of the covariance error matrix, which is sensitive to permutations in the columns of the array steering matrix, to update the current DOA estimates of targets without the data association problem; but, this method can only be used in the case of different signal powers. Inspired by such a property reported in [67], the authors of [68] simplified the objective function as the distance between the corresponding elements of the previously estimated and current covariance matrix. Satish and Kashyap [69] derived a maximum likelihood (ML)-based technique for optimal determination of the current DOA and range estimates for slowly changing targets based on the second-order approximation of the inverse of the array covariance matrix. The authors of [70] then introduced a recursive expectation and maximization algorithm to reduce the computational burden of the traditional ML-based technique. The ML-based tracking technique suggested in [71] makes use of the target motion state to improve the tracking performance, where the DOA estimates are updated at each time frame and refining through Kalman filtering. Zhou et al. [72] obtained the DOA tracking through updating the information of the target motion state described

as the multiple target state (MTS). Recently, the authors of [73] presented a signal selective DOA tracking technique by using the special features of cyclostationary sources to improve the tracking performance for wideband multiple moving sources without combating the association problem.

1.3 Organization and Contributions

The organization of the thesis along with the main contributions of each chapter is presented as follows.

In Chapter 2, accurate DOA estimation methods for noncooperative signals are investigated with a special focus on the estimators applicable to sparse linear arrays (SLA) and sparse L-shaped arrays that are constructed by two linear arrays perpendicular with each other. First, we present a new array geometry named minimum redundancy sparse subarray (MRSSA) that is considered by uniform linear subarray (ULS) according to the principle of MRLA, where the inter-subarray spacing is much larger than the half wavelength, and each ULS is composed of the ULA with inter-element spacing less than or equal to half of the wavelength. An extended correlation matrix is constructed from Kronecker Steering Vectors (KSVs) each of which contains the ambiguous angle and the corresponding unambiguous angle. Subsequently, we propose a new array geometry named nonuniform linear sparse array(NLSA), that is composed of uniform linear sparse subarrays with the inter-sensor spacing in the same subarray much larger than the half wavelength. However, the minimum distance among sensors must be less than or equal to the half wavelength to avoid the

multi-value ambiguity. Furthermore, we propose a joint elevation and azimuth DOA estimation method using the L-shaped array geometry constructed by two linear arrays that are perpendicular to each other. The signal subspace of the elevation angle can be obtained directly by the property of CCM without the effects of an unknown noise field. The ESPRIT-like algorithm can be employed to estimate the elevation angles without finding polynomial roots or searching over parameter space. Next, we derive a computational efficient modified TLS method to estimate the azimuth angle by employing the estimated waveforms and elevation angles of the incident signals

In Chapter 3, Kalman filter technique is used to estimate and track the DOA of AR modeled source signals. First, a novel DOA estimation method for AR modeled source signals impinging on SLA is proposed. Since each sensor can be considered as a dynamic model of the time-varying AR sources where each regression coefficient contains the information of DOA, we employ Kalman filter to obtain the source signal estimates and then the TLS technique is used to derive an approximate optimal estimator for the DOA of signal. In addition, we propose a new DOA estimation and tracking method for AR modeled signals based on symmetric sparse array.

In Chapter 4, we propose a novel DOA estimation method of multiple signals with known source waveforms and unknown gains based on SLAs. By using linear regression analysis, the proposed algorithm is presented as an optimal estimator for simultaneous DOA and complex gain estimation. The output of each sensor of the antenna array, as a combination of the received signals of interest, is expressed as a linear regression model where each regression coefficient contains the information

of DOA and the corresponding complex gain. A new technique for unwrapping ambiguity by the Chinese remainder theorem (CRT) is then presented to extract the angle information from the estimated complex coefficients. Further, the well-known generalized LS technique is used to obtain asymptotically-optimal estimate of DOAs without requiring heavy computation. Based on the idea of the proposed array geometry design and DOA estimation method, we also derive two LS-based schemes for moving targets, to update the coefficient changes of each sensor at successive time intervals: one is based on the block QRD-RLS technique and the other on the block regularized LS technique.

Finally, Chapter 5 contains conclusions and provides directions for future work.

Chapter 2

Accurate DOA Estimation by Sparse Arrays

2.1 Introduction

Designing nonuniform linear arrays to obtain the accurate angle estimate is very popular in radar systems such as airborne surveillance radar, ground based radar systems, and shipborne radar systems [76] because of low cost and complexity. Most of these techniques are considered to construct the "no holes" covariance matrix, which still limits the extension of array aperture [1]. In this chapter, we focus on the DOA estimation with sparse arrays, where the inter-element spacings can be much larger than the half-wavelength. In Section 2.2, the first sparse array called the minimum redundancy sparse subarray is designed to estimate 1-D DOA, where uniform linear subarrays are employed to construct the whole arrays according to the principle of MRLA. Kronecker steering vectors (KSVs) are constructed using the

relationship between the subarrays, and then a modified ESPRIT approach is used to find all the KSVs. Finally, the accurate DOAs can be obtained by solving a simple algebraic problem. The proposed method enjoys two advantages in comparison to some of the existing methods. First, the cyclic ambiguity can be resolved by the one-to-one mapping between unambiguous and ambiguous angles without requiring additional algorithms such as MUSIC or MODE. Second, it can deal with the case of different unambiguous angles with the same ambiguous angle, which might not be possible to deal with by using the previous schemes [24, 25]. However, the proposed method cannot deal with the cases of correlated sources. Therefore, in Section 2.3, we propose another sparse linear array and its corresponding method to handle the shortcoming of the MRSSA. The second sparse linear array is constructed by using two sparse ULAs, where each sparse ULA is constructed by interleaving sensors. We first estimate the rough DOA by the generalized ESPRIT method, and then employ the alternating null-steering technique to estimate the fine DOA. In Section 2.4, a sparse L-shaped array is designed to estimate 2-D DOA. Here, the ULA along with the ESPRIT-based method is used to estimate the elevation angle, and then the signal waveform is obtained by the estimated elevational angle. Since the elevation angle and the waveform have been obtained, each sensor of the SLA on the x-axis can be considered as a linear regression model with respect to the phase information containing the azimuth angle, which can be obtained by exploiting the modified TLS and GLS techniques.

2.2 DOA Estimation by Minimum Redundancy Sparse Subarrays

2.2.1 Proposed MRSSA Pattern

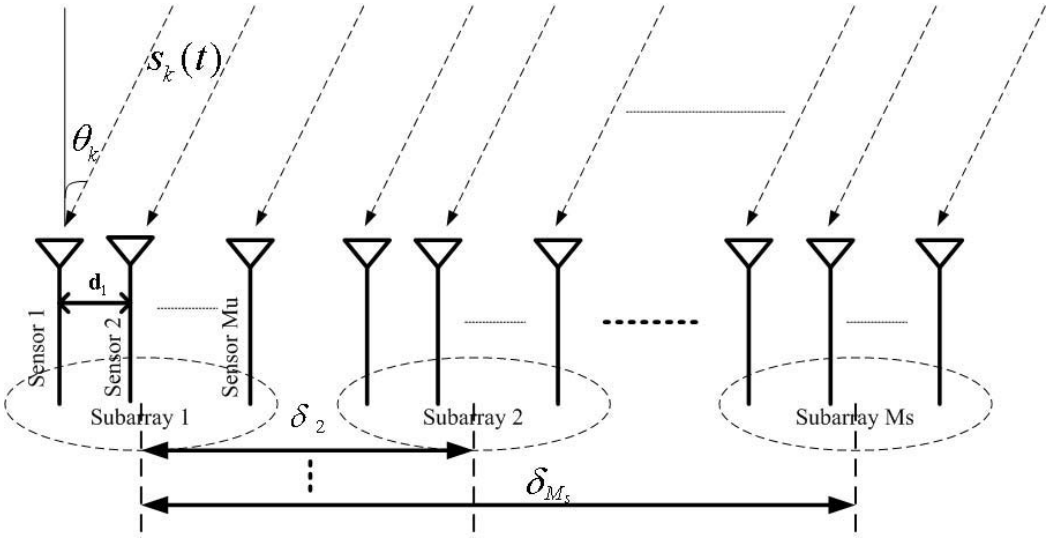


Figure 2.1: Structure of the MRSSA

Let us consider K narrowband signals with the same wavelength, say λ, that impinge on a minimum redundancy sparse subarray (MRSSA), shown in Fig.2.1. The MRSSA consists of M sensors with M_s subarrays placed according to the philosophy of MRLA [20], where each subarray contains $M_u = M/M_s$ sensors with the inter-element spacing being $d_1 \leq 1$, and the smallest intersubarray spacing between the two consecutive subarray centers is $d_2 > (M_u + 1)d_1$. Note that d_1 and d_2 are normalized distances in terms of the half wavelength. The observed signals in the p^{th} ($p = 1, \cdots, M_s$) subarray at time t can be represented by an $M_u \times 1$ complex vector as

$$\boldsymbol{x}_p(t) = \boldsymbol{A}_p \boldsymbol{s}(t) + \boldsymbol{n}_p(t) = \boldsymbol{A}_1 \boldsymbol{B}_p \boldsymbol{s}(t) + \boldsymbol{n}_p(t), \tag{2.1}$$

where $\boldsymbol{s} = [s_1(t), s_2(t), \cdots, s_K(t)]^T$ is the signal sources from K different directions, and the elements of $\boldsymbol{n}_p(t)$ are the white Guassian random processes with zero-mean and variance σ^2 . The matrix $\boldsymbol{A}_1$ is the steering matrix of the first subarray

$$\begin{aligned}\boldsymbol{A}_1 &= [\boldsymbol{a}_1(\theta_1), \boldsymbol{a}_1(\theta_2), \cdots, \boldsymbol{a}_1\theta_K)] \\ &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-jd_1\alpha_1} & e^{-jd_1\alpha_2} & \cdots & e^{-jd_1\alpha_K} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j(M_u-1)d_1\alpha_1} & e^{-j(M_u-1)d_1\alpha_2} & \cdots & e^{-j(M_u-1)d_1\alpha_K} \end{bmatrix},\end{aligned} \tag{2.2}$$

where $\boldsymbol{a}_1(\theta_k) = [1, e^{-jd_1\alpha_k}, \cdots, e^{-jd_1(M_u-1)\alpha_k}]^T$ with $\alpha_k = \pi \sin\theta_k$.

The matrix $\boldsymbol{B}_p$ in (2.1) is given by

$$\boldsymbol{B}_p = diag(e^{-j\beta_{p,1}}, e^{-j\beta_{p,2}}, \cdots, e^{-j\beta_{p,K}}), \tag{2.3}$$

with $\beta_{p,k} = \delta_p\alpha_k$ and δ_p being the spacing between the centre of the p^{th} subarray and the centre of the first subarray. To simplify the problem formulation, we assume that there are three subarrays, giving $\delta_0 = 0, \delta_1 = d_2, \delta_2 = 3d_2$. We also assume that the sources are uncorrelated so that the source correlation matrix is diagonal, namely,$\boldsymbol{R}_s = diag([\rho_1, \rho_2, \cdots, \rho_K])$, where ρ_k is the power of the k^{th} incident signal. Then the correlation matrices of the subarrays are given by

$$\begin{aligned}\boldsymbol{R}_x(d_2^i - d_2^j) \triangleq E[\boldsymbol{x}_i(t)\boldsymbol{x}_j^H(t)] &= \boldsymbol{A}_1\boldsymbol{B}_i\boldsymbol{R}_s\boldsymbol{B}_j^H\boldsymbol{A}_1^H + \sigma^2\boldsymbol{I}_{M_u}\delta(i-j) \\ &= \boldsymbol{A}_1\boldsymbol{B}_i\boldsymbol{B}_j^H\boldsymbol{R}_s\boldsymbol{A}_1^H + \sigma^2\boldsymbol{I}_{M_u}\delta(i-j)(i,j=1,2,3)\end{aligned}. \tag{2.4}$$

2.2.2 DOA Estimation Technique

From (2.4), we can obtain the following correlation matrices

$$\begin{aligned}
\boldsymbol{R}_x(0) &= \boldsymbol{A}_1\boldsymbol{R}_s\boldsymbol{A}_1^H + \sigma^2\boldsymbol{I}_{M_u} \quad (i=j) \\
\boldsymbol{R}_x(d_2) &= \boldsymbol{A}_1\boldsymbol{B}\boldsymbol{R}_s\boldsymbol{A}_1^H = \boldsymbol{R}_x(-d_2)^H \quad (i=1, j=2) \\
\boldsymbol{R}_x(2d_2) &= \boldsymbol{A}_1\boldsymbol{B}^2\boldsymbol{R}_s\boldsymbol{A}_1^H = \boldsymbol{R}_x(-2d_2)^H \quad (i=2, j=3) \\
\boldsymbol{R}_x(3d_2) &= \boldsymbol{A}_1\boldsymbol{B}^3\boldsymbol{R}_s\boldsymbol{A}_1^H = \boldsymbol{R}_x(-3d_2)^H \quad (i=1, j=3)
\end{aligned}, \tag{2.5}$$

where $\boldsymbol{B} = diag([e^{-j\beta_1}, e^{-j\beta_2}, \cdots, e^{-j\beta_K}])$ with $\beta_k = d_2\alpha_k$.

Next, let us construct a block Toeplitz matrix $\boldsymbol{R}$ of size $(4M/3) \times (4M/3)$ as follows

$$\begin{aligned}
\boldsymbol{R} &= \begin{bmatrix}
\boldsymbol{R}_x(0) & \boldsymbol{R}_x(d_2) & \boldsymbol{R}_x(2d_2) & \boldsymbol{R}_x(3d_2) \\
\boldsymbol{R}_x(-d_2) & \boldsymbol{R}_x(0) & \boldsymbol{R}_x(d_2) & \boldsymbol{R}_x(2d_2) \\
\boldsymbol{R}_x(-2d_2) & \boldsymbol{R}_x(-d_2) & \boldsymbol{R}_x(0) & \boldsymbol{R}_x(d_2) \\
\boldsymbol{R}_x(-3d_2) & \boldsymbol{R}_x(-2d_2) & \boldsymbol{R}_x(-d_2) & \boldsymbol{R}_x(0)
\end{bmatrix}, \\
&\triangleq (\boldsymbol{B}_s * \boldsymbol{A}_1)\boldsymbol{R}_s(\boldsymbol{B}_s * \boldsymbol{A}_1)^H + \sigma^2\boldsymbol{I}_{4M/3}
\end{aligned} \tag{2.6}$$

where $\boldsymbol{B}_s$ is given by

$$\boldsymbol{B}_s = [\boldsymbol{b}(\theta_1), \boldsymbol{b}(\theta_2), \cdots, \boldsymbol{b}(\theta_K)], \quad \text{with} \quad \boldsymbol{b}(\theta_k) = [1, e^{-j\beta_k}, e^{-2j\beta_k}, e^{-3j\beta_k}]. \tag{2.7}$$

Further, we define $\boldsymbol{b}(\theta_k) \otimes \boldsymbol{a}_1(\theta_k)(k = 1, 2, \cdots, K)$ as the Kronecker steering vectors, each containing a pair of unambiguous and ambiguous angles. It is worth noting that our method can detect at most $4M/3 - 1$ source signals if we use MUSIC-like method. In order to reduce the computational cost, next we will introduce a

ESPRIT-like method without searching the parameter space. Performing the eigenvalue decomposition for $\boldsymbol{R}$ in (2.6) results in

$$\boldsymbol{R} = [\boldsymbol{U}_s \quad \boldsymbol{U}_n]\boldsymbol{\Sigma}[\boldsymbol{U}_s \quad \boldsymbol{U}_n]^H, \tag{2.8}$$

where $\boldsymbol{\Sigma} = diag([\sigma_1, \cdots, \sigma_K, \sigma_{K+1}, \cdots, \sigma_{4M/3}])$ with $\sigma_1 \geq \cdots \geq \sigma_K \geq \sigma_{K+1} = \cdots = \sigma_{4M/3}$ is the diagonal matrix containing the eigenvalues of $\boldsymbol{R}$, and $\boldsymbol{U}_s = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_K]$ of size $(4M/3) \times K$ is the signal subspace or signal plus noise subspace,and $\boldsymbol{U}_n = [\boldsymbol{u}_{K+1}, \cdots, \boldsymbol{u}_{4M/3}]$ is the noise subspace. According to the principle of subspace-based methods, there exists a nonsingular $K \times K$ matrix $\boldsymbol{T}$ such that

$$\boldsymbol{U}_s = (\boldsymbol{B}_s * \boldsymbol{A}_1)\boldsymbol{T}. \tag{2.9}$$

In order to obtain the unambiguity angles first, herein we define a permutation matrix [77]

$$\boldsymbol{\Gamma} = [\boldsymbol{I}_4 \otimes e_1, \boldsymbol{I}_4 \otimes e_2, \cdots, \boldsymbol{I}_4 \otimes e_M], \tag{2.10}$$

where e_i is the i^{th} column of $\boldsymbol{I}_{M_u}$. We then have

$$\boldsymbol{U}_p \triangleq \boldsymbol{\Gamma}\boldsymbol{U}_s = (\boldsymbol{A}_1 * \boldsymbol{B}_s)\boldsymbol{T}. \tag{2.11}$$

Let us partition $\boldsymbol{U}_p$ into two $4(M_u - 1) \times K$ submatrices as

$$\begin{aligned} \boldsymbol{U}_{p1} &= \boldsymbol{U}_p(1 : 4(M_u - 1), :) \triangleq (\boldsymbol{A}_{11} * \boldsymbol{B}_s)\boldsymbol{T} \\ \boldsymbol{U}_{p2} &= \boldsymbol{U}_p(5 : 4M_u, :) \triangleq (\boldsymbol{A}_{11} * \boldsymbol{B}_s)\boldsymbol{\Psi}\boldsymbol{T}, \end{aligned} \tag{2.12}$$

where $\boldsymbol{A}_{11}$ is the first $(M_u - 1)$ rows of the array response matrix $\boldsymbol{A}_1$ and the diagonal matrix $\boldsymbol{\Psi}$ is of the form $\boldsymbol{\Psi} = diag(e^{-jd_1\alpha_1}, e^{-jd_1\alpha_2}, \cdots, e^{-jd_1\alpha_K})$. From (2.12), we can obtain

$$\boldsymbol{U}_{p2} = (\boldsymbol{A}_{11} * \boldsymbol{B}_s)\boldsymbol{T}\boldsymbol{T}^{-1}\boldsymbol{\Psi}\boldsymbol{T} = \boldsymbol{U}_{p1}\boldsymbol{T}^{-1}\boldsymbol{\Psi}\boldsymbol{T}. \tag{2.13}$$

Therefore, we can establish a one-to-one relation between the rough DOA without ambiguity and the rotationally invariant counterpart of KSVs by utilizing the eigenvalue decomposition (EVD) for the matrix $\boldsymbol{F} \triangleq \boldsymbol{U}_{p2}^{\dagger}\boldsymbol{U}_{p1} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}_{-1}$, where $\boldsymbol{V}$ denotes the eigenvectors of $\boldsymbol{F}$ and the diagonal elements of $\boldsymbol{\Lambda} = diag([\lambda_1, \lambda_2, \cdots, \lambda_K])$ are the corresponding eigenvalues and the estimated diagonal matrix $\boldsymbol{\Psi}$. Thus, λ_k contains the unambiguous DOA information and then the rough angels θ_k can be found by using the following function as

$$\hat{\theta}_k = \arcsin[-\arg(\lambda_k)/(\pi d_1)], k = 1, \cdots, K. \tag{2.14}$$

Moreover, it is easy to verify [77, 78] that $\boldsymbol{V}$ and $\boldsymbol{T}$ satisfy the relationship $\boldsymbol{v} = \boldsymbol{T}^{-1}\boldsymbol{E}\boldsymbol{P}^{-1}$, where $\boldsymbol{E}$ is the real diagonal matrix and $\boldsymbol{P}$ is so-called generalized permutation matrix. Therefore, we can obtain an estimate of the $4M_u \times K$ KSV matrix $\boldsymbol{B}_s * \boldsymbol{A}_1$,

$$\hat{\boldsymbol{D}} = \boldsymbol{U}_s\boldsymbol{V}\boldsymbol{P}\boldsymbol{\Gamma}^{-1}. \tag{2.15}$$

Herein, we employ the simple arithmetic operation suggested in [77] to obtain the estimate of β_k, i.e.,

$$\hat{\beta}_k = \arg(\hat{\boldsymbol{D}}((M/3+1) : 4M/3, k)^H \hat{\boldsymbol{D}}(1 : M, k)). \tag{2.16}$$

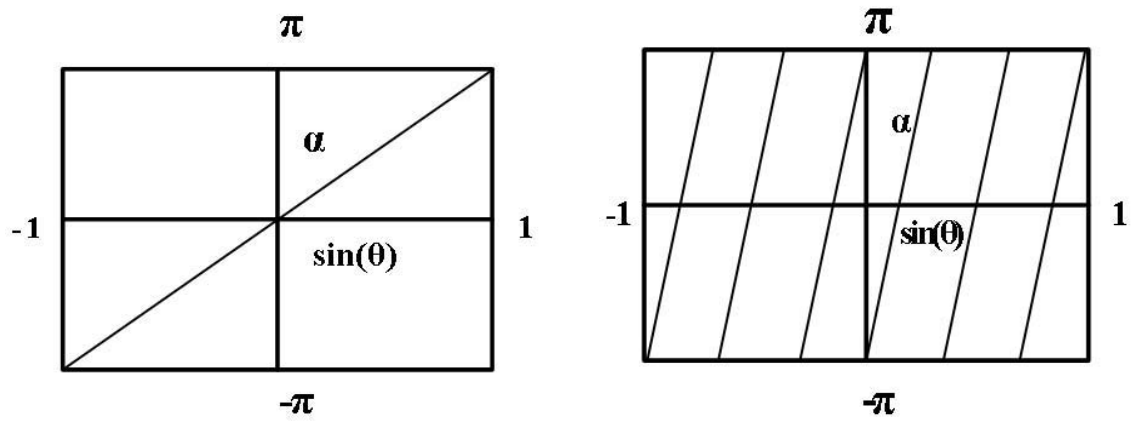
Obviously, if $d_2 \gg 1$, there may exist different angels for the same β_k according to the cyclically ambiguous for sine estimates. To better understand this concept, let us consider the situation without ambiguity. Assume that the distance between the sensors is $d = 1$. Under this condition, there is no ambiguity for $-\pi \leq \alpha = \pi \sin\theta \leq \pi$. The $\alpha \sim \theta$ plot is shown in Fig.2.2(a) indicating a one-to-one mapping between α and θ for the case of $d = 1$. When $d > 1$, for example, $d = 6$, then $\alpha = 6\pi \sin\theta \quad mod \quad 2\pi$, which gives 5 ambiguities as shown in Fig.2.2(b). The relationship between multiple-ambiguity and the true value can be expressed as

$$\alpha_k^l = (\beta_k + 2\pi l)/d_2, \quad \lceil -d_2/2 - \beta_k/(2\pi) \rceil \leq l \leq \lfloor d_2/2 - \beta_k/(2\pi) \rfloor. \tag{2.17}$$

Therefore, making use of the previously estimated rough angles in (2.14) and (2.17), the unambiguous accurate estimate for the angle can be obtained as

$$\hat{\hat{\theta}}_k = \arcsin(\arg\min_{\alpha_k^l} |\alpha_k^l - \pi \sin\hat{\theta}_k|/\pi). \tag{2.18}$$

Note that the proposed method obtains the accurate estimate by searching for the closet value to the corresponding rough estimate and therefore, it gives an improved estimation performance as compared to the previous method [24], where different rough DOA values correspond to the same fine DOA value. In addition, the proposed method can also yield a better performance than the traditional methods do due to the fact that MRLA can extend the effective array aperture.



(a) (b)

Figure 2.2: $\alpha \sim \theta$ plots for (a) $d = 1$ and (b) $d = 6$

2.2.3 Simulation Results for MRSSA

We now present some results from Monte Carlo simulation of the proposed MRSSA. We assume that there are 15 elements in the array system and in each ULA there are $M_u = 5$ sensors, which are apart by $d_1 = 1$. A total of 1000 independent realizations of the received data are adopted. In this first example, the proposed method is evaluated as compared to the method in [24] for two equal power signals with the incident directions $[\theta_1, \theta_2] = [45°, 60°]$ and the SNR varying from -10 to 20 dB. The smallest inter-subarray spacing is 10 times half-wavelength. The root mean square error (RMSE) as defined by

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{\hat{\theta}} - \theta)^2} \tag{2.19}$$

is used as the objective performance metric, where T is the number of independent trials. Fig.2.3(a) and 2.3(b) show the RMSE of $\theta_1 = 45°$ and $\theta_2 = 60°$, respectively.

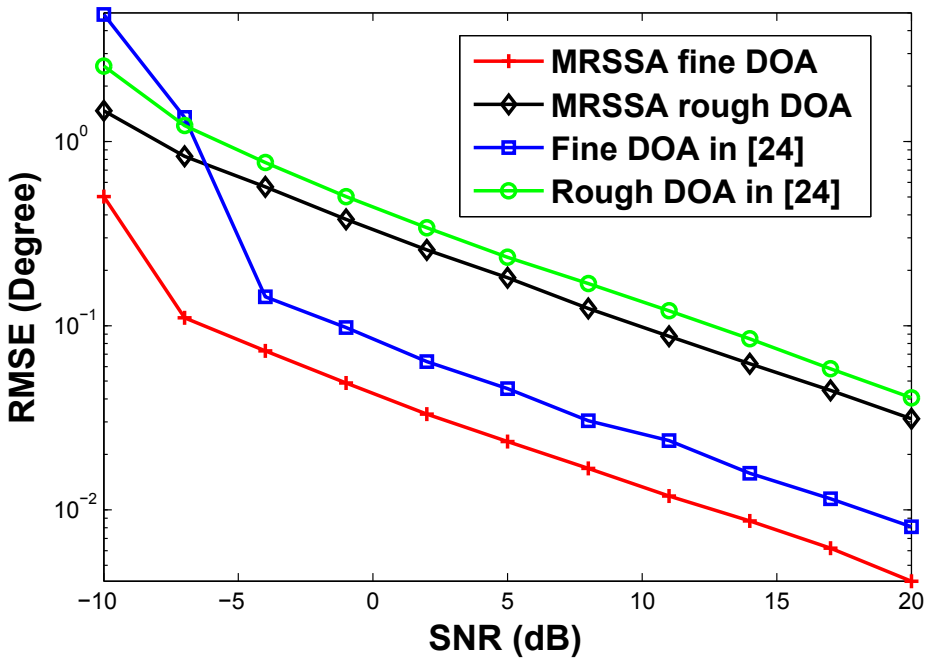
It is observed that the proposed method delivers more accurate DOA estimates, especially at low SNR. This is beacuse the proposed method uses the principle of MRLA to extend the array aperture and implements the rough and fine DOA estimation without requiring pair matching. At low SNR, the method in [24] may produce the errors between the rough angle and the accurate one, while our method does not need pair matching in this case.

In the second example, we demonstrate our method as well as the method suggested in [24] for different inter-subarray spacing. To have a fair comparison, we assume that there are two equal-power uncorrelated signals from the DOAs $[\theta_1, \theta_2] = [34^\circ, 36^\circ]$, the SNR is set to 0dB and 200 snapshots are employed. The RMSE results are shown in Fig.2.4(a) and 2.4(b). It is seen that the estimation performance of the method [24] may decrease with the increase of array aperture in some situations, while our method can still increase the estimation performance owing to the improved array aperture and automatic pair matching, which is not the case for the method suggested in [24]. This is because our method estimates the accurate and rough DOAs simultaneously, while the method in [24] estimates the rough and accurate angles individually.

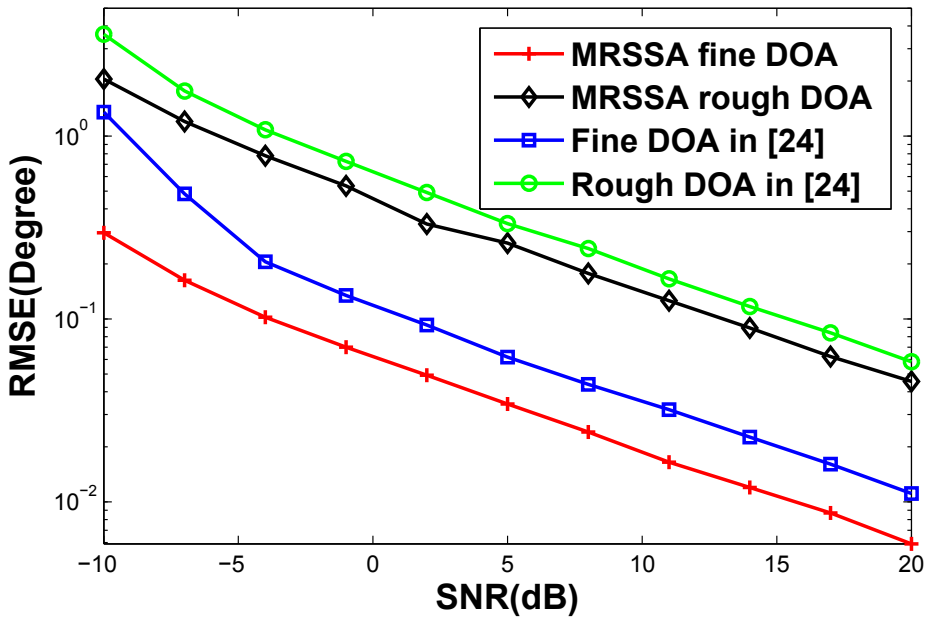
2.3 DOA Estimation by Sparse ULAs

2.3.1 Proposed Two Sparse ULAs

Different from the MRSSA mentioned above, here a nonuniform linear sparse array that consists of $2M + 1$ sensors arranged along two sparse ULAs is proposed, as shown in Fig.2.5. The inter-element spacing for the two ULAs is d_1 and d_2 satisfying

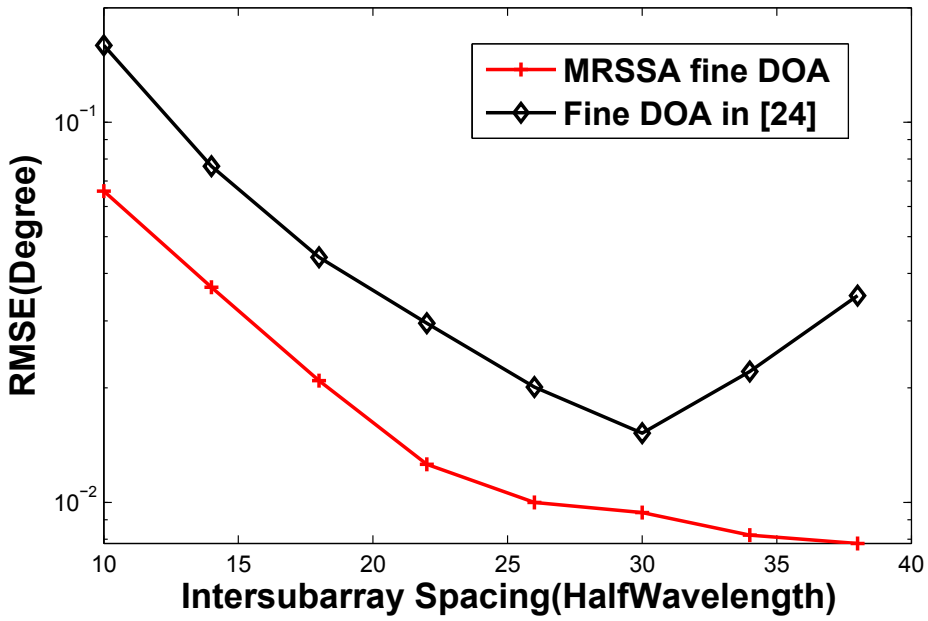


(a)

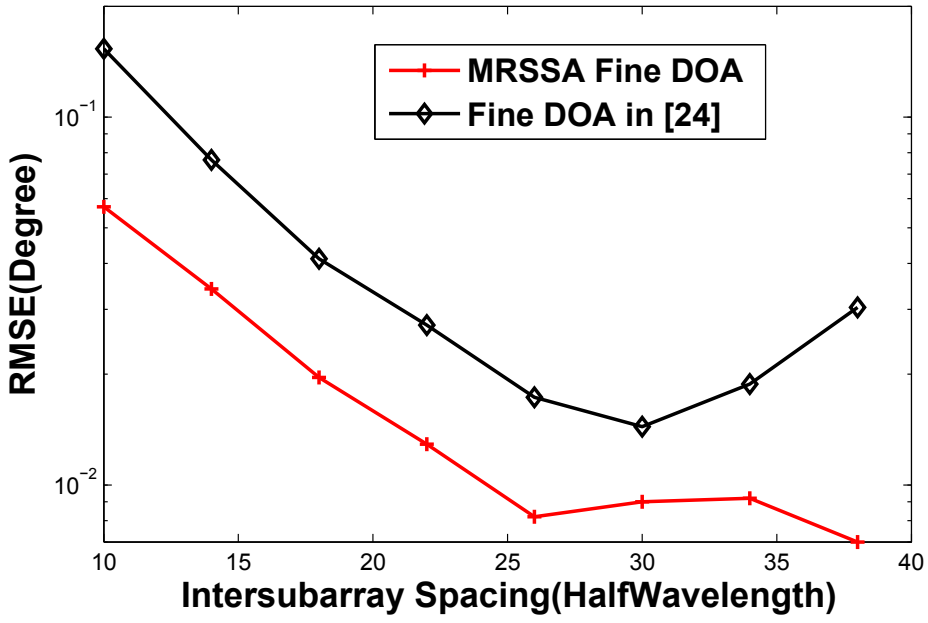


(b)

Figure 2.3: DOA estimation performance of MRSSA with three ULAs for two sources with $[\theta_1, \theta_2] = [45°, 60°]$: (a) RMSE for $\theta_1 = 45°$ (b) RMSE for $\theta_2 = 60°$. The number of snapshots is 200.



(a)



(b)

Figure 2.4: DOA estimation performance of MRSSA with three ULAs for two sources with $[\theta_1, \theta_2] = [34°, 36°]$: (a) RMSE for $\theta_1 = 34°$ (b) RMSE for $\theta_2 = 36°$. The SNR is 0dB, and the number of snapshots is 200.

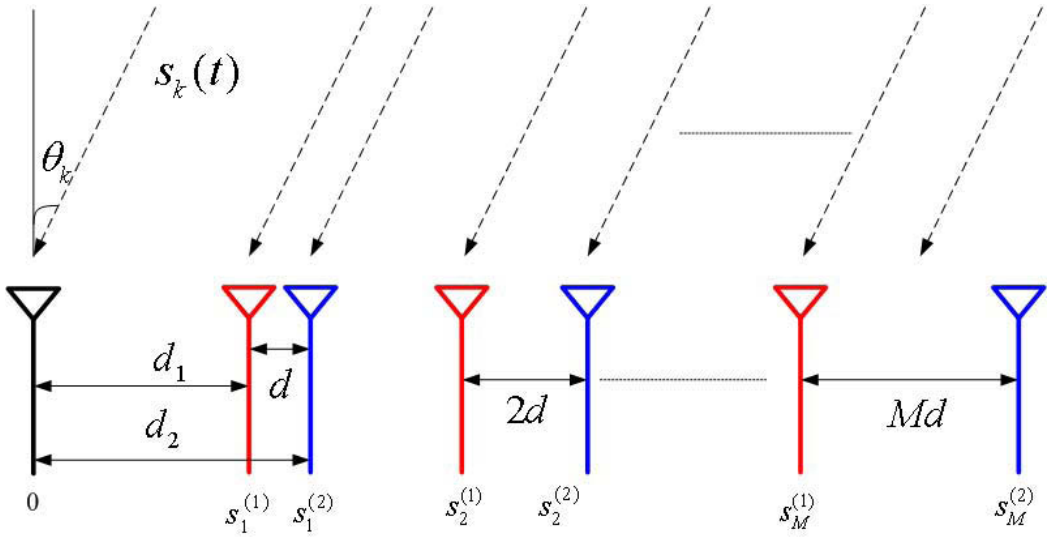


Figure 2.5: Structure of two sparse ULAs

$d_2 - d_1 = d$ and $d_1 \gg d$. The element placed at the origin is common for the two ULAs as reference. Let $m_i(m = 1, \cdots, M, i = 1, 2)$ denote the index of the m^{th} sensor along the i^{th} subarray. The observed signals in the $i^{th}(i = 1, 2)$ subarray at time t can be expressed as an $(M + 1) \times 1$ complex vector

$$\boldsymbol{x}_i(t) = \boldsymbol{A}_i \boldsymbol{s}(t) + \boldsymbol{n}_i(t) \tag{2.20}$$

where $\boldsymbol{A}_i = [\boldsymbol{a}_i(\theta_1), \boldsymbol{a}_i(\theta_2), \cdots, \boldsymbol{a}_i(\theta_K)]$ is the steering matrix of the i^{th} subarray with $\boldsymbol{a}_i(\theta_k) = [1, e^{-jd_i\alpha_k}, \cdots, e^{-j(M-1)d_i\alpha_k}]^T$. In (2.20), $\boldsymbol{n}_i(t)$ is additive noise as in MRSSA. Then, the signal received by the entire array can be written as

$$\boldsymbol{x}(t) = [\boldsymbol{A}_1^T \boldsymbol{\Pi}, \boldsymbol{A}_2^T]^T \boldsymbol{s}(t) + [\boldsymbol{n}_1^T(t)\boldsymbol{\Pi}, \boldsymbol{n}_2^T]^T, \tag{2.21}$$

where $\boldsymbol{\Pi} = [\boldsymbol{0}_{M\times 1}, \boldsymbol{J}_M]^T$ is an $(M + 1) \times M$ matrix with $\boldsymbol{J}_M$ being the $M \times M$ exchange matrix with ones on its anti-diagonal and zeros elsewhere. To simplify the

expressions, by rewriting $\boldsymbol{x}(t)$ as $\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{n}(t)$ with $\boldsymbol{A} \triangleq [\boldsymbol{A}_1^T\boldsymbol{\Pi}, \boldsymbol{A}_2^T]$, the whole array steering vectors can be given as

$$\boldsymbol{a}_k = [e^{-j(M-1)d_1\alpha_k}, \cdots, e^{-jd_1\alpha_k}, 1, e^{-jd_2\alpha_k}, \cdots, e^{-j(M-1)d_2\alpha_k}]^T. \tag{2.22}$$

From (2.22), it is easy to verify that there exists some special relationship between $\boldsymbol{a}_k(2M-p+2:2M+1)$ and $\boldsymbol{a}_k(1:p)$, i.e., $\boldsymbol{J}_p\boldsymbol{a}_k(2M-p+2:2M+1) = \boldsymbol{\Psi}_k^p\boldsymbol{a}_k(1:p)$, where

$$\boldsymbol{\Psi}_k^p = \begin{cases} diag(e^{-j(M-1)d\alpha_k}, \cdots, e^{-j(M-p)d\alpha_k}) & 1 \leqslant p \leqslant M \\ diag(e^{-j(M-1)d\alpha_k}, \cdots, 1, \cdots, e^{j(p-M)d\alpha_k}) & p \geq M+1 \end{cases} \tag{2.23}$$

is a $p \times p$ diagonal matrix. It is worth-noting that we can avoid the phase ambiguity with DOA information for the diagonal element only if $d \leq 1$.

2.3.2 DOA Estimation Technique

Differing from the DOA estimation method for the MRSSA, here the proposed method can be utilized in the case of correlated sources. Let us consider the auto-covariance matrix of $\boldsymbol{x}(t)$, namely,

$$\boldsymbol{R} = E[\boldsymbol{x}(t)\boldsymbol{x}^H(t)] = \boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H + \sigma^2\boldsymbol{I}_{2M+1}. \tag{2.24}$$

In a manner similar to that used in obtaining (2.8) and (2.9), we have $\boldsymbol{U}_s = \boldsymbol{AT}$. According to the array configuration described above, we can get the following two signal subspaces

$$\begin{aligned} \boldsymbol{U}_1 &\triangleq \boldsymbol{U}_s(1:p,:) = \boldsymbol{A}(1:p,:)\boldsymbol{T} \\ \boldsymbol{U}_2 &\triangleq \boldsymbol{U}_s(2M-p+2:2M+1,:) = \boldsymbol{A}(2M-p+2:2M+1,:)\boldsymbol{T} \end{aligned} . \tag{2.25}$$

Similar to the generalized ESPRIT algorithm suggested in [79], we can define the following matrix

$$\boldsymbol{J}_p\boldsymbol{U}_2 - \boldsymbol{\Phi}_p(\theta)\boldsymbol{U}_1 = (\boldsymbol{J}\boldsymbol{A}(2M-p+2:2M+1,:) - \boldsymbol{\Phi}_p(\theta)\boldsymbol{A}(1:p,:))\boldsymbol{T} \tag{2.26}$$

where $\boldsymbol{\Phi}_p(\theta)$ has the same expression as (2.23). It is easy to prove that both sides of (2.26) would no longer be a full column rank matrix when $\theta = \theta_k$. That is to say, by left-multiplying $\boldsymbol{J}_p\boldsymbol{U}_2 - \boldsymbol{\Phi}_p(\theta)\boldsymbol{U}_1$ with any $K \times p$ full row rank matrix $\boldsymbol{D}$, we can find that the matrix $\boldsymbol{D}(\boldsymbol{J}_p\boldsymbol{U}_2 - \boldsymbol{\Phi}_p(\theta)\boldsymbol{U}_1)$ would be singular for $\theta = \theta_k, (k = 1, \cdots, K)$, while nonsingular for other angles.

Therefore, we can implement the following spectral function to find the rough DOAs without ambiguity

$$F_r(\theta) = \frac{1}{[\det(\boldsymbol{D}(\boldsymbol{J}_p\boldsymbol{U}_2 - \boldsymbol{\Phi}_p(\theta)\boldsymbol{U}_1))]^2} \tag{2.27}$$

We will employ the generalized ESPRIT method to estimate the rough DOA based on our proposed array geometry to deal with the problem of the multi-value ambiguity, which may occur in implementing the traditional ESPRIT for sparse ULAs

directly. We will then introduce a novel alternating null-steering technique (ANST) to estimate the accurate DOA based on the estimated rough DOA in (2.27).

Now, let us define the following matrix

$$\boldsymbol{B}(\boldsymbol{\theta}_{(k)}, \theta) = [\boldsymbol{A}_{(k)} \quad \boldsymbol{w}(\theta)], \tag{2.28}$$

where $\boldsymbol{B}(\boldsymbol{\theta}_{(k)}, \theta)$ is constructed using $\boldsymbol{A}_{(k)}$, the array steering matrix including all the DOAs in (2.27) except for the k^{th} DOA, and the steering vector $\boldsymbol{w}(\theta)$, where θ is an arbitrary angle with the k^{th} angle selection set $[\theta_k - \Delta\theta \quad \theta_k + \Delta\theta]$, $\Delta\theta$ being dependent on the cyclically ambiguous period; for example, $\Delta\theta = \pi/12$ for the case in Fig.2.2(b), and $\boldsymbol{w}(\theta)$ is the corresponding steering vector of θ. Following the idea of the projection matrix decomposition technique suggested in [80], we can obtain

$$P_{\boldsymbol{B}(\boldsymbol{\theta}_{(k)},\theta)} = P_{\boldsymbol{A}_{(k)}} + \boldsymbol{w}_P(\theta)\boldsymbol{w}_P^H(\theta)/(\boldsymbol{w}_P^H(\theta)\boldsymbol{w}_P(\theta)), \tag{2.29}$$

where $\boldsymbol{w}_P(\theta) \triangleq P^{\perp}_{\boldsymbol{A}_{(k)}}\boldsymbol{w}(\theta)$ is the projection steering vector on the orthogonal subspace of $P_{\boldsymbol{A}_{(k)}}$. Note that the goal for implementing such a projection decomposition is to construct two orthogonal subspaces: one is formed from the previously estimated DOAs except for the k^{th} DOA and the other is constructed by the updated fine DOA of the k^{th} signal source. From (2.29) it is easy to obtain the relationship

$$P_{\boldsymbol{B}(\boldsymbol{\theta}_{(k)},\theta)}\boldsymbol{w}_P(\theta) = \boldsymbol{w}_P(\theta), \tag{2.30}$$

which implies that $\boldsymbol{w}_P(\theta)$ belongs to the projection subspace of the steering vectors

$\boldsymbol{B}(\boldsymbol{\theta}_{(k)}, \theta)$.

Theorem 1. *Assume that there does not exist ambiguities for* θ_k *in the* k^{th} *angle selection set,that is,* $\boldsymbol{w}_P(\theta_a) \neq \boldsymbol{w}_P(\theta_b), a \neq b$. *Then the function* $L(\boldsymbol{\theta}_{(k)}, \theta) \triangleq \boldsymbol{w}_P^H(\theta)(\boldsymbol{I}_{2M+1} - P_{\boldsymbol{U}_s})\boldsymbol{w}_P(\theta) = 0$ *have a unique nontrivial solution* $\theta = \theta_k$,*i.e.,*$\boldsymbol{w}_P(\theta) \neq \boldsymbol{0}$, *where* $\boldsymbol{U}_s$ *is the signal subspace of (2.24).*

Proof: From (2.9) and (2.24), we can obtain the relationship $\boldsymbol{U}_s = \boldsymbol{AT} = \boldsymbol{B}(\boldsymbol{\theta}_{(k)}, \theta_k)\boldsymbol{T}$. Therefore,

$$\begin{aligned} L(\boldsymbol{\theta}_{(k)}, \theta) &\triangleq \boldsymbol{w}_P^H(\theta)(\boldsymbol{I}_{2M+1} - P_{\boldsymbol{U}_s})\boldsymbol{w}_P(\theta) \\ &= \boldsymbol{w}_P^H(\theta)(\boldsymbol{I}_{2M+1} - \boldsymbol{B}(\boldsymbol{\theta}_{(k)}, \theta_k)\boldsymbol{T} \\ &\quad (\boldsymbol{T}^H\boldsymbol{B}^H(\boldsymbol{\theta}_{(k)}, \theta_k)\boldsymbol{B}(\boldsymbol{\theta}_{(k)}, \theta_k)\boldsymbol{T})^{-1}\boldsymbol{T}^H\boldsymbol{B}^H(\boldsymbol{\theta}_{(k)}, \theta_k))\boldsymbol{w}_P(\theta) \\ &= \boldsymbol{w}_P^H(\theta)(\boldsymbol{I}_{2M+1} - P_{\boldsymbol{B}(\boldsymbol{\theta}_{(k)}, \theta_k)})\boldsymbol{w}_P(\theta) \end{aligned} . \tag{2.31}$$

Clearly, when $\theta = \theta_k$, $L(\boldsymbol{\theta}_{(k)}, \theta_k)$ in (2.31) equals zero. Assume that there exists another solution for $L(\boldsymbol{\theta}_{(k)}, \theta) = 0$, say θ_x, then exploiting (2.29) and (2.30) we can obtain the relationship that $\|\boldsymbol{w}_P(\theta_k)\|^2\|\boldsymbol{w}_P(\theta_x)\|^2 = |\boldsymbol{w}_P^H(\theta_k)\boldsymbol{w}_P(\theta_x)|^2$. Obviously, the equation holds only if $\theta_x = \theta_k$. □

Note that to avoid a trivial solution, we can get rid of the DOA value in advance because this DOA can be considered as one of the ambiguous values in the set of $\boldsymbol{\theta}_{(k)}$ or we can implement the following criterion to avoid this phenomenon

$$\theta_k = \arg \max_{\theta \in [\theta_k - \Delta\theta \quad \theta_k + \Delta\theta]} \boldsymbol{w}_P^H(\theta) P_{\boldsymbol{U}_s} \boldsymbol{w}_P(\theta). \tag{2.32}$$

Therefore, the alternating null-steering technique can be summarized as follows:

1. Implement the generalized ESPRIT technique of (2.27) to estimate rough DOA estimates of the sources and denote them as $[\hat{\theta}_k^{(0)}]_{k=1}^K$

2. Find a set of angle selection intervals in which a fine search will be performed: $\bigcup_{k=1}^{K}[\hat{\theta}_k^{(0)} - \Delta\theta \quad \hat{\theta}_k^{(0)} + \Delta\theta]$ and then implement the following iterative steps.

3. For $k = 1 : K$, compute the following optimization criterion in the k^{th} angle selection set of the i^{th} iterations:

$$\hat{\theta}_k^{(i)} = \arg \min_{\theta \in [\hat{\theta}_k^{(i-1)} - \Delta\theta \quad \hat{\theta}_k^{(i-1)} + \Delta\theta]} \boldsymbol{w}_P^H(\theta)(\boldsymbol{I}_{2M+1} - P_{\boldsymbol{U}_s})\boldsymbol{w}_P(\theta) \tag{2.33}$$

or

$$\hat{\theta}_k^{(i)} = \arg \max_{\theta \in [\hat{\theta}_k^{(i-1)} - \Delta\theta \quad \hat{\theta}_k^{(i-1)} + \Delta\theta]} \boldsymbol{w}_P^H(\theta) P_{\boldsymbol{U}_s} \boldsymbol{w}_P(\theta) \tag{2.34}$$

end

4. Compare $[\hat{\theta}_k^{(i)}]_{k=1}^K$ with $[\hat{\theta}_k^{(i-1)}]_{k=1}^K$ by checking if $\epsilon \geq \sqrt{\frac{\sum_{k=1}^{K}(\hat{\theta}_k^{(i)} - \hat{\theta}_k^{(i-1)})^2}{K}}$, where ϵ is set by user to get the accurate estimates. If the inequality holds, terminate the iteration and output the final DOA estimates. Otherwise set $i = 1 + 1$ and update the selection set $\bigcup_{k=1}^{K}[\hat{\theta}_k^{(i)} - \Delta\theta, \hat{\theta}_k^{(i)} + \Delta\theta]$, and go back to step 3.

2.3.3 Simulation Results for ANST

This subsection presents two examples to illustrate the performance of the proposed DOA estimation method in comparison with the MUSIC technique in terms of both accuracy and resolution of the estimated angle. In order to obtain the maximun

capability of estimating the number of sources, herein, we set $d = 1$ and $p = 2M + 1$. We use 9 elements in the array system, i.e., $M = 4$, for the first example, and 5 elements, i.e., $M = 2$ for the second example. The two ULAs have elements spaced by $d_2 = 8$ and $d_1 = 7$, respectively. Monte Carlo simulations with 800 independent realizations of the received data are carried out.

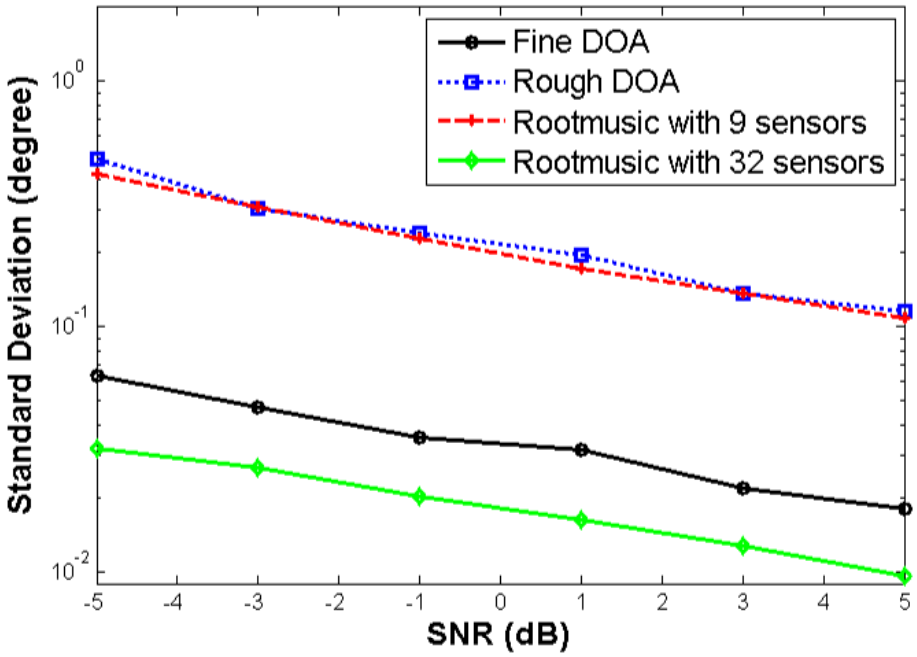


Figure 2.6: Two sources at the DOAs $[10°, 20°]$ impinge on a NULA system with two ULAs for the larger intersubarray of 8 half wavelength. 200 snapshots used.

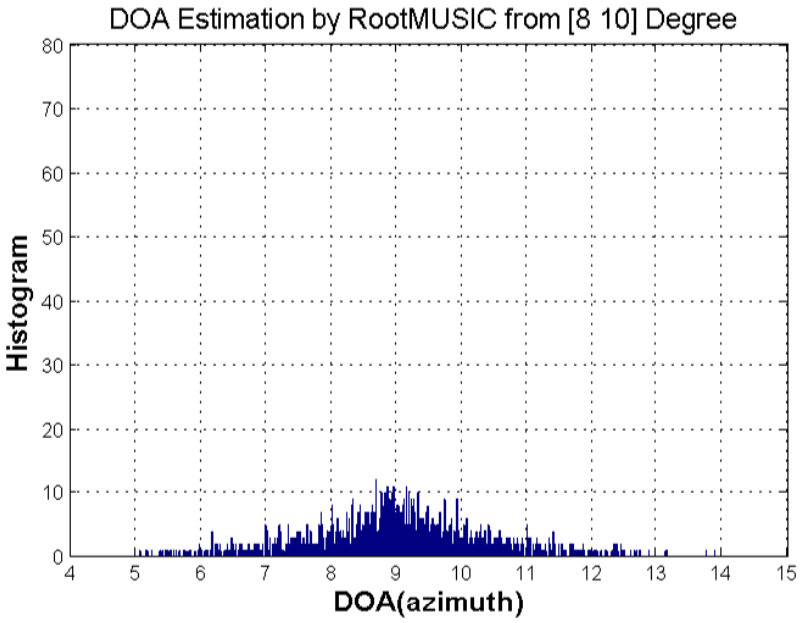
In the first example, we evaluate the DOA performance of the proposed and the MUSIC methods baseds on 9-sensor ULA and 32-sensor ULA with two equal-power signals whose incident directions are $[\theta_1, \theta_2] = [10°, 20°]$ and the SNR varies from -5 to 5dB. It is observed from Fig.2.6 that the proposed method provides a better accuracy than the traditional RootMUSIC method, using the same sensors. This is perhaps

because our proposed method makes use of the array pattern to largely extend the array aperture and implements the estimation of both rough and accurate angles. It is also found that the estimation accuracy of our proposed method is a little worse than that of the RootMUSIC with the same array aperture. This is because the latter exploits more sensors to reduce the additive noise.

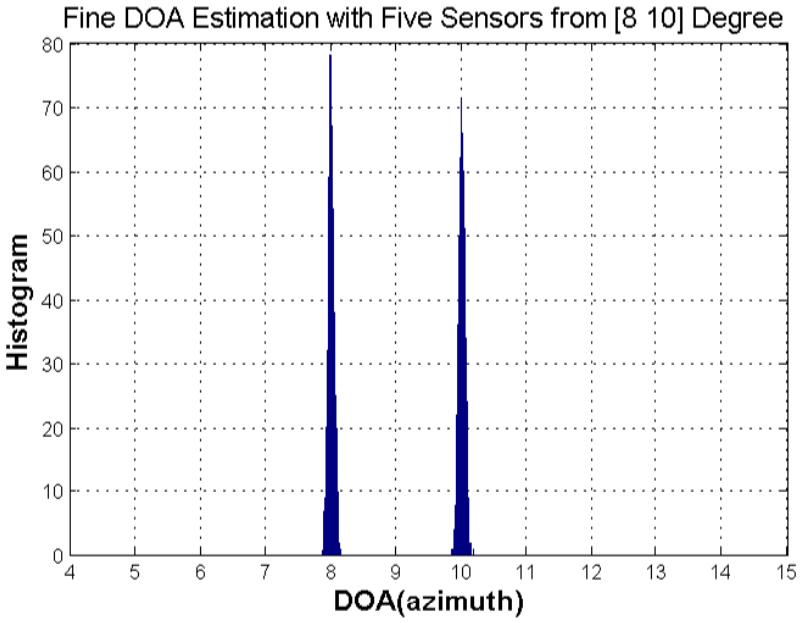
In the second example, we demonstrate the resolution of the two DOA estimation methods. In order to compare them under the same conditions, we assume that there are two equal power uncorrelated signals from DOAs of $[\theta_1, \theta_2] = [8°, 10°]$, and the SNR and the number of snapshots are set to 15dB and 200, respectively. The results are shown in Fig.2.7(a) and 2.7(b). From Fig.2.7(a), we see that the ULA using Root-MUSIC gives only one small peak around two arrival angles of the signals, i.e., two peaks merged into one peak, while Fig.2.7(b) shows clearly two peaks without any ambiguity produced by the proposed method. This means that among most of trials, the ULA-based root-MUSIC algorithm fails to give an accurate DOA estimation or resolve two close DOAs. Nevertheless, our proposed method can provide a more accurate DOA estimate and a higher resolution than the traditional methods such as MUSIC.

2.4 Joint Elevation and Azimuth Angle Estimation Based on Sparse L-shaped Array

In this section, we introduce a novel estimator for joint elevation and azimuth angle estimation based on sparse L-shaped array in the $x-z$ plane as shown in Fig.2.8, which



(a)



(b)

Figure 2.7: Histogram of two uncorrelated sources with DOAs of 8° and 10°, and the SNR=15dB, and 200 snapshots (a) ULA by Root-MUSIC, (b) Proposed technique.

is constructed by one ULA with half-wavelength inter-elements along the z axis and one SLA along the x axis as shown in Fig.2.9, where $d_1, d_2, \cdots, d_{M_x}$ are inter-element spacings in terms of the half wavelength and at least one of them is less than or equal to unity to solve the problem of ambiguity. Note that the element placed at the origin is a reference sensor. The proposed method estimates the elevation angles based on the signal subspace formed by a linear operation of the matrix from the cross-correlations between sensor data, and the array geometry and shift invariance property. Then, the azimuth angles can be estimated on the estimated signal waveforms obtained as a linear combination of the array outputs in which the weights are computed from the estimated elevation angles.

2.4.1 Proposed Array Model

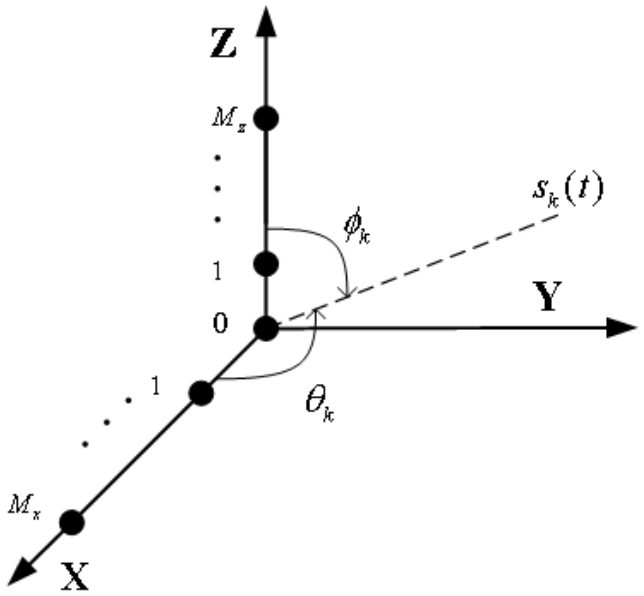


Figure 2.8: Array configuration for 2-D DOA estimation

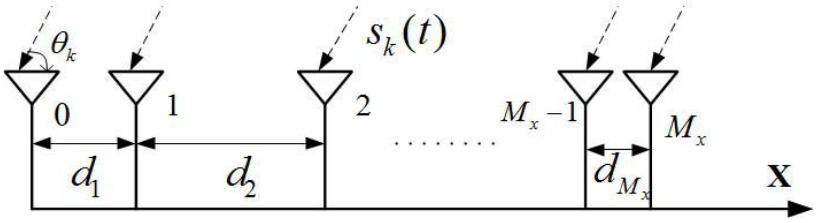


Figure 2.9: Geometry of sparse linear array

Suppose that there are K narrowband sources with the same wavelength impinging on the array from different azimuth and elevation directions, i.e., $[\theta_k \quad \phi_k]_{k=1}^{K}$. These sources are assumed to be in the far-field with respect to the sensor location. The observed signals at the ULA along the z axis and the SLA along the x axis are given by

$$\begin{aligned} \boldsymbol{z}(t) &= \boldsymbol{A}_z \boldsymbol{s}(t) + \boldsymbol{n}_z(t) \\ \boldsymbol{x}(t) &= \boldsymbol{A}_x \boldsymbol{s}(t) + \boldsymbol{n}_x(t) \end{aligned} . \tag{2.35}$$

The matrices and vectors in (2.35) have the following definitions. Both of the observed signals $\boldsymbol{z}(t) = [z_0(t), z_1(t), \cdots, z_{M_z}(t)]^T$ and $\boldsymbol{x}(t) = [x_0(t), x_1(t), \cdots, x_{M_x}(t)]^T$ are $(M_z+1)\times 1$ and $(M_x+1)\times 1$ vectors and functions of the snapshot t, respectively. $\boldsymbol{n}_z(t) = [n_{z,0}(t), n_{z,1}, \cdots, n_{z,M_z}]$ and $\boldsymbol{n}_x(t) = [n_{x,0}(t), n_{x,1}, \cdots, n_{x,M_x}]$ are i.i.d additive noise vectors, whose elements have zero mean and variance σ^2. Note that $x_0(t) = z_0(t)$ and $n_{x,0}(t) = n_{z,0}(t)$ denote the measurement and noise of the reference sensor, respectively. $\boldsymbol{A}_z$ has the same expression as $\boldsymbol{A}_1$ in (2.2) except that $\sin\theta = \cos\phi$, $d_1 = 1$, and $(M_u - 1) = M_z$. $\boldsymbol{A}_x$ is the $(M_x+1)\times K$ array manifold matrix of SLA including

the reference sensor as given by

$$\begin{aligned}\boldsymbol{A}_x &= [\boldsymbol{a}_x(\theta_1), \boldsymbol{a}_x(\theta_2), \cdots, \boldsymbol{a}_x\theta_K)] \\ &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\varphi_1^{(1)}} & e^{-j\varphi_2^{(1)}} & \cdots & e^{-j\varphi_K^{(1)}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j\varphi_1^{M_x}} & e^{-j\varphi_2^{M_x}} & \cdots & e^{-j\varphi_K^{M_x}} \end{bmatrix},\end{aligned} \tag{2.36}$$

where $\varphi_k^{(m)} = \sum_{i=1}^{m} d_i\psi_k, (m = 1, 2, \cdots, M_x)$ with $\psi_k = \pi\cos\theta_k$ is the phase difference between the m^{th} sensor of the x axis and the reference sensor for the $k^{th}(1, 2, \cdots, K)$ signal.

2.4.2 Elevation Angle Estimation

As is well known, the CCM, a noise-free correlation matrix [40, 45, 46], can be used to improve the estimation performance and reduce the computational complexity. Therefore, we will exploit the CCM to estimate the elevation angles by constructing the signal subspace without EVD and implementing ESPRIT-like technique. Now let us define the correlation vectors $\boldsymbol{r}_{z,x}^{(m)}, m = 1, 2, \cdots, M_x$ between $\boldsymbol{z}(t)$ and the m^{th} element $x_m(t)$ by

$$\boldsymbol{r}_{z,x}^{(m)} = E[\boldsymbol{z}(t)x_m^*(t)] = \boldsymbol{A}_z\boldsymbol{R}_s\boldsymbol{b}_m^*. \tag{2.37}$$

where $\boldsymbol{b}_m \triangleq \boldsymbol{A}_x(m+1, :)$ is the $(m+1)^{th}$ row of $\boldsymbol{A}$. Note that under the assumptions for the signals and the additive noises above, we can show that the correlation vectors

in (2.37) are not affected by the additive noises. Also, it can be viewed as a linear combination of the columns of the array response matrix $\boldsymbol{A}_z$, By concatenating the correlation vectors of $\boldsymbol{r}_{z,x}^{(m)}$, we construct the CCM

$$\boldsymbol{R}_{z,x}^{(m)} = [\boldsymbol{r}_{z,x}^{(1)}\boldsymbol{r}_{z,x}^{(2)}, \cdots, \boldsymbol{r}_{z,x}^{(M_x)}] = \boldsymbol{A}_z\boldsymbol{R}_s\boldsymbol{B}^*. \tag{2.38}$$

where $\boldsymbol{B} \triangleq [\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_{M_x}]$. Note that this matrix defines the cross correlation between the received data for the ULA including the common element and for the SLA along the x axis. If we assume that $\boldsymbol{B}$ is a row full rank matrix, we can extract any K columns, say the first K columns, from $\boldsymbol{R}_{z,x}^{(m)}$ to compose a new full column rank matrix named the signal subspace $\boldsymbol{U}_z$ with $\text{Rank}(\boldsymbol{U}_z) = \text{Rank}(\boldsymbol{A}_z)$. It can be seen that each column of $\boldsymbol{U}_z$ is a combination of rows of $\boldsymbol{A}_z$. Hence, following the same principle in (2.9), there exists a nonsingular $K \times K$ linear transform matrix $\boldsymbol{T}$ that satisfies the relation

$$\boldsymbol{U}_z = \boldsymbol{A}_z\boldsymbol{T}. \tag{2.39}$$

Since we have assumed that the subarray along z axis is ULA, we can obtain the elevation angles by employing the ESPRIT-like method instead of the MUSIC-like method without searching the parameter spaces. Therefore, we can use the same steps from (2.11) to (2.14) to estimate the elevation angle ϕ_k, i.e.,

$$\hat{\phi}_k = \arccos[-\arg(\lambda_{z,k})/\pi], k = 1, \cdots, K, \tag{2.40}$$

where $\lambda_{z,k}$ is the k^{th} eigenvalue of the matrix $\boldsymbol{F} \triangleq \boldsymbol{U}_z^{\dagger}(1 : M_z, :)\boldsymbol{U}_z(2 : M_z + 1, :)$.

2.4.3 Azimuth Angle Estimation

It is well known that under the model represented by (2.35), the estimates of the signal azimuth angles can be obtained by implementing the same technique as for the elevation angles above without explicitly estimating the signal waveform. However, there exist two shortcomings: one is the pair-matching problem and the other is computational burden. Here, we introduce a simple TLS-based method to estimate the azimuth angles on the estimated signal waveform from the ULA. Let us denote the matrix $\boldsymbol{A}_z$, evaluated using the estimated elevation angles of (2.40) by $\hat{\boldsymbol{A}}_z$. The estimate of $\boldsymbol{s}(t)$ is then given by

$$\hat{\boldsymbol{s}}(t) = \arg\min_{\boldsymbol{s}(t)} \sum_{t=1}^{N} \|\boldsymbol{z}(t) - \hat{\boldsymbol{A}}_z \boldsymbol{s}(t)\|^2, \tag{2.41}$$

where N is the number of snapshots. Thus, according to the assumption for the additive noise, the desired waveform estimate is given by

$$\hat{\boldsymbol{s}}(t) = [\hat{\boldsymbol{A}}_z^H \hat{\boldsymbol{A}}_z]^{-1} \hat{\boldsymbol{A}}_z^H \boldsymbol{z}(t). \tag{2.42}$$

Note that the estimated elevation angles in (2.40) and the estimated signal waveforms in (2.42) satisfy the one-to-one relationship. Therefore, the choice for constructing the $\boldsymbol{A}_z$ does not affect the generality of the azimuth estimates. For a perfect estimation of the elevation angles, (2.42) yields

$$\hat{\boldsymbol{s}}(t) = [\boldsymbol{A}_z^H \boldsymbol{A}_z]^{-1} \boldsymbol{A}_z^H \boldsymbol{z}(t) = \boldsymbol{s}(t) + [\boldsymbol{A}_z^H \boldsymbol{A}_z]^{-1} \boldsymbol{A}_z^H \boldsymbol{n}_z(t) = \boldsymbol{s}(t) + \boldsymbol{n}_s(t)\cdot \tag{2.43}$$

where $\boldsymbol{n}_s(t)$ with zero mean and covariance matrix $\boldsymbol{C}_{n_s} = \sigma^2[\boldsymbol{A}_z^H \boldsymbol{A}_z]^{-1}$ is the estimation error of signal waveforms due to the effect of the additive noise along z axis. Now, let us consider the output of each sensor in the x axis as a linear combination of the signal waveforms in (2.43), that is,

$$x_m(t) = \hat{\boldsymbol{s}}^T(t)\boldsymbol{b}_m = \boldsymbol{s}^T(t)\boldsymbol{b}_m + n_{x,m}(t) \quad t = 1, 2, \cdots, N. \tag{2.44}$$

Clearly, when the signal waveforms are known or estimated by noise-free array measurements, i.e., $\boldsymbol{n}_s(t) = \boldsymbol{0}$, the K-dimensional vector $\boldsymbol{b}_m$ can be obtained as the least squares solution to the minimization problem $\min\limits_{\boldsymbol{b}_m} \sum\limits_{t=1}^{N} |\boldsymbol{s}^T(t)\boldsymbol{b}_m - x_m(t)|^2$, which is identical to the maximum-likelihood one [81]. In practice, however, the measurements are noisy and the estimation of the elevation angles is not perfect, $\boldsymbol{b}_{m_{\mathrm{LS}}} = (\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}})^{-1}\hat{\boldsymbol{S}}^H\boldsymbol{x}_m$, where $\hat{\boldsymbol{S}} = [\hat{\boldsymbol{s}}(1), \hat{\boldsymbol{s}}(2), \cdots, \hat{\boldsymbol{s}}(N)]^T$ and $\boldsymbol{x}_m = [x_m^T(1), x_m^T(2), \cdots, x_m^T(N)]^T$, is no longer optimal from a statistical point of view and it suffers from bias and increased covariance due to the accumulation of noise errors [82]. To deal with this problem, TLS [83, 84] solution, that is, $\boldsymbol{b}_{m_{\mathrm{TLS}}} = (\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}} - \sigma_{K+1}^2\boldsymbol{I}_K)^{-1}\hat{\boldsymbol{S}}^H\boldsymbol{x}_m$, where σ_{K+1}^2 is the smallest singular value of $[\hat{\boldsymbol{S}} \quad \boldsymbol{x}_m]$, can be used to obtain more consistent estimation by removing the noise in $\hat{\boldsymbol{S}}$. Obviously, the elements of estimation error $\boldsymbol{n}_s(t)$ in (2.43) are not i.i.d and zero mean,$\boldsymbol{b}_{m_{\mathrm{TLS}}}$ is still a biased estimation. Therefore, we modify the TLS solution as

$$\boldsymbol{b}_{m_{\mathrm{MTLS}}} = (\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}} - N\boldsymbol{C}_{n_s})^{-1}\hat{\boldsymbol{S}}^H\boldsymbol{x}_m \tag{2.45}$$

to obtain an unbiased estimate of $\boldsymbol{b}_m$ and its covariance is $\sigma^2\boldsymbol{R}_s^{-1}(\boldsymbol{R}_s + N\boldsymbol{C}_{n_s})\boldsymbol{R}_s^{-H}$(see Appendix A), where $\boldsymbol{R}_s \triangleq \boldsymbol{S}^H\boldsymbol{S}$. Note that the order of $\boldsymbol{b}_{m_{\mathrm{MTLS}}}$ is the same as that

of the estimated signal waveforms and the elevation angles; therefore, our method can obtain the paired azimuth and elevation angles without additional steps. Since we have obtained the entire $\boldsymbol{b}_m^*, (m = 1, 2, \cdots, M_x)$, each of which contains a set of DOAs, we can use any $\boldsymbol{b}_m^*$ to estimate the final DOA. However, there exists the problem of ambiguity when $\sum_{i=1}^{m} d_i > 1$; herein, we assume $d_1 \leq 1$ to obtain the rough DOA without ambiguity, and then using (2.17) to obtain $\hat{\varphi}_k^{(m)}$, i.e., the estimate of $\varphi_k^{(m)}$. Finally, the final θ_k can be estimated as

$$\hat{\theta}_k = \arccos\Big(\sum_{m=1}^{M} \hat{\varphi}_k^{(m)} \Big/ \pi \sum_{m=1}^{M} \sum_{i=1}^{m} d_i\Big). \tag{2.46}$$

2.4.4 Complexity Analysis

The implementation of the proposed method requires three major steps:

1. Computation of the cross-correlation matrices $\boldsymbol{R}_{z,x}^{(m)}$ to form the signal subspace $\boldsymbol{U}_z$ using (2.38) and (2.39).

2. Estimation of the elevation angle ϕ_k in the way similar to ESPRIT-like method.

3. Estimation of the source waveforms using (2.42) and estimate the azimuth angle based on TLS method using (4.2) and (2.46).

The number of flops needed to form $\boldsymbol{U}_z$ is $2N(M_z + 1)K$, since it requires approximately $2N(M_z + 1)$ flops to obtain each $\boldsymbol{r}_{z,x}^{(m)}$. The flop is defined as a floating-point addition/multiplication operation. According to the algorithm suggested in [85], the computation of $\boldsymbol{F}$ takes about $4M_zK^2 + o(K^3)$ flops. The calculation of $\hat{\boldsymbol{s}}(t)$ requires roughly $2N(M_z + 1)K + 4(M_z + 1)K^2 + o(K^3)$ flops and the entire $\{\boldsymbol{b}_m\}$ takes about

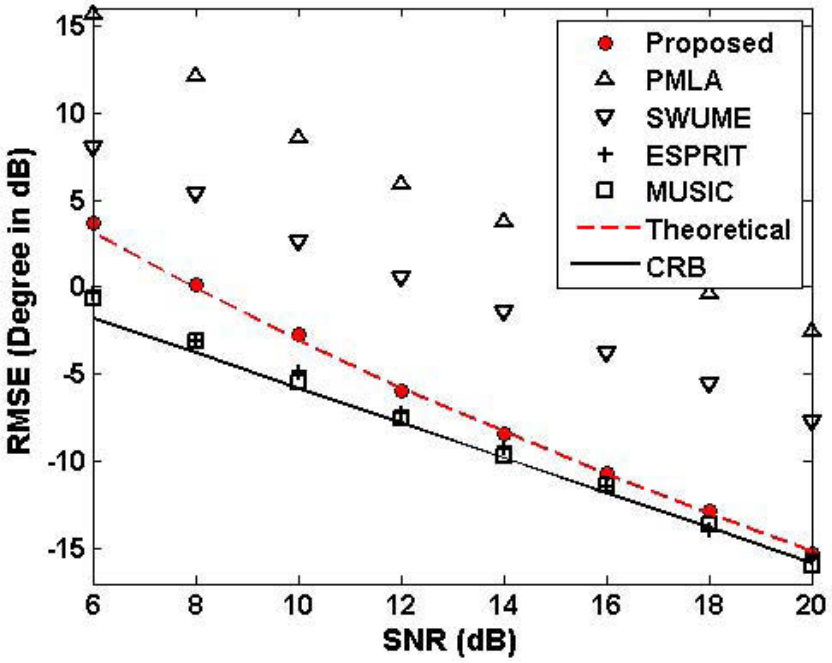
$4NK^2+o(K^3)+2NKM_x$ flops, respectively. Thus, the number of flops required by the proposed algorithm is roughly $2(M_x+M_z+1)NK$ in total when $N \gg M_z \geq M_x \gg K$, which occurs often in practical applications of DOA estimation. In comparison, the fast 2-D algorithm of PMLA suggested in [37] requires nearly $2(M_x + M_z + 2)NK$ flops and the Kikuchi algorithm requires another $(M_x+M_z+2)NK$ flops for the pair-matching process [39]. Obviously, although the SUWME does not need to compute the decomposition of the array covariance, it has also a very heavy computational load in finding the roots or searching the parameter spaces.

2.4.5 Simulation Results for 2-D DOA Estimation

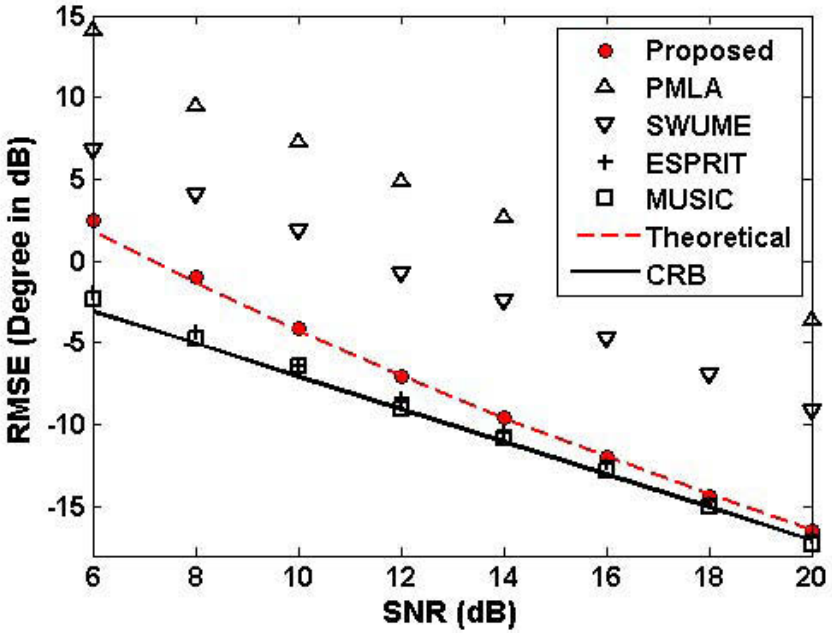
In this subsection, simulation results are presented to validate the proposed method and to illustrate its performance. In the first and second examples, the sensor displacement d between the adjacent elements in each uniform linear array is taken to be half the wavelength of the signal waveform. The total number of array elements is 9, i.e., $M_x = M_z = 4$. In the third example, an L-shaped array is constructed by a four-sensor ULA along the z-axis and a two-sensor SLA with the different array aperture along the x-axis, i.e., $M_x = 2, M_z = 4$. The fourth example is to show the performance with respect to the number of sensors of SLA along the x-axis. For all the tests, 1000 independent trials are carried out.

Example 1: *Performance of 2-D DOA estimation versus SNR*

In this example, the L-shaped array is constructed by two ULA each of which consists of four isotropic antennas spaced by half a wavelength. We consider two uncorrelated signals of 2-D DOA $[\theta_1 \quad \phi_1] = [45^\circ \quad 70^\circ]$ and $[\theta_2 \quad \phi_2] = [55^\circ \quad 80^\circ]$ with

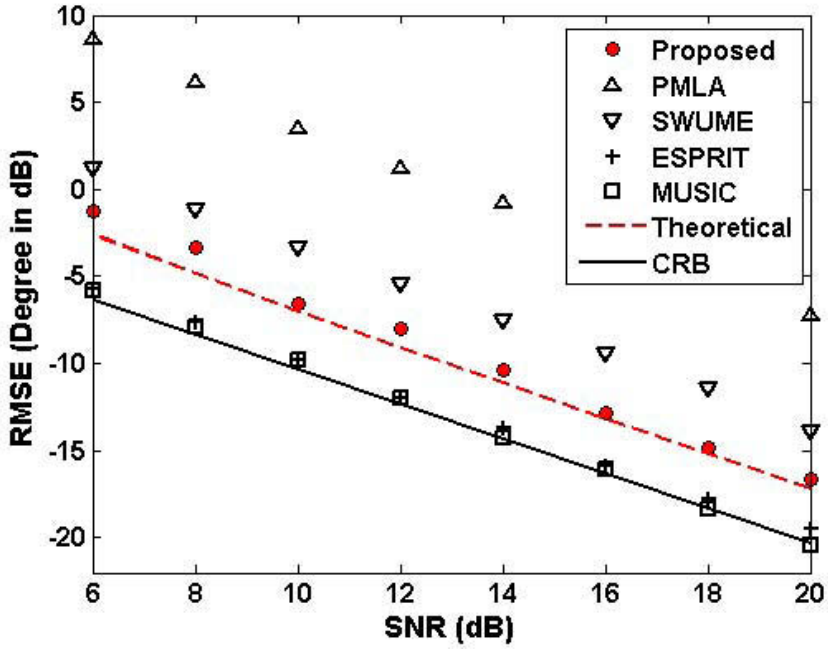


(a)

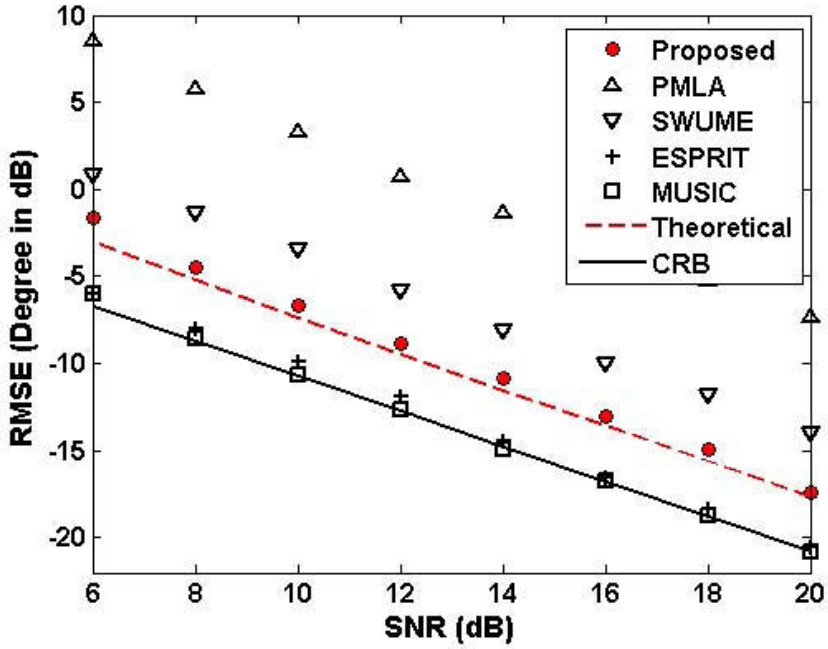


(b)

Figure 2.10: Elevation angle estimation performance with respect to SNR using L-shaped array constructed by two four-sensor ULAs for two sources with $[\phi_1, \phi_2] = [45°, 55°]$: (a) RMSE for $\phi_1 = 45°$ (b) RMSE for $\phi_2 = 55°$. The number of snapshots is 200.



(a)



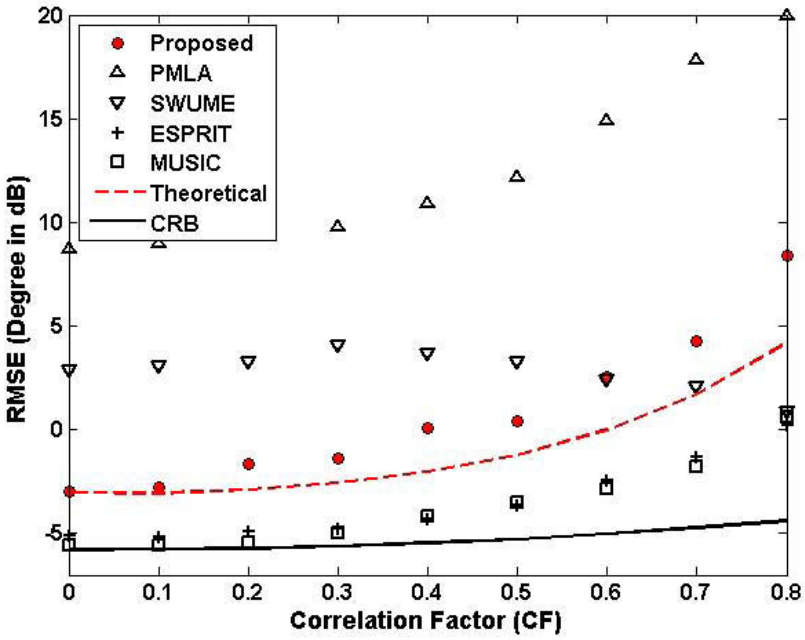
(b)

Figure 2.11: Azimuth angle estimation performance with respect to SNR using L-shaped array constructed by two four-sensor ULAs for two sources with $[\theta_1, \theta_2] = [70°, 80°]$: (a) RMSE for $\theta_1 = 70°$ (b) RMSE for $\theta_2 = 80°$. The number of snapshots is 200.

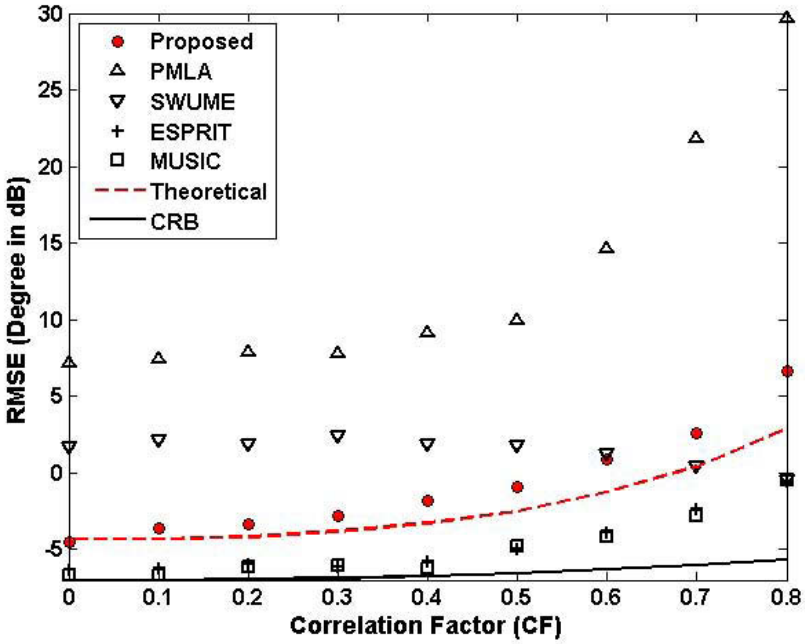
identical powers and white Gaussian additive noises. For performance comparison, we evaluate the proposed method against PMLA [37], SUMWE [44], ESPRIT, MUSIC, and the Cramer-Rao bound (CRB) [86]; these are shown in Fig.2.10 and Fig.2.11. Notice that the proposed method, PMLA, and SUMWE are computationally efficient algorithms that do not require the eigen-decomposition for the array covariance matrix, but both the PMLA and SUMWE have to deal with the pair-matching problem. In contrast, our proposed method takes advantage of the estimated elevation angles to obtain the pair matching information without additional procedures. The results indicate that the performance of the proposed method is better than that of the PMLA and SUMWE algorithms in both the elevation and the azimuth angle estimation, especially for medium and high levels of SNR. It is also shown that the estimated elevation angle from the proposed method is almost the same as that of the subspace methods based on EVD and SVD, which have a heavy computational burden (e.g., Root-MUSIC and ESPRIT), and its RMSE curve is also identical to the CRB at higher SNR. A possible reason for this is that our method exploits the noise-free CCM to obtain the signal subspace, and in this way the effect of noise can be reduced without implementing eigen-decomposition. From the figures, we also see that the RMSEs curve of the proposed method nearly coincide with the theoretical RMSEs for both the azimuth and elevation angles.

Example 2: *Performance of 2-D DOA estimation versus correlation factor*

This example uses a similar data model as in the first example except for the CF and SNR, herein we set SNR=10dB and the 2-D DOA estimation performance curves are plotted with respect to the CF from 0 to 0.8. From Fig.2.12 and 2.13, we can

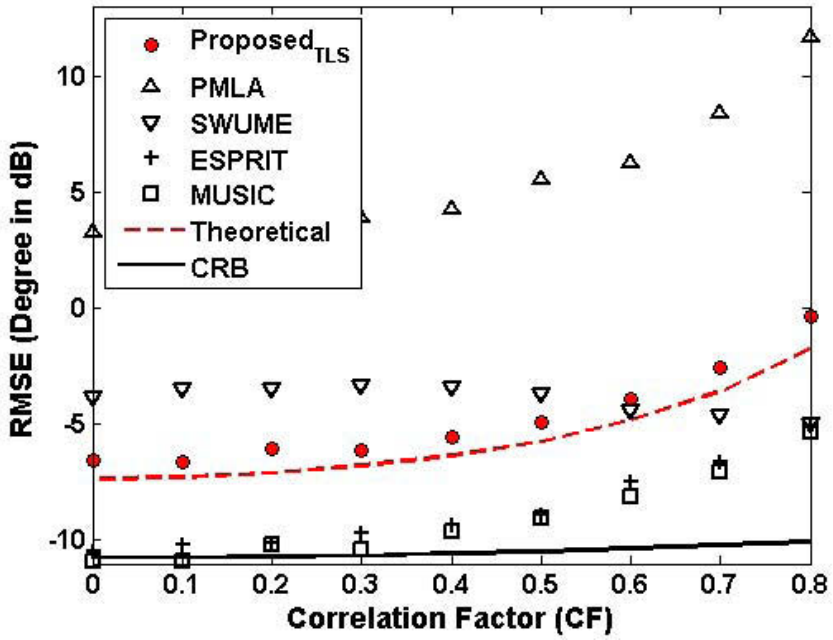


(a)

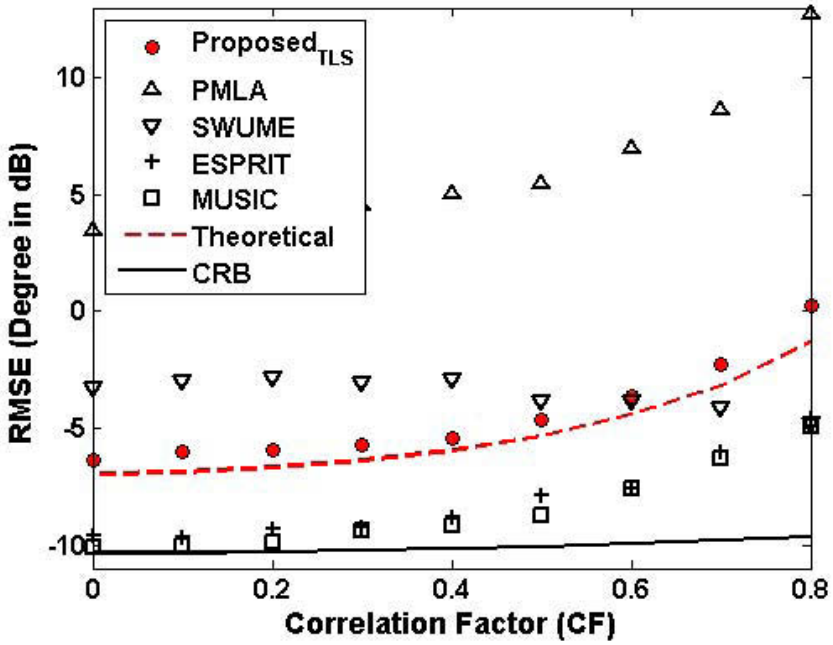


(b)

Figure 2.12: Elevation angle estimation performance with respect to correlation factor (CF) using L-shaped array constructed by two four-sensor ULAs for two sources with $[\phi_1, \phi_2] = [45°, 55°]$: (a) RMSE for $\phi_1 = 45°$ (b) RMSE for $\phi_2 = 55°$. The SNR is set to 10dB and the number of snapshots is 200.



(a)



(b)

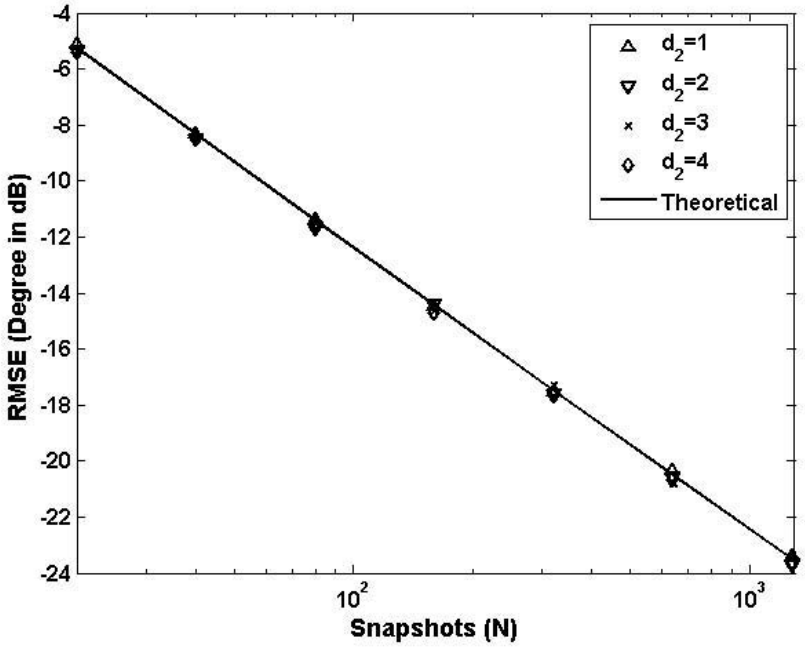
Figure 2.13: Azimuth angle estimation performance with respect to correlation factor (CF) using L-shaped array constructed by two four-sensor ULAs for two sources with $[\theta_1, \theta_2] = [70°, 80°]$: (a) RMSE for $\theta_1 = 70°$ (b) RMSE for $\theta_2 = 80°$. The SNR is set to 10dB and the number of snapshots is 200.

find that the proposed method has the best performance among the fast algorithms at low and media CF while the SUWME has the best estimation results at high CF due to the spatial smoothing technique. It can also be seen that the spatial smoothing technique is good for high CF or coherent signal sources but not for uncorrelated or low CF signal sources. The results shown in Fig.2.13(a) and 2.13(b) told us that the estimation performance can be improved significantly by reducing the effect of noise.

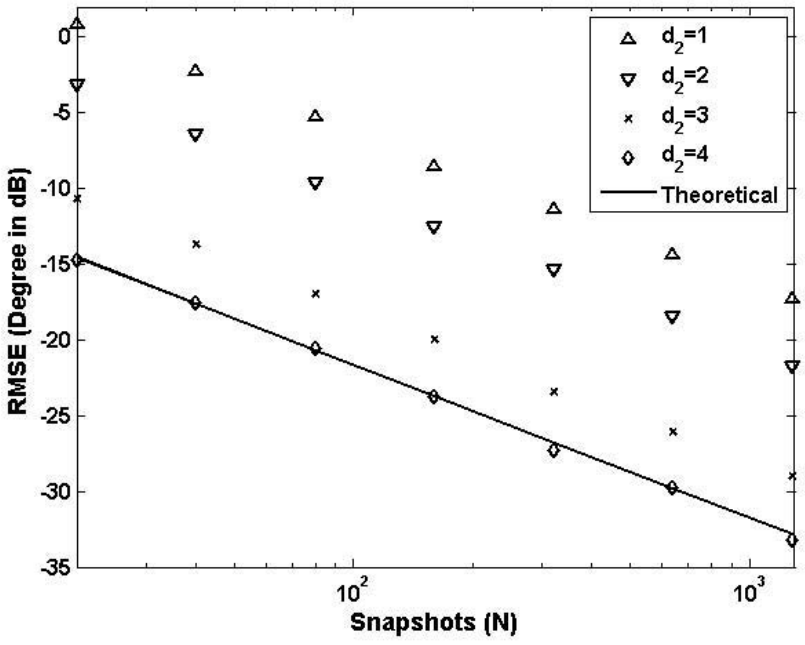
Example 3: *Performance of 2-D DOA estimation versus snapshots based on S-LA* In the third example, we consider two uncorrelated signals with the elevation angles $[\phi_1, \phi_2] = [80°, 90°]$ and azimuth angles $[\theta_1, \theta_2] = [85°, 105°]$ impinging into the L-shaped array with three sensors SLA along x-axis. Herein there are five cases with different array apertures, i.e. $d_2 = 1, 2, 3, 4, 7$, to be compared with respect to the number of snapshots varying from 20 to 1280 as shown in Fig.2.14(a) and 2.14(b).From the figures, we can find the same estimation performance for the elevation angles with the same number of sensors along x-axis while different azimuth estimation performances with different array apertures. Obviously, the larger the array aperture, the better performance is. Therefore, we can improve the azimuth performance significantly by designing SLA with the same number of sensors without loss of the elevation estimation performance.

Example 4: *Performance of 2-D DOA estimation versus SNR based on SLA*

This example shows the 2-D DOA performance with the same array aperture and different number of sensors along x-axis with respect to SNR ranging from 6dB to 20dB. The 3-sensor SLA with $d_1 = 1, d_2 = 5$, the 4-sensor SLA with $d_1 = d_3 = 1, d_2 = 4$, and the 5-sensor SLA with $d_1 = d_3 = d_4 = 1, d_2 = 4$ are adopted in this

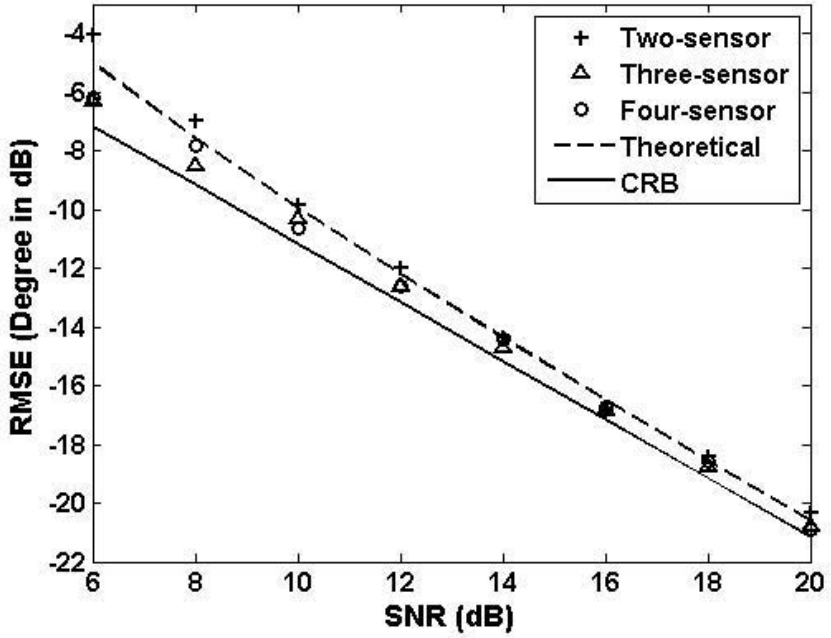


(a)

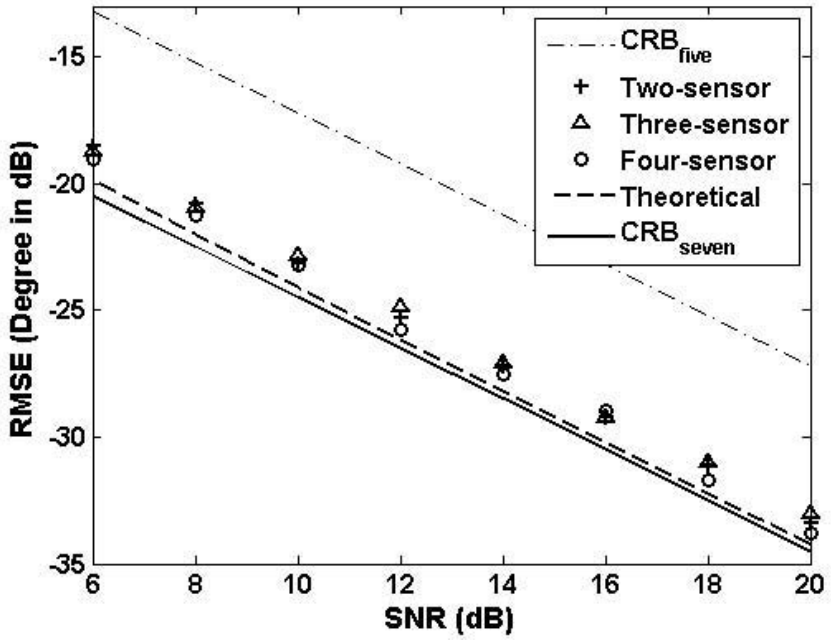


(b)

Figure 2.14: 2-D DOA estimation performance with respect to number of snapshots using L-shaped array constructed by a 4-sensor ULA along z-axis and a 2-sensor SLA along x-axis for two sources with $[\theta_1, \theta_2] = [85°, 105°]$ and $[\phi_1, \phi_2] = [80°, 90°]$: (a) RMSE for $\phi_2 = 90°$ (b) RMSE for $\theta_2 = 105°$. The SNR is set to 15dB and the number of snapshots is from 20 to 1280.



(a)



(b)

Figure 2.15: 2-D DOA estimation performance with respect to SNR using L-shaped array constructed by a 4-sensor ULA along z-axis and a SLA with 2 to 4 sensors along x-axis for two sources with $[\theta_1, \theta_2] = [85°, 105°]$ and $[\phi_1, \phi_2] = [80°, 90°]$: (a) RMSE for $\phi_2 = 90°$ (b) RMSE for $\theta_2 = 105°$. The SNR is from 6dB to 20dB and the number of snapshots is 200.

example. The number of snapshots is 200. Other conditions are the same as those in Example 3. Notice that we obtain the elevation angles by averaging the multiple estimation results using different signal subspace cross-correlation. For example, there are two signals and three sensors along x-axis which can be used to construct two different signal subspace, one is constructed by the first two sensors and the other by the last two sensors. The final estimate of the elevation angle is averaged by two results from different signal subspaces. Obviously, by this way we can enhance the performance of the elevation angel estimation and ultimately improve the azimuth estimation performance.

2.5 Conclusion

Two novel schemes to improve the estimation accuracy and the angle resolution for 1-D DOA estimation with sparse sensors have been presented. One is named minimum redundancy sparse subarray (MRSSA), which improves the performance of DOA estimation by designing multiple subarrays based on the principle of minimum redundancy linear array and eliminates the ambiguity by employing the idea of Kronecker steering vectors to obtain the one-to-one mapping for the rough angle and the fine angle. The other provides a new array design strategy for trading off the unambiguity, rough DOA and fine DOA estimates by designing two sparse uniform linear arrays with the minimum inter-sensor spacing less than half wavelength. Inspired by the idea of the generalized ESPRIT, we have obtained the rough DOA without ambiguity, and then designed the alternating null-steering technique (ANST) to select the true fine value and cancel the ambiguity set in the same angle section. Note that we

can only carry out our proposed method for uncorrelated sources using the MRSSA while the second proposed method can be used in the situation of correlated sources. Furthermore, we have extended the spare array to the case of 2-D DOA estimation for correlated sources. In this method, we have designed an L-shaped array constructed by one uniform linear array along the z-axis and a sparse linear array along the x-axis, and then an efficient and effective total least squares-based algorithm is proposed to estimate the azimuth and elevation angles without requiring pair-matching. Finally, simulation results for the MRSSA, ANST and 2-D DOA techniques justifying the effectiveness of the proposed method in comparison to an existing method have been presented.

Chapter 3

DOA Estimation and Tracking for AR Model-based Signals with Unknown Waveform

3.1 Introduction

In Chapter 2, we have studied accurate 1-D and 2-D DOA estimation techniques by meticulously designing sparse arrays. Obviously, these proposed configurations can be used to obtain excellent estimation results without incorporating temporal information. In some applications, such as speech processing and mobile communications, the signals can be described by autoregressive (AR) models. Hence, in this chapter, we first present a nontraditional approach for estimating and tracking the signal DOA using an array of sensors. The proposed method consists of two stages: first, the sources modeled by AR processes are estimated by the celebrated Kalman filter

and then, the QR-decomposition-based recursive least square (QRD-RLS) technique is employed to estimate the DOAs and AR coefficients in each observed time interval. The AR-modeled sources can provide useful temporal information to handle cases, where the number of sources is larger than the number of antennas. The symmetric array enables one to transfer a complex-valued nonlinear problem to a real-valued linear one, which can reduce the computational complexity. Moreover, we also propose a DOA estimation method for AR-modeled sources based on general SLA. Simulation results demonstrate the superior performance of the algorithm for estimating and tracking DOA under different scenarios.

3.2 DOA Estimation and Tracking Based on Symmetric Arrays

3.2.1 Signal and System Model

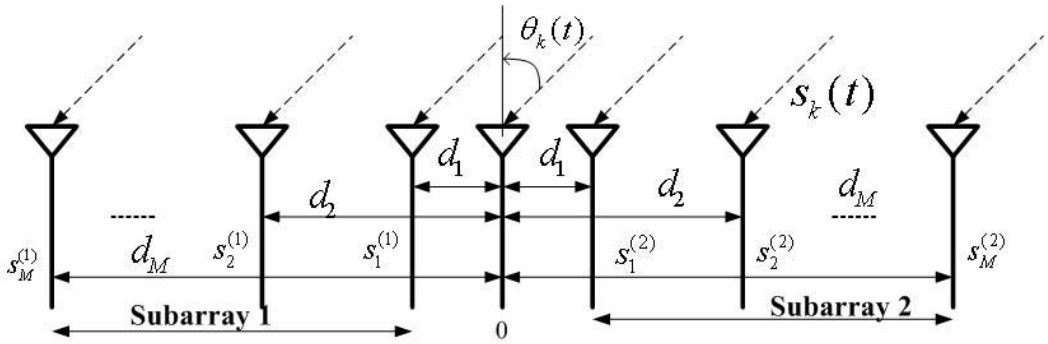


Figure 3.1: Structure of the symmetric linear array

Suppose that there are K narrowband moving sources with the same wavelength, which impinge onto a symmetric linear array as shown in Fig.3.1. The whole array is

assumed to be divided into two subarrays with inter-element spacings $d_1, d_2, \cdots, d_M$ in terms of half wavelength. In order to cope with the problem of ambiguity, we also assume $d_1 \leq 1$, and then the sensor element at the origin is used as reference. We assume that all the sources are independent order-L_k AR processes and the complex envelope at time t of the k^{th} AR source is

$$p_k(t) = \sum_{i=1}^{L_k} a_{k,i}(t) p_k(t-i) + v_k(t) \tag{3.1}$$

where $v_k(t)$ $(k = 1, 2, \cdots, K)$ is the excitation of the k^{th} AR process, which is a white Gaussian noise with zero mean and variance $\sigma^2_{v_k}(t)$, and $a_{k,i}(t)$ are real-valued AR coefficients [18, 87]. Here, we have assumed that the AR coefficients are real to simplify our discussion in sequel. Howvever, it is not difficult to extend this technique for complex AR coefficients. Eq.(3.1) can be rewritten as the following state-space representation,

$$\boldsymbol{p}_k(t) = \boldsymbol{F}_k(t)\boldsymbol{p}_k(t-1) + \boldsymbol{v}_k(t), \tag{3.2}$$

where $\boldsymbol{F}_k(t) = \left[\boldsymbol{a}_k^T(t) \quad \left[\boldsymbol{I}_{(L_k-1)} \quad \boldsymbol{0}_{1\times(L_k-1)}\right]^T\right]^T$ with $\boldsymbol{a}_k(t) = [a_{k,1}(t) \cdots a_{k,L_k}(t)]$, and $\boldsymbol{v}_k(t) = \left[v_k(t) \quad \boldsymbol{0}_{1\times(L_k-1)}\right]^T$. By using (3.2), the K AR sources can be written as the following state equation

$$\boldsymbol{p}(t) = \begin{bmatrix} \boldsymbol{p}_1(t) \\ \vdots \\ \boldsymbol{p}_K(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_1(t) & \boldsymbol{0} \\ & \ddots \\ \boldsymbol{0} & \boldsymbol{F}_K(t) \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_1(t-1) \\ \vdots \\ \boldsymbol{p}_K(t-1) \end{bmatrix} + \begin{bmatrix} \boldsymbol{v}_1(t) \\ \vdots \\ \boldsymbol{v}_K(t) \end{bmatrix} = \boldsymbol{F}(t)\boldsymbol{p}(t-1) + \boldsymbol{v}(t), \tag{3.3}$$

where $\boldsymbol{F}(t) \in \mathbb{R}^{(\sum_{k=1}^{K} L_k)\times(\sum_{k=1}^{K} L_k)}$ is a block diagonal matrix, and $\boldsymbol{p}(t)$ and $\boldsymbol{v}(t)$ are $\sum_{k=1}^{K} L_k$-dimensional complex vectors. Since $v_k(t)(k = 1, \cdots, K)$ are independent, the covariance matrix of $\boldsymbol{v}(t)$ is a block diagonal matrix, which can be expressed as $\boldsymbol{Q}(t) = diag\left(\boldsymbol{Q}_1(t), \cdots, \boldsymbol{Q}_K(t)\right)$. At time t, let $\theta_1, \theta_2, \cdots, \theta_K$ denote the DOAs of the K targets, and $x_m^{(i)}(t)$ denote the complex signals of the m^{th} sensor of the $i^{th}(i = 1, 2)$ subarray with $x_0(t)$ being the data received by the reference element. Then, the observed output complex signals of the $2M+1$ sensors, denoted as $\boldsymbol{x}(t) \in \mathbb{C}^{2M+1}$, can be written as

$$\begin{aligned} \boldsymbol{x}(t) &\triangleq [\boldsymbol{x}_1^T(t) x_0(t) \boldsymbol{x}_2^T(t)]^T \\ &= \begin{bmatrix} \boldsymbol{A}_1^T(t) & \mathbf{1} & \boldsymbol{A}_2^T(t) \end{bmatrix}^T \boldsymbol{s}(t) + \boldsymbol{e}(t) \\ &\triangleq \boldsymbol{A}(t)\boldsymbol{s}(t) + \boldsymbol{e}(t), \end{aligned} \tag{3.4}$$

where $\boldsymbol{x}_i(t) = [x_1^{(i)}(t), \cdots, x_M^{(i)}(t)]^T (i = 1, 2)$ is an M-dimensional vector of complex signals at the i^{th} subarray output, $\boldsymbol{s}(t) \triangleq [p_1(t), p_2(t), \cdots, p_K(t)]^T = \boldsymbol{\Gamma}\boldsymbol{p}(t)$ is a K-dimensional vector of the target sources with $\boldsymbol{\Gamma}$ being a $K \times \sum_{k=1}^{K} L_k$ dimensional selection matrix whose entries are 1 on the position $\left(k, \sum_{p=1}^{k} L_p - L_k + 1\right)$ and 0 elsewhere, and $\boldsymbol{e}(t) = [e_M^{(1)}(t), \cdots, e_1^{(1)}(t), e_0(t), e_1^{(2)}(t), \cdots, e_M^{(2)}(t)]^T$ is a $(2M + 1)$-dimensional vector of complex white measurement noises with zero-mean and covariance matrix $\boldsymbol{R}_{\boldsymbol{e}}(t)$. Finally, $\boldsymbol{A}$ is the $(2M - 1) \times K$ array manifold matrix of the whole symmetric array and $\boldsymbol{A}_i(t), (i = 1, 2)$ is the $M \times K$ array manifold matrix of the i^{th} subarray as

given by

$$\boldsymbol{A}_i(t) = \begin{bmatrix} e^{-(-1)^i j d_M \psi_1(t)} & e^{-(-1)^i j d_M \psi_2(t)} & \dots & e^{-(-1)^i j d_M \psi_K(t)} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-(-1)^i j d_1 \psi_1(t)} & e^{-(-1)^i j d_1 \psi_2(t)} & \dots & e^{-(-1)^i j d_1 \psi_K(t)} \end{bmatrix} \tag{3.5}$$

where $\psi_k(t) = \pi \sin\theta_k(t)$ and the arrangement of d_m in the reverse order of m is for convenience. It is easy to verify that $\boldsymbol{A}_1(t) = (\boldsymbol{A}_2(t))^*$. Therefore, the matrix $\boldsymbol{A}(t)$ in (3.4) can be rewritten as

$$\boldsymbol{A}(t) = \begin{bmatrix} \boldsymbol{A}_1^T(t) & \mathbf{1}_{1\times M}^T & \boldsymbol{J}_M \boldsymbol{A}_1^H(t) \end{bmatrix}^T. \tag{3.6}$$

It can be seen that each column of the matrix $\boldsymbol{A}(t)$ is conjugate symmetric. Note that rearranging $\boldsymbol{A}(t)$ and the array measurements does not affect the results of DOA estimation. In the following subsection, we will introduce a novel algorithm to estimate the waveforms of the signals of interest and their AR coefficients by employing the celebrated Kalman filter technique.

3.2.2 Kalman Filter and AR Coefficient Estimation

Herein we assume that both $\boldsymbol{\theta}(t)$ and the AR coefficients are slowly time-varying in the observed interval [71]. In particular, we assume that over each time interval the

change in both $\boldsymbol{\theta}(t)$ and $\boldsymbol{F}(t)$ is small enough so that

$$\boldsymbol{\theta}(t) \approx \boldsymbol{\theta}(nT), \quad \boldsymbol{a}_k(t) \approx \boldsymbol{a}_k(nT), t \in ((n-1)T, nT], (n = 1, 2, \cdots; k = 1, \cdots, K) \tag{3.7}$$

It is also assumed that there are N snapshots or signal samples available to process the received data and estimate the AR coefficients and DOAs over each interval $((n-1)T, nT]$. Consequently, the N snapshots in the n^{th} time interval can be approximately expressed as

$$\boldsymbol{x}(n,i)=\boldsymbol{A}(n)\boldsymbol{s}(n,i) + \boldsymbol{e}(n,i), (i = 1, \cdots, N) \tag{3.8}$$

Note that (3.8) can be considered as the discrete-time version of (3.4) in the n^{th} interval. To simplify the expressions, we use $\boldsymbol{x}(i), \boldsymbol{s}(i)$, and $\boldsymbol{e}(j)$ instead of $\boldsymbol{x}(n, j), \boldsymbol{s}(n, j)$, and $\boldsymbol{e}(n, j)$ in the sequel. As mentioned earlier, the proposed algorithm is to explore the dynamics of the source signals and jointly estimate the DOA and AR parameters of the source signals. To this end, we shall track the source state vector $\boldsymbol{p}(t)$ using KF given initial DOAs and then update the DOAs using a regularized QRD-RLS algorithm. In general, KF is an optimal minimum mean squares state estimator for a linear state space system [88–90], while the regularized QRD-RLS algorithm has good numerical stability and low estimation variance [91, 92]. The required state space model at the n^{th} interval can be obtained from the state dynamical equation in (3.3) and the sensor measurement equations in (3.8) as follows

$$\boldsymbol{p}(i)=\boldsymbol{F}(n)\boldsymbol{p}(i-1) + \boldsymbol{v}(i), \tag{3.9}$$

$$\boldsymbol{x}(i)=\boldsymbol{A}(n)\boldsymbol{\Gamma}\boldsymbol{p}(i)+\boldsymbol{e}(i), \tag{3.10}$$

where the system matrices $\boldsymbol{F}$ and $\boldsymbol{A}$ depend only on n and can be considered as constants in each interval, and $\boldsymbol{v}(i)$ and $\boldsymbol{e}(i)$ are assumed to be uncorrelated, i.e., $E[\boldsymbol{v}(p)\boldsymbol{e}^T(q)]=\boldsymbol{0}$, for all p and q. Then, the state of the system can be recursively estimated using the KF [89] as

$$\hat{\boldsymbol{p}}(i\,|i)=\hat{\boldsymbol{p}}(i\,|i-1)+\boldsymbol{\kappa}(i)(\boldsymbol{x}(i)-\boldsymbol{A}(n-1)\boldsymbol{\Gamma}\hat{\boldsymbol{p}}(i\,|i-1)) \tag{3.11}$$

where $\hat{\boldsymbol{p}}(i\,|i)$ is the update or posteriori estimate and $\hat{\boldsymbol{p}}(i\,|i-1)$ is the best estimate prior to assimilating the measurement $\boldsymbol{x}(i)$, and $\boldsymbol{\kappa}(i)$ is Kalman gain to be determined by the following Kalman recursion

$$P(i\,|i-1)=\boldsymbol{F}(n-1)P(i-1\,|i-1)\boldsymbol{F}^T(n-1)+\boldsymbol{Q}(i-1), \tag{3.12}$$

$$\boldsymbol{\kappa}(i)=P(i\,|i-1)\boldsymbol{\Gamma}^T\boldsymbol{A}^H(n-1)[\boldsymbol{A}(n-1)\boldsymbol{\Gamma}\,P(i\,|i-1)\boldsymbol{\Gamma}^T\boldsymbol{A}^H(n-1)+\boldsymbol{R}_{\boldsymbol{e}}(i)]^{-1}, \tag{3.13}$$

$$P(i\,|i)=(\boldsymbol{I}-\boldsymbol{\kappa}(i)\boldsymbol{A}(n-1)\boldsymbol{\Gamma})\,P(i\,|i-1) \tag{3.14}$$

where $P(i\,|i-1)$ and $P(i\,|i)$ are the error covariance matrices associated with the priori estimate $\hat{\boldsymbol{p}}(i\,|i-1)$ and the posteriori estimate $\hat{\boldsymbol{p}}(i\,|i)$, respectively. Since all states of the system have been estimated by (3.11), it is easy to separate the sources, giving the k^{th} AR source expressed in the linear regression model

$$p_k(i)=\boldsymbol{p}_k^T(i-1)\boldsymbol{a}_k(n)+v_k(i), \tag{3.15}$$

where $\boldsymbol{a}_k(n)$ are the AR coefficients for the n^{th} interval to be estimated. Since $\boldsymbol{a}_k(n)$ is real-valued while $p_k(i)$ and $v_k(i)$ are complex valued, we can transform (3.15) to a real-valued expression to reduce the computational load by separating the real and imaginary parts, leading to

$$\boldsymbol{z}_k(i) = \boldsymbol{B}_k^T(i)\boldsymbol{a}_k(n) + \boldsymbol{w}_k(i), \tag{3.16}$$

where

$$\boldsymbol{z}_k(i)= \begin{bmatrix} \mathrm{Re}[p_k(i)] \\ \mathrm{Im}[p_k(i)] \end{bmatrix}, \quad \boldsymbol{B}_k(i)= \begin{bmatrix} \mathrm{Re}[\boldsymbol{p}_k^T(i-1)] \\ \mathrm{Im}[\boldsymbol{p}_k^T(i-1)] \end{bmatrix}^T, \quad \boldsymbol{w}_k(i)= \begin{bmatrix} \mathrm{Re}\,[v_k(i)] \\ \mathrm{Im}\,[v_k(i)] \end{bmatrix}. \tag{3.17}$$

We now estimate $\boldsymbol{a}_k(n)$ by employing the regularized QRD-RLS algorithm described in Table 3.1. As compared to RLS, QRD-RLS has a better performance, lower computational complexity, as well as better numerical stability in finite word-length implementation. In the next section, we will consider the tracking and estimation of the DOAs.

3.2.3 DOA Estimation and Tracking

We implement the DOA estimation in the real-valued domain in view of its excellent accuracy and substantial reduction of the computational burden over conventional DOA estimation techniques. In particular, we derive a real-valued DOA estimation and tracking method based also on the regularized QRD-RLS algorithm. First of all,

Table 3.1: QRD-RLS Algorithm

Initialization:

For $k = 1, 2, \cdots, K$

$\mathbf{H}_k(0) = \delta \boldsymbol{I}$, , δ is a positive constant, say $\delta = 1$

$\boldsymbol{U}_k(0) = \mathbf{0}$.

end

Recursion:

For $n = 1, 2, \cdots$

For $i = 1, 2, \cdots, N$

Given $\mathbf{H}_k(i-1), \boldsymbol{U}_k(i-1), \boldsymbol{z}_k(i), \boldsymbol{B}_k(i)$

$$\boldsymbol{Q}_k(i) \begin{bmatrix} \sqrt{\beta}\mathbf{H}_k(i-1) & \sqrt{\beta}\boldsymbol{U}_k(i-1) \\ \boldsymbol{B}_k^T(i) & \boldsymbol{z}_k(i) \end{bmatrix} = \begin{bmatrix} \mathbf{H}_k(i) & \boldsymbol{U}_k(i) \\ \mathbf{0} & \tilde{\boldsymbol{\delta}}_k(i) \end{bmatrix}$$

where $\boldsymbol{Q}_k(i)$ is an unitary Givens rotation matrix of the i^{th} recursion for the k^{th} AR modeled source and β is the forgetting factor.

end

$\boldsymbol{a}_k(n) = (\mathbf{H}_k(N))^{-1}\boldsymbol{U}_k(N)$.

end

let us introduce an odd unitary transformation defined in [93]

$$\mathbf{T}_{2M+1} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I}_M & \mathbf{0}_{M\times 1} & j\boldsymbol{I}_M \\ \mathbf{0}_{1\times M} & \sqrt{2} & \mathbf{0}_{1\times M} \\ \boldsymbol{J}_M & \mathbf{0}_{M\times 1} & -j\boldsymbol{J}_M \end{bmatrix}. \tag{3.18}$$

By left-multiplying (3.10) with the above unitary transformation, one gets

$$\begin{aligned} \boldsymbol{x}_T(i) &\triangleq \mathbf{T}_{2M+1}^H \boldsymbol{A}(n)\boldsymbol{\Gamma}\boldsymbol{p}(i) + \mathbf{T}_{2M+1}^H \boldsymbol{e}(i) \\ &\triangleq [x_T^{-M}(i), x_T^{-M+1}(i), \cdots, \sum_{k=1}^{K} p_k(i), \cdots, x_T^{M-1}(i), x_T^{M}(i)]^T \end{aligned} \tag{3.19}$$

Let $\boldsymbol{A}_T(n) \triangleq \mathbf{T}_{2M+1}^H \boldsymbol{A}(n)$. It is easy to show that $\boldsymbol{A}_T(n)$ is a real-valued matrix as given by

$$\boldsymbol{A}_T(n) = \sqrt{2} \begin{bmatrix} \mathrm{Re}[\boldsymbol{A}_1(n)] \\ \mathbf{1}_{1\times M} \\ \mathrm{Im}[\boldsymbol{J}_M \boldsymbol{A}_1(n)] \end{bmatrix}, \tag{3.20}$$

where

$$\mathrm{Re}[\boldsymbol{A}_1(n)] = \begin{bmatrix} \cos(d_M\psi_1(n)) & \cos(d_M\psi_2(n)) & \cdots & \cos(d_M\psi_K(n)) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(d_1\psi_1(n)) & \cos(d_1\psi_2(n)) & \cdots & \cos(d_1\psi_K(n)) \end{bmatrix}$$

$$\mathrm{Im}[\boldsymbol{J}_M\boldsymbol{A}_1(n)] = \begin{bmatrix} \sin(d_1\psi_1(n)) & \sin(d_1\psi_2(n)) & \cdots & \sin(d_1\psi_K(n)) \\ \vdots & \vdots & \ddots & \vdots \\ \sin(d_M\psi_1(n)) & \sin(d_M\psi_2(n)) & \cdots & \sin(d_M\psi_K(n)) \end{bmatrix}$$

Using (3.19) and (3.20), we can find that each element of $\boldsymbol{x}_T(i)$ can be described by a linear regression model as

$$x_T^m(i) = (\boldsymbol{\Gamma}\boldsymbol{p}(i))^T(\boldsymbol{A}_T^m(n))^T + \boldsymbol{e}_T^m(i), (m{=}\pm 1, \pm 2, \cdots, \pm M), \tag{3.21}$$

where $\boldsymbol{A}_T^m(n)$ denotes the m^{th} row of matrix $\boldsymbol{A}_T(n)$. Note that the auto-covariance matrix of $\boldsymbol{e}_T^m(i)$ is the same as those of $\boldsymbol{e}^m(i)$ due to its unitary invariance. Similar to the estimation of the AR coefficients, we can again employ the QRD-RLS algorithm to estimate all the row vectors of $\boldsymbol{A}_T(n)$ except for the middle one, namely, $\boldsymbol{A}_T^m(n), (m{=}\pm 1, \pm 2, \cdots, \pm M)$, Further, we can combine $\boldsymbol{A}_T^{+m}(n)$ and $\boldsymbol{A}_T^{-m}(n)$ as one group, giving

$$\text{Group} \quad m: \quad \left\{ \begin{array}{l} \sin(d_m\psi_1(n)) \cdots \sin(d_m\psi_K(n)) \\ \cos(d_m\psi_1(n)) \cdots \cos(d_m\psi_K(n)) \end{array} \right\}. \tag{3.22}$$

Obviously, one could exploit either the sine or cosine vector of group 1 to obtain the DOA estimate, but this sine or cosine calculation may fail to obtain the DOA estimate when the argument of the inverse sine or cosine function is greater than unity. To overcome this limitation, we can combine the sine and cosine vectors to get a new set of tangent vector for each group, i.e.,

$$\text{Group} \quad m: \quad \{\tan(d_m\psi_1(n)) \cdots \tan(d_m\psi_K(n))\}. \tag{3.23}$$

Note that from (3.21) shows us that the AR modeled sources $\boldsymbol{p}(i)$ and the regression coefficients $\boldsymbol{A}_T^m(n)$ satisfy one-to-one relationship, and there is no need to deal with the pair-matching problem for the different DOAs. From (3.22) and (3.23), we can find that any group contains the entire information of the angles, implying that we can achieve DOA estimation by just using any pair of sensors from the different subarrays. Therefore, our proposed method is also suitable for the case, where the number of sensors is much less than that of the sources. As is well-known, for $d_m > 1, (m = 2, \cdots, M)$, an ambiguity may happen due to the inability of the tangent function being able to deal with more than 180 degrees, where the arc-tangent function is to select the right DOA from $|d_m \psi_k(n)| > \pi/2$. Herein, we employ a technique similar to that suggested in Section 2.2 to handle this ambiguity problem. According to the assumption $d_1 \leq 1$, we can obtain the rough DOA estimate without ambiguity by using the first group of data. There exist multiple values for the same tangent value, i.e.,

$$\psi_{m,k}^{l_m} = \psi_{1,k} + \pi l_m d_m \left\lceil -(\pi + \psi_{m,k}) d_m / \pi \right\rceil \leq l_m \leq \left\lfloor (\pi - \psi_{m,k}) d_m / \pi \right\rfloor, \tag{3.24}$$

where $\psi_{1,k} \triangleq \pi \sin(\hat{\theta}_{(1),k}(n))$ is estimated by the first group in (3.22). So the unambiguous estimate for the k^{th} angle can be obtained as

$$\hat{\theta}_{(m),k}(n) = \arctan(\arg\min_{\psi_{m,k}^{l_m}} \left| \arcsin(\psi_{m,k}^{l_m}/\pi) - \hat{\theta}_{(m-1),k}(n) \right|) \tag{3.25}$$

where $\hat{\theta}_{(m),k}(n)$ is the m^{th} estimated value based on the $(m-1)^{th}$ unambiguous estimate $\hat{\theta}_{(m-1),k}(n)$. Note that it is also possible to use two subarrays to estimate the DOAs without requiring the reference element, in which case an even unitary transformation matrix [93] should be utilized.

3.2.4 Simulation Results for DOA Estimation and Tracking

In this subsection, some simulation results are presented to show the effectiveness of the proposed DOA estimation and tracking method. Here, we assume that the AR model order of the source signals is fixed and known. In practical applications, it can be determined by standard model order estimation techniques. Note that the reference sensor is not used to carry out DOA estimation in the following simulations in order to reduce the cost of array system for practical considerations.

Example 1: *Performance of joint AR coefficients and DOA estimation*

In this example, the sensor array consists of three isotropic antennas spaced by half a wavelength, i.e., $d_1 = 1/2$. The sources are two second-order AR stationary signals with coefficients $\boldsymbol{a}_1 = [0.872 \quad -0.550]$ and $\boldsymbol{a}_2 = [1.096 \quad -0.870]$ and each has a signal to noise ratio (SNR) of 30dB. Here, the SNR of the k^{th} signal is defined as the ratio of the k^{th} signal power to the average power of noise $\boldsymbol{e}(t)$. The DOAs are chosen to be 0° and 20°. The other parameters are chosen as follows: the initial guess for the two DOAs is 5° and 25°, the initial AR coefficients are $[0.772 \quad -0.450]$ and $[0.96 \quad -0.77]$, and N=30 snapshots are used. As shown in Fig.3.2 and Fig.3.3, the proposed method can estimate the DOAs of the two sources and the AR coefficients very well using only two sensors spaced by half wavelength, which is almost impossible

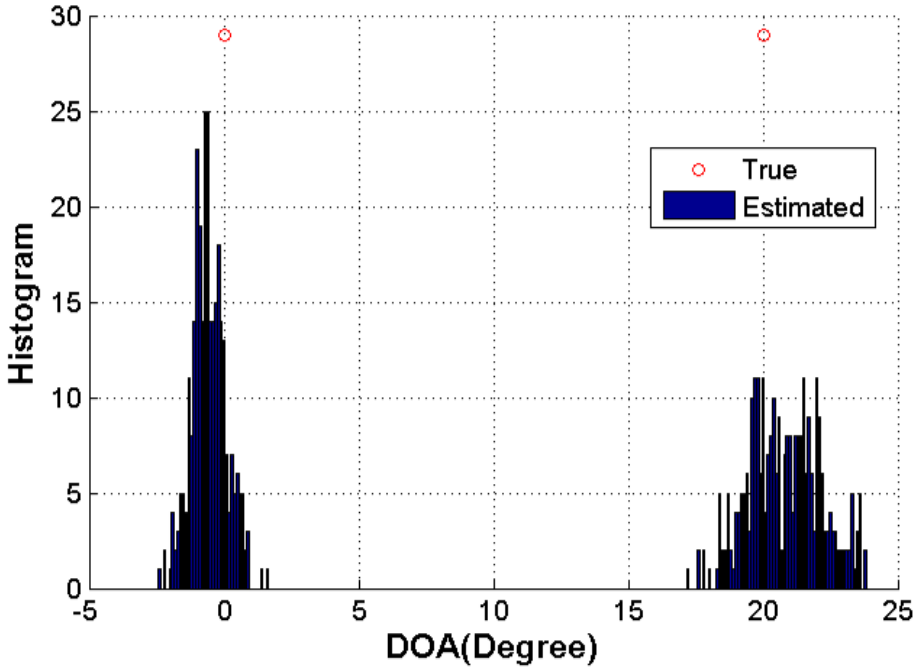
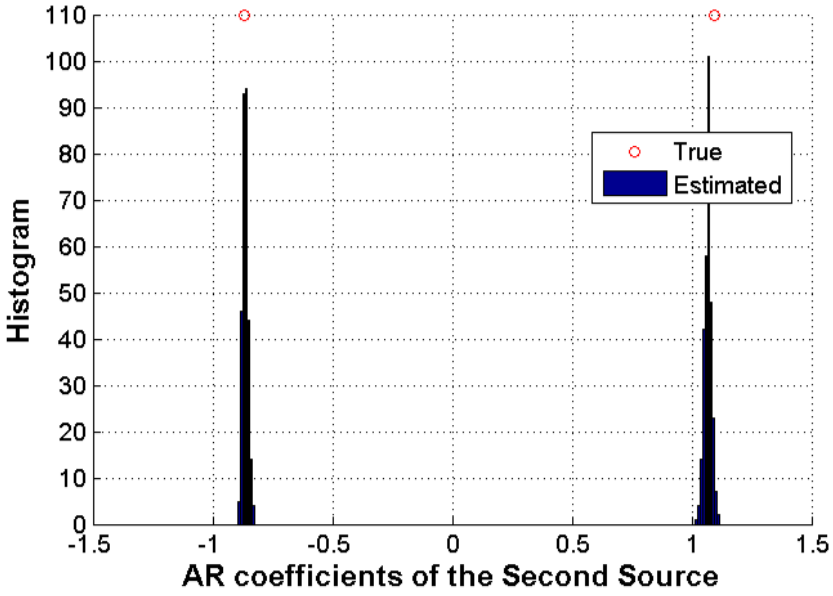


Figure 3.2: Histogram of DOA estimates for two AR modeled sources of DOA at $[0^\circ \quad 20^\circ]$ using two sensors with $d_1 = 1$.

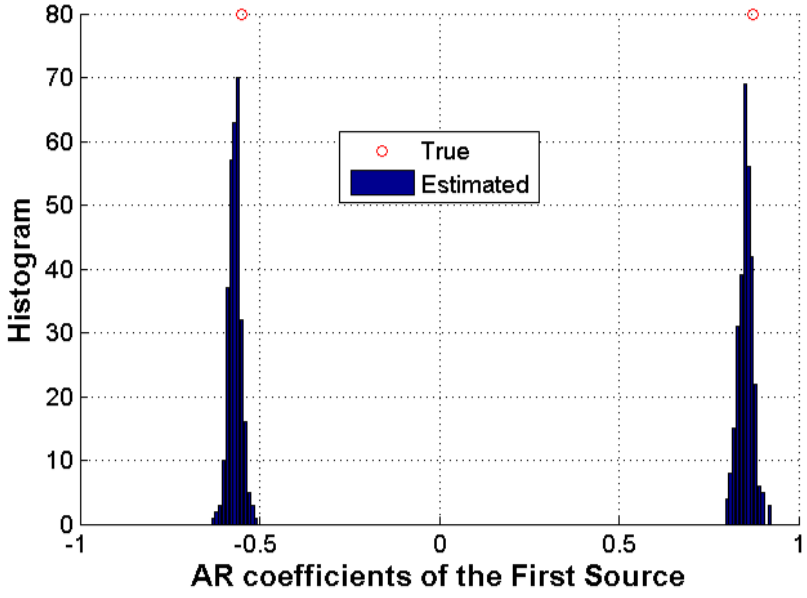
for conventional subspace-based method to do. This confirms the effectiveness of the proposed method in exploiting the dynamic information of the AR modeled sources.

Example 2: *Performance of DOA tracking*

To assess the DOA tracking performance of the proposed method, we set the simulation conditions as follows: AR coefficients are the same as those in Example 1 and herein these values are assumed known previously and SNR is set at 10dB. Five sensors with $d_1 = 1$ and $d_2 = 2$ are used. Fig.3.4 depicts the DOA tracking result of the proposed method in comparison with that obtained using the PAST method in [60] when the DOAs of the two sources are 0° and -20° and N=3. Fig.3.5 shows

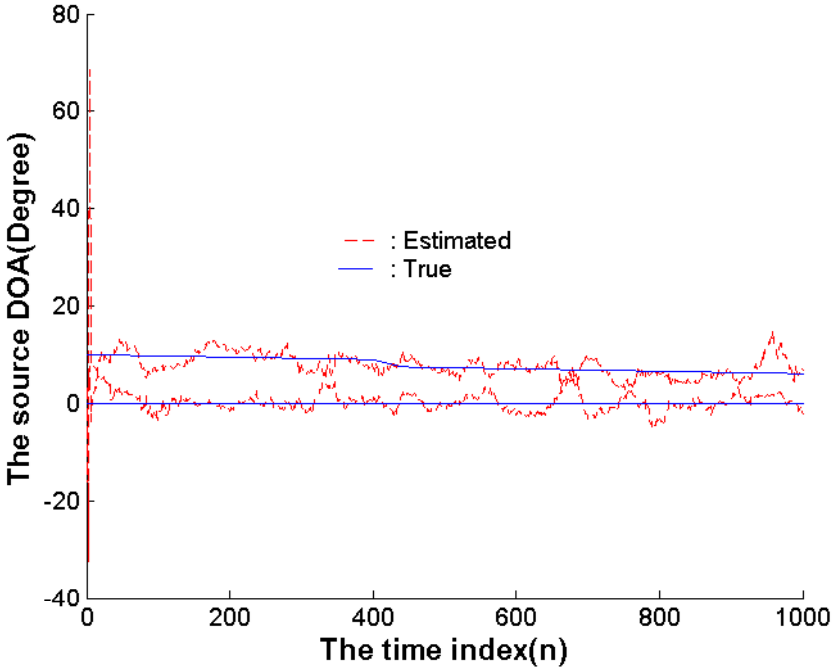


(a)

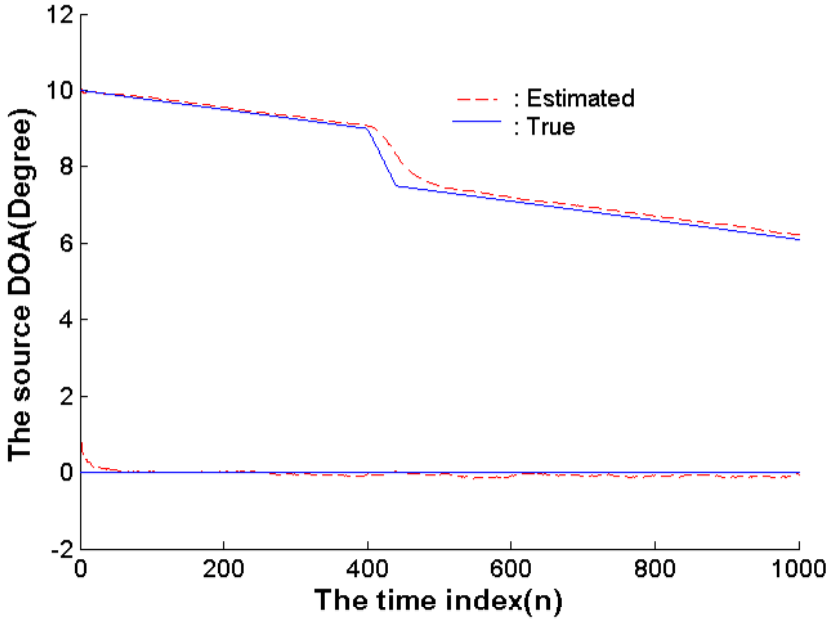


(b)

Figure 3.3: Histogram of AR coefficients estimation for two AR sources with DOA at $[0^\circ \quad 20^\circ]$ and AR coefficients $[0.872 \quad -550]$ and $[1.096 \quad -0.870]$ using two sensors with $d_1 = 1/2$.

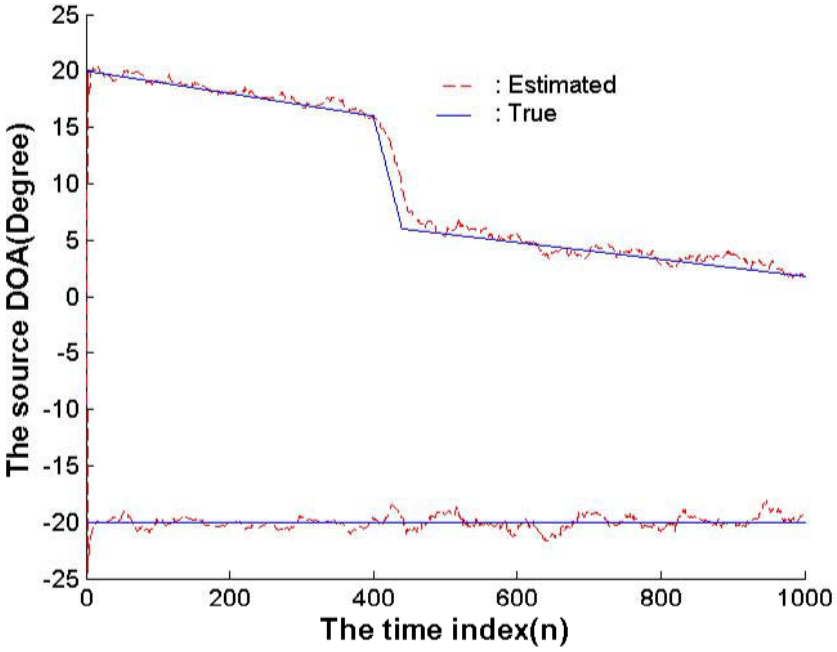


(a)

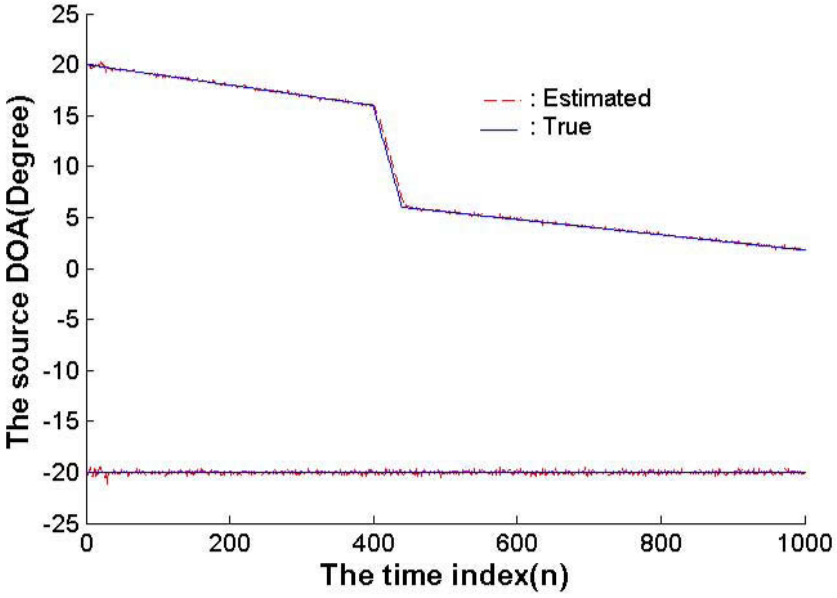


(b)

Figure 3.4: Tracking two sources: one fixed DOA of 00 and the other moving from 10° to 6.1° with N=3 for each interval. (a) PAST and (b) proposed method.

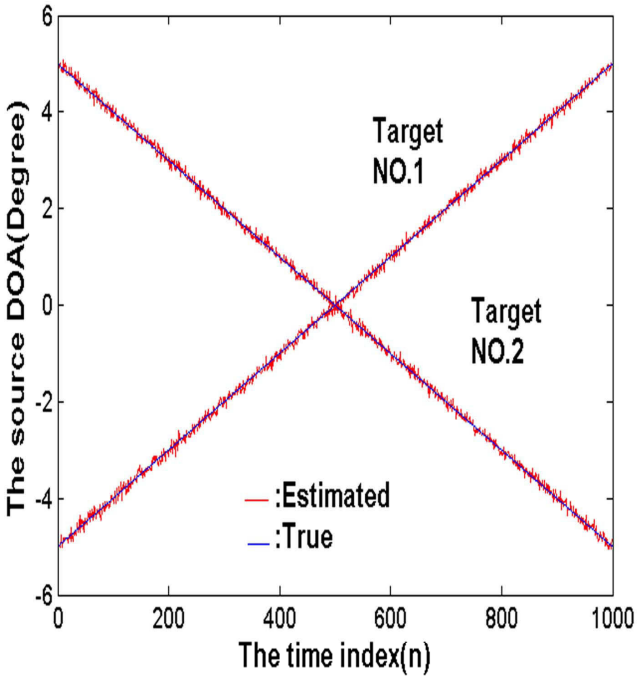


(a)

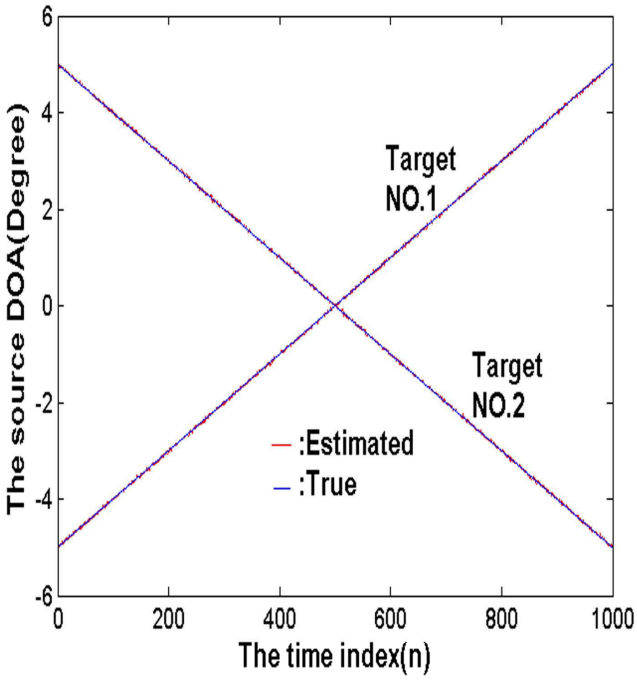


(b)

Figure 3.5: Tracking two sources: one fixed DOA of -200and the other moving from 20° to 1.8° with N=30 for each interval. (a) PAST, and (b) proposed method.



(a)



(b)

Figure 3.6: Tracking two sources via five sensors with $d_1 = 1$ and $d2 = 4$ by the proposed method (a) group 1, i.e., $d_1 = 1$, and (b) group 2, i.e., $d_2 = 4$.

the simulation results of both the proposed and the PAST methods for the DOA that is governed by

$$\text{DOA} = \begin{cases} 10 - n/400 \quad \text{and} \quad 20 - n/100 & n \in [1 \quad 400] \\ 9 - 1.5(n-400)/40 \quad \text{and} \quad 16 - 2.5(n-400)/10 & n \in [401 \quad 440] \\ 7.5 - 2.5(n-440)/1000 \quad \text{and} \quad 6 - 7.5(n-440)/1000 & n \in [441 \quad 1000] \end{cases} \tag{3.26}$$

when N=30. Both of the initial DOAs are 2° and 12°. From Fig.3.4, we see that the subspace-based method using PAST is unable to resolve the closely-spaced angles in this case, while the proposed method yields satisfactory tracking results. Fig.3.5 shows that although the PAST can track the separated DOAs, it is hard to track the fast moving target while our method can handle this situation satisfactorily. Two possible explanations are 1) our method takes advantage of the temporal information and 2) the subspace swap and leakage between the signal and noise subspaces degrades considerably the performance of the subspace-based methods in the case of closely spaced DOAs. It is also clear that the more snapshots used, better performance for fast moving targets can be achieved. Moreover, through a large number of simulations, we have also found that the tracking performance is not sensitive to the initial guess of DOA values.

Example 3: *Performance of DOA tracking for two moving targets* The tracking performance of the proposed method for two moving targets is now examined. The simulation conditions are similar to those in Example 2, except that the AR coefficients of the first source is a non-stationary AR signal with $\boldsymbol{a}_1 = [-0.450 \quad 0.772]$, $d_1 = 1$

and $d_2 = 4$, and the DOAs of the two sources are generated according to $[-5 + n/100 \quad 5 - n/100], n \in [1 \quad 1000]$. The averaged results by 50 independent trials are shown in Fig.3.6. Clearly, the proposed method provides a good tracking performance for two moving targets, with one source being a stationary signal source and the other a non-stationary signal source.

3.3 DOA Estimation based on General SLA

In this section, we introduce a novel method to estimate AR-modeled source waveforms and their DOA using the Kalman filter and TLS techniques based on sparse linear arrays. The key idea of our method lies in that each sensor of arrays is considered as a subsystem to obtain the angle information to estimate the DOA, and then the whole information combined to derive an optimal estimate for the angles. Unlike the method proposed in the previous section, which assumed that the AR coefficients are real numbers and the array is symmetric sparse array, herein the proposed method will be used in more general cases for complex AR coefficients and SLA. Moreover, in the previous section we exploited the QRD-RLS techniques to estimate the AR coefficients and DOAs, which however suffer from the effects of the bias due to the errors in estimating the state values [82]. Therefore, this section introduces an unbiased estimator for AR coefficients and DOAs.

3.3.1 Initial State Values

In this section, the SLA array shown in Fig.2.9 is used and the data model is the same as in the previous section except for the real AR coefficients. Note that the directions have the same definition as in Fig.3.1. It is well-known from the KF theory that the procedures for estimating the source waveforms in the previous section are optimal, if we know the accurate model parameters such as $\boldsymbol{p}(1)$, $\boldsymbol{Q}$, $\boldsymbol{R_e}$, $\boldsymbol{F}$, and $\boldsymbol{A}$. However, in some practical applications we have to estimate these parameters only using the measurements; in the following context, we propose methods to obtain these parameters.

Now let us show as to how to estimate the the initial $\boldsymbol{p}(1)$ and $\boldsymbol{Q}$. Consider again the model of (3.9) and (3.10), which can be rewritten as

$$\begin{aligned} \boldsymbol{x}(i) &= \boldsymbol{A\Gamma F p}(i-1) + \boldsymbol{A\Gamma v}(i-1) + \boldsymbol{e}(i) \\ &= \boldsymbol{A\Gamma F}^{i-1}\boldsymbol{p}(1) + \boldsymbol{A\Gamma v}(i-1) + \cdots + \boldsymbol{A\Gamma F}^{i-2}\boldsymbol{v}(1) + \boldsymbol{e}(i). \end{aligned} \tag{3.27}$$

Therefore, by collecting the first i equations of (3.27), we may write them compactly as

$$\begin{aligned} \boldsymbol{x}_i &\triangleq \left[\boldsymbol{x}^T(1), \boldsymbol{x}^T(2), \cdots, \boldsymbol{x}^T(i)\right]^T \\ &= \mathbf{B}\boldsymbol{p}(1) + \boldsymbol{n}_i, \end{aligned} \tag{3.28}$$

where

$$\mathbf{B} \triangleq \begin{bmatrix} \boldsymbol{A\Gamma} \\ \boldsymbol{A\Gamma F} \\ \vdots \\ \boldsymbol{A\Gamma F}^{i-1} \end{bmatrix}, \quad \boldsymbol{n}_i \triangleq \begin{bmatrix} 0 & \cdots & 0 \\ \boldsymbol{A\Gamma} & 0 & \cdots \\ \vdots & \ddots & \vdots \\ \boldsymbol{A\Gamma F}^{i-2} & \cdots & \boldsymbol{A\Gamma} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(1) \\ \boldsymbol{v}(2) \\ \vdots \\ \boldsymbol{v}(i-1) \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}(1) \\ \boldsymbol{e}(2) \\ \vdots \\ \boldsymbol{e}(i) \end{bmatrix}. \tag{3.29}$$

It is easy to verify that $\boldsymbol{n}_i$ is a $\left(\sum_{k=1}^{K} L_k\right)$-dimensional Gaussian random vector with zero-mean and the covariance matrix as follows,

$$\boldsymbol{C}_{\boldsymbol{n}_i} \triangleq \begin{bmatrix} 0 & \cdots & 0 \\ \boldsymbol{A\Gamma Q} & 0 & \cdots \\ \vdots & \ddots & \vdots \\ \boldsymbol{A\Gamma F}^{i-2}\boldsymbol{Q} & \cdots & \boldsymbol{A\Gamma Q} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ \boldsymbol{A\Gamma} & 0 & \cdots \\ \vdots & \ddots & \vdots \\ \boldsymbol{A\Gamma F}^{i-2} & \cdots & \boldsymbol{A\Gamma} \end{bmatrix}^H + \boldsymbol{R_e} \begin{bmatrix} \boldsymbol{I}_{\sum_{k=1}^{K} L_k} \\ \boldsymbol{I}_{\sum_{k=1}^{K} L_k} \\ \vdots \\ \boldsymbol{I}_{\sum_{k=1}^{K} L_k} \end{bmatrix}. \tag{3.30}$$

Using the principle of the generalized Least squares [94], we can obtain the optimal estimate for the j^{th} iteration of $\boldsymbol{p}(1)$ as

$$\boldsymbol{p}^{(j)}(1) = \left(\mathbf{B}^{(j-1)H}\boldsymbol{C}_{\boldsymbol{n}_i}^{(j-1)-1}\mathbf{B}^{(j-1)}\right)^{-1}\mathbf{B}^{(j-1)H}\boldsymbol{C}_{\boldsymbol{n}_i}^{(j-1)-1}\boldsymbol{x}_i. \tag{3.31}$$

It is easy to prove that (3.31) is an unbiased estimator to obtain $\boldsymbol{p}^{(j)}(1)$ with the covariance $\boldsymbol{C}_{\boldsymbol{p}^{(j)}(1)} = (\mathbf{B}^{(j-1)H}\boldsymbol{C}_{\boldsymbol{n}_i}^{(j-1)-1}\mathbf{B}^{(j-1)})^{-1}$. In this way, we use $\boldsymbol{Q}^{(j)}(\sum_{p=1}^{k} L_p - L_k + 1, \sum_{p=1}^{k} L_p - L_k + 1) = \boldsymbol{C}_{\boldsymbol{p}^{(j)}(1)}(\sum_{p=1}^{k} L_p - L_k + 1, \sum_{p=1}^{k} L_p - L_k + 1)$ for the next iteration. In

some applications, the covariance of the additive array noise $\boldsymbol{R_e}$ can be known a priori or estimated during the non-signal periods. Thus for the first iteration, there is no information of $\boldsymbol{Q}$ which can be used to estimate $\boldsymbol{p}^{(1)}(1)$; we only employ the first array sample to estimate $\boldsymbol{s}^{(1)}(1) = \boldsymbol{\Gamma p}^{(1)}(1)$, that is,

$$\boldsymbol{s}^{(1)}(1) = \begin{cases} \left(\boldsymbol{A}^H \boldsymbol{R_e}^{-1} \boldsymbol{A}\right)^{-1} \boldsymbol{A}^H \boldsymbol{R_e}^{-1} \boldsymbol{x}_i & \text{for} \quad K \leq (M+1) \\ \boldsymbol{A}^{\dagger} \boldsymbol{x}_i & \text{for} \quad K > (M+1) \end{cases} . \tag{3.32}$$

3.3.2 Unbiased Estimation for AR Coefficients Based on Observations

Using the same philosophy as in (3.21), we can consider each sensor of SLA as a subsystem

$$x^m(i) = \boldsymbol{A}^m \boldsymbol{\Gamma p}(i) + e^m(i), \quad (m = 0, 1, \cdots, M). \tag{3.33}$$

Recalling (3.1), (3.33) can be rewritten as

$$x^m(i) = \sum_{k=1}^{K} a^m(\theta_k) \left(\sum_{l_k=1}^{L_k} a_{l_k}^k p_k(i - l_k) + v_k(i) \right) + e^m(i), \tag{3.34}$$

where $a^m(\theta_k)$ denotes the m^{th} element of the steering vector of the k^{th} signal. Now, let us define the autocovariance values of the measurements of the m^{th} sensor and the source waveforms at lag l by

$$r_l^m = E\left[x^m(i) x^{m*}(i-l)\right] \quad \text{and} \quad r_l^k = E\left[p_k(i) p_k^*(i-l)\right]. \tag{3.35}$$

Theorem 2. *Assume that the signal waveforms are uncorrelated with each other, i.e.,* $E\left[p_a(i)p_b^*(j)\right] = 0, \quad a \neq b$, *and let* $l_M \triangleq \max\left\{L_k\right\}_{k=1}^K$. *Then,* $r_{l_M+s}^m, s = 1, 2, \cdots, S$ *is immune to additive noise, and the unbiased estimate for the AR coefficients is given by*

$$[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_K]^T = \left(\boldsymbol{R}_S^H \boldsymbol{R}_S\right)^{-1} \boldsymbol{R}_S^H \boldsymbol{r}_S^m, \tag{3.36}$$

where $\boldsymbol{R}_S \triangleq [\boldsymbol{r}_1^T, \boldsymbol{r}_2^T, \cdots, \boldsymbol{r}_S^T]^T$ *and* $\boldsymbol{r}_s \triangleq [r_{l_M+s-1}^1, \cdots, r_{l_M+s-l_1}^1, r_{l_M+s-1}^2, \cdots, r_{l_M+s-l_2}^2, \cdots, r_{l_M+s-1}^K, \cdots, r_{l_M+s-l_K}^K]^T$, *and* $\boldsymbol{r}_S^m \triangleq [r_{l_M+1}^m, r_{l_M+2}^m, \cdots, r_{l_M+S}^m]^T$.

Proof: Combining (3.34) and (3.35), we can derive that

$$\begin{aligned} r_{l_M+s}^m &= E\left[x^m(i)x^{m*}(i-l_M-s)\right] \\ &= E\left[\left(\sum_{k=1}^K a^m(\theta_k)\left(\sum_{l_k=1}^{L_k} a_{l_k}^k p_k(i-l_k) + v_k(i)\right) + e^m(i)\right)\right. \\ &\qquad \left.\left(\sum_{k=1}^K a^m(\theta_k) p_k(i-l_M-s) + e^m(i-l_M-s)\right)^*\right] \\ &= \sum_{k=1}^K \left(\sum_{l_k=1}^{L_k} a_{l_k}^k r_{l_M+s-l_k}^k\right) \\ &\triangleq \boldsymbol{r}_s^T [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_K]^T. \end{aligned} \tag{3.37}$$

where $\boldsymbol{r}_s \triangleq [r_{l_M+s-1}^1, \cdots, r_{l_M+s-l_1}^1, r_{l_M+s-1}^2, \cdots, r_{l_M+s-l_2}^2, \cdots, r_{l_M+s-1}^K, \cdots, r_{l_M+s-l_K}^K]^T$. Clearly, the vector $\boldsymbol{r}_s$ is not affected by the additive noise. Therefore, $r_{l_M+s}^m$ is also free from additive noise. Then by combining $S \geqslant \sum_{k=1}^K L_k$ equations of (3.37), the AR coefficients can be obtained by using (3.36) based on the LS technique. □

Suppose that there are N snapshots available to estimate the AR coefficients.

Then, we replace r_l^m and r_l^k in (3.35) by their estimates

$$\hat{r}_l^m = \frac{1}{N - l_M + 1} \sum_{i=l_M}^{N} [x^m(i) x^{m*}(i-l)] \quad \text{and} \quad \hat{r}_l^k = \frac{1}{N - l_k + 1} \sum_{i=l_k}^{N} [p_k(i) p_k^*(i-l)] \tag{3.38}$$

Owing to the estimation errors in $\left\{\hat{r}_l^m, \hat{r}_l^k\right\}$, the matrix equation (3.36) does not hold exactly in general. Thus, we solve (3.37) for AR coefficients in a TLS sense [95]. Since each sensor is a subsystem with the same AR coefficients, we can use the mean of all the estimated values as the final estimate.

3.3.3 DOA Estimation Using TLS

Eq.(3.33) can be rewritten as

$$x^m(i) = \sum_{k=1}^{K} a^m(\theta_k) p_k(i) + e^m(i) = \boldsymbol{s}^T(i) \boldsymbol{A}^{mT} + e^m(i), \quad m = 0, 1, \cdots, M \tag{3.39}$$

where $\boldsymbol{s}(i) \triangleq [p_1(i), p_2(i), \cdots, p_K(i)]^T$ denotes the source waveforms of the i^{th} snapshot. Then, rewrite the whole N snapshots compactly as

$$\boldsymbol{x}^m \triangleq \boldsymbol{S}^T \boldsymbol{A}^{mT} + e^m(i) = [\boldsymbol{s}(1), \boldsymbol{s}(2), \cdots, \boldsymbol{s}(N)]^T \boldsymbol{A}^{mT} + e^m(i), \tag{3.40}$$

where $\boldsymbol{x}^m \triangleq [x^m(1), x^m(2), \cdots, x^m(N)]^T$ is the N snapshots measured by the m^{th} sensor of SLA.

It is well known that when the noise in $\boldsymbol{S}$ is zero and the noise in $\boldsymbol{x}^m$ is zero mean Gaussian, the LS solution $\boldsymbol{A}_{LS}^m$ is identical to the maximum-likelihood one [94].

Unfortunately, herein $\boldsymbol{S}$ is unknown and obtained using (3.11), which is affected by the estimation errors and additive noise; therefore, $\boldsymbol{A}_{LS}^{m} = \left(\left(\boldsymbol{S}^{*} \boldsymbol{S}^{T} \right)^{-1} \boldsymbol{S}^{*} \boldsymbol{x}^{m} \right)^{T}$ is no longer optimal from a statistical point of view and it suffers from bias and increased covariance due to the accumulation of noise errors in $(\boldsymbol{S}^{*}\boldsymbol{S}^{T})$ [82]. To cope with this problem, in the following, we will introduce a TLS-sense method.

From (3.11)-(3.14), we can obtain the estimated value $\hat{\boldsymbol{s}}(i) = \boldsymbol{\Gamma}\hat{\boldsymbol{p}}(i|i)$ which is a unbiased estimate with the covariance $\boldsymbol{C}_{\hat{\boldsymbol{s}}(i)} \triangleq \boldsymbol{\Gamma} P(i|i) \boldsymbol{\Gamma}^{T}$. Therefore, we propose the TLS estimator as

$$\boldsymbol{A}_{TLS}^{m} = \left((\hat{\boldsymbol{S}}^{*} \hat{\boldsymbol{S}}^{T} - \sum_{i=1}^{N} \boldsymbol{C}_{\hat{\boldsymbol{s}}(i)})^{-1} \hat{\boldsymbol{S}}^{*} \boldsymbol{x}^{m} \right)^{T} \tag{3.41}$$

to obtain an approximately unbiased estimate of $\boldsymbol{A}^{m}$ and its auto-covariance matrix is $\boldsymbol{C}_{\boldsymbol{A}_{TLS}^{m}} \triangleq \boldsymbol{R}_{S}^{-1}(N) \left(\boldsymbol{R}_{S}(N) + \sum_{i=1}^{N} \boldsymbol{C}_{\hat{\boldsymbol{s}}(i)} \right) \boldsymbol{R}_{S}^{-H}(N) \sigma_{m}^{2}$, where $\boldsymbol{R}_{S}(N) \triangleq \frac{\boldsymbol{S}^{*}\boldsymbol{S}^{T}}{N}$, and the cross-covariances equal to $\boldsymbol{C}_{\boldsymbol{A}_{TLS}^{pT}, \boldsymbol{A}_{TLS}^{qT}} \triangleq \boldsymbol{R}_{\boldsymbol{e}}(p,q) \boldsymbol{R}_{S}^{-1}(N) \left(\boldsymbol{R}_{S}(N) + \sum_{i=1}^{N} \boldsymbol{C}_{\hat{\boldsymbol{s}}(i)} \right) \boldsymbol{R}_{S}^{-H}(N)$, $p \neq q$ (see Appendix B). Note that when the spatial noise is Gaussian white noise, the covariance matrix $\boldsymbol{R}_{\boldsymbol{e}}$ is a diagonal matrix. Then the cross-covariances are zero, which is a reasonable assumption in array signal processing, especially for SLA. Therefore, without loss of generality, we assume $\boldsymbol{R}_{\boldsymbol{e}} = diag([\sigma_{0}^{2}, \sigma_{1}^{2}, \cdots, \sigma_{M}^{2}])$ in the following analysis.

Obviously, the order of DOA information of $\boldsymbol{A}_{TLS}^{m}$ depends on the arrangement of the signals on the estimate of source waveform, i.e., the k^{th} element of $\boldsymbol{s}(i)$. Therefore, we can obtain $\boldsymbol{a}_{TLS}(\theta_{k}) = [a_{TLS}^{0}(\theta_{k}), a_{TLS}^{1}(\theta_{k}), \cdots, a_{TLS}^{M}(\theta_{k})]^{T}$ as the k^{th} spatial feature, which does affect the generality of the analysis. Now, we make the following

partition of $\boldsymbol{a}_{TLS}(\theta_k)$ with two vectors as

$$\begin{aligned} \boldsymbol{a}_{TLS}^{(1)}(\theta_k) &\triangleq [a_{TLS}^0(\theta_k), a_{TLS}^1(\theta_k), \cdots, a_{TLS}^{M-1}(\theta_k)]^T \\ \boldsymbol{a}_{TLS}^{(2)}(\theta_k) &\triangleq [a_{TLS}^1(\theta_k), a_{TLS}^2(\theta_k), \cdots, a_{TLS}^{M}(\theta_k)]^T \end{aligned} . \tag{3.42}$$

We define the vector

$$\hat{\boldsymbol{d}}_k \triangleq \boldsymbol{a}_{TLS}^{(1)}(\theta_k) \circ \boldsymbol{a}_{TLS}^{(1)}(\theta_k)^* \triangleq \boldsymbol{d}_k + \triangle \boldsymbol{d}_k, \tag{3.43}$$

where $\boldsymbol{d}_k \triangleq [e^{jd_1\psi_k}, e^{jd_2\psi_k}, \cdots, e^{jd_M\psi_k}]^T$ and $\triangle \boldsymbol{d}_k$ is a Gaussian random vector with zero mean and the covariance matrix $\boldsymbol{C}_{\hat{\boldsymbol{d}}_k} = diag\left([\sigma^2_{\triangle d_k^1}, \sigma^2_{\triangle d_k^2}, \cdots, \sigma^2_{\triangle d_k^M}]\right)$, where $\sigma^2_{\triangle d_k^m} \triangleq \sigma^2_{\triangle a^m(\theta_k)} + \sigma^2_{\triangle a^{m-1}(\theta_k)}$ and $\sigma^2_{\triangle a^m(\theta_k)} = [\boldsymbol{C}_{\boldsymbol{A}_{TLS}^m}]_{k,k}$ (see Appendix C). Clearly, the MLE of θ_k, which is equivalent to the following GLS problem, is found by minimizing [96]

$$L(\theta_k) = \left(\boldsymbol{d}_k - \hat{\boldsymbol{d}}_k\right)^T \boldsymbol{C}_{\hat{\boldsymbol{d}}_k}^{-1} \left(\boldsymbol{d}_k - \hat{\boldsymbol{d}}_k\right) \tag{3.44}$$

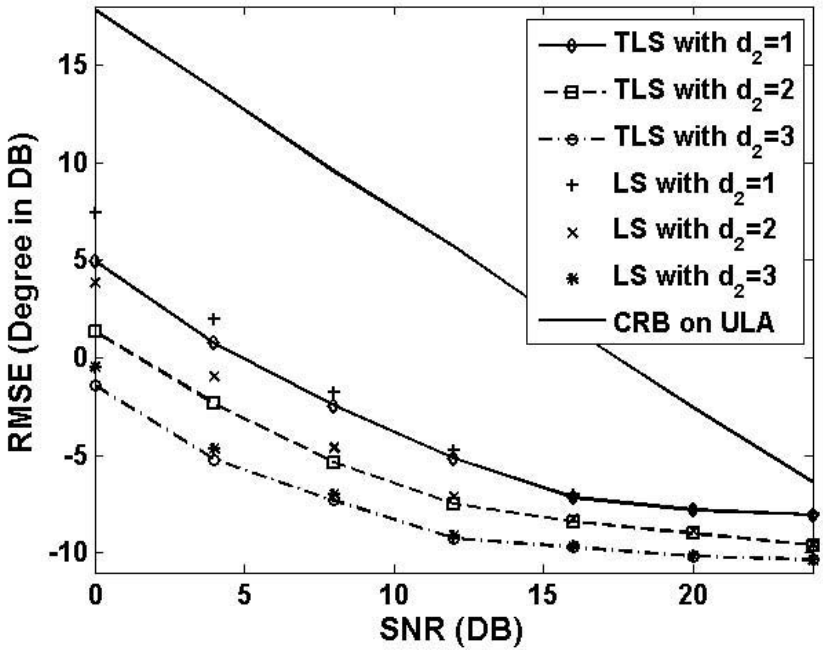
Since SLA is used in this section, there may exist the problem of ambiguity by implementing (3.44) directly. Similar to the method in Section 2.3, we determine the search range by the unambiguous DOA and unambiguous interval. If we assume that one of the element-spacings $\{d_m\}_{m=1}^M$ is less than or equal to unity, say $d_1 \leq 1$, which can be used to estimate the unambiguous angle value $\hat{\theta}_k^{(1)}$, then the unambiguous interval can be obtained by the largest element-spacing, i.e., $d_p = \max_m d_m$. Hence, the search range is $[\hat{\theta}_k^{(1)} - \triangle\theta_k^p \quad \hat{\theta}_k^{(1)} + \triangle\theta_k^p]$. Therefore, the final DOA estimation of the k^{th} signal can be obtained by computing (3.44) in the searching space of interest.

3.3.4 Simulation Results for DOA Estimation Based on Kalman Filter and TLS Techniques

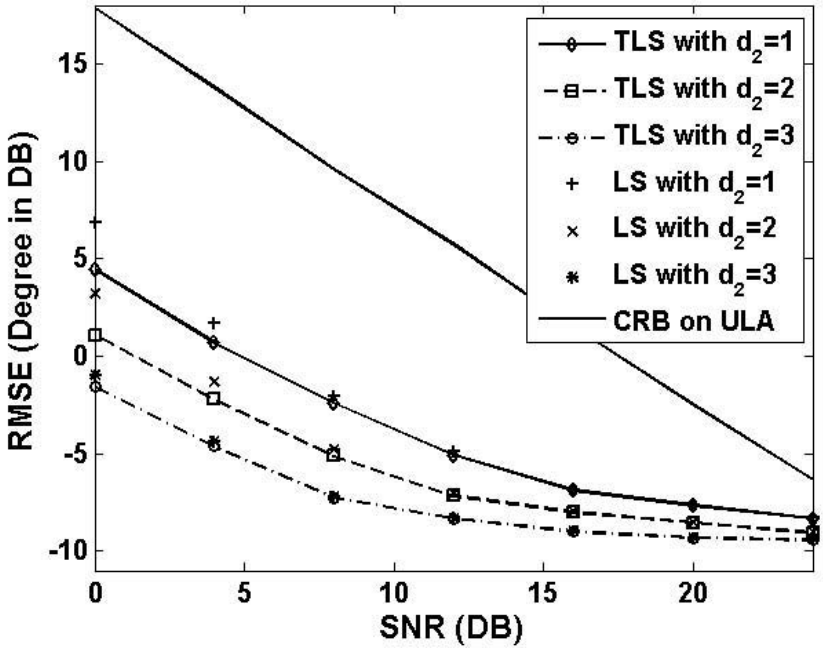
In this section, simulation results are presented to validate the proposed method and illustrate its performance. In the first two examples, three sensors will be used to estimate the DOA of two AR modeled signals, where three SLA have the same value of $d_1 = 1$ and the different values of d_2, i.e., $d_2 = 1, 2$, and 3. In the third example, we consider the case, where three-sensor SLA is exploited to estimate three AR modeled sources. For all tests, 1000 independent trials are carried out.

Example 1: *Performance of DOA estimation versus SNR*

In the first example, a three-sensor SLA system is used to estimate the DOA of two AR modeled sources with coefficients $\boldsymbol{a}_1 = [0.872 \quad -0.550]$ and $\boldsymbol{a}_2 = [1.096 \quad -0.870]$ incident from $[\theta_1 \quad \theta_2] = [-5^\circ \quad 5^\circ]$ with identical powers and white Gaussian additive noises. For performance comparison, we evaluate the proposed method with respect to SNR from 0dB to 24dB for three SLA cases, that is, $d_2 = 1, 2$ and 3 with the same $d_1 = 1$, and the Cramer-Rao bound (CRB) [86] for a 3-sensor ULA shown in Fig.3.7. Notice that the proposed method is implemented on LS and TLS based techniques. We find that the estimation performance of the proposed TLS-based method is superior to that of the LS-based one, especially at low and medium SNR. The possible reason for this phenomenon lies in the fact that there exists estimated error and noise in the estimated AR source waveform producing a bias when LS-based method is exploited, while TLS-based method can reduce the effects of the error and noise. It is also seen that the RMSE of our proposed method is much smaller than the CRB due to the



(a)



(b)

Figure 3.7: RMSE of DOA estimation for two AR modeled sources versus SNR at (a)$\theta_1 = -5^\circ$ and (b)$\theta_2 = 5^\circ$ with 100 snapshots.

fact that our proposed method makes use of the temporal information to improve the performance of DOA estimation significantly.

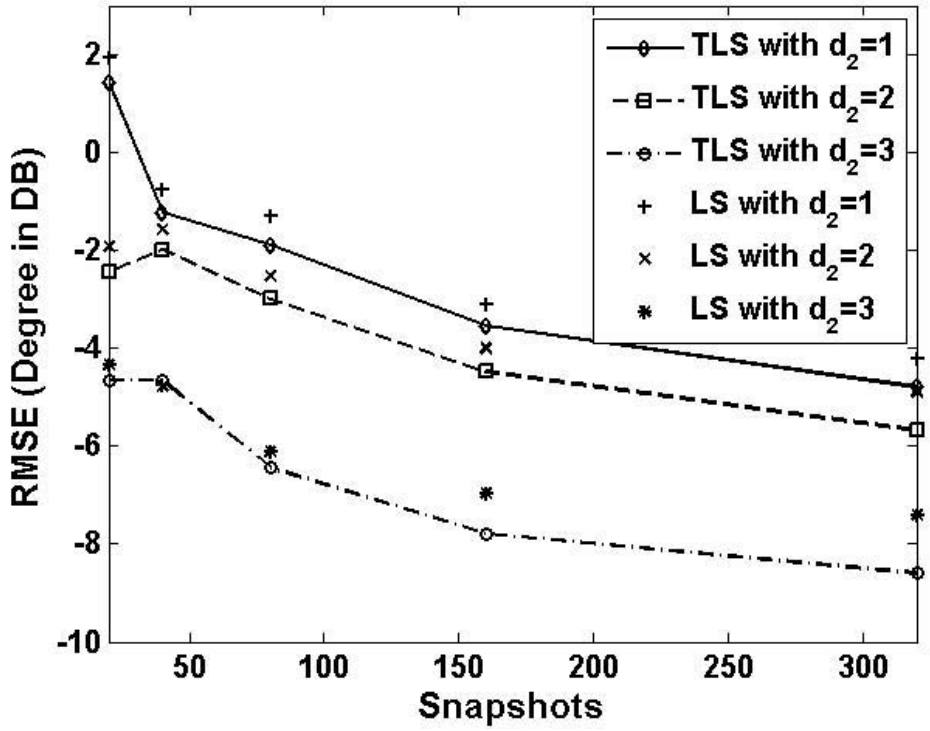


Figure 3.8: RMSE of DOA estimation for two AR modeled sources versus snapshots at $\theta_1 = -5^\circ$ with SNR=10dB.

Example 2: *Performance of DOA estimation versus the number of snapshots*

In this example, the same array system as in the first example is used except that the number of snapshots is varied and SNR is fixed in 10dB. The plots are shown in Fig.3.8, from which we can see that the RMSE of DOA estimation decreases with increasing number of snapshots. We also see that the larger the array aperture, the better the estimation performance.

Example 3: *Performance of DOA estimation for the same number of sensors and signals*

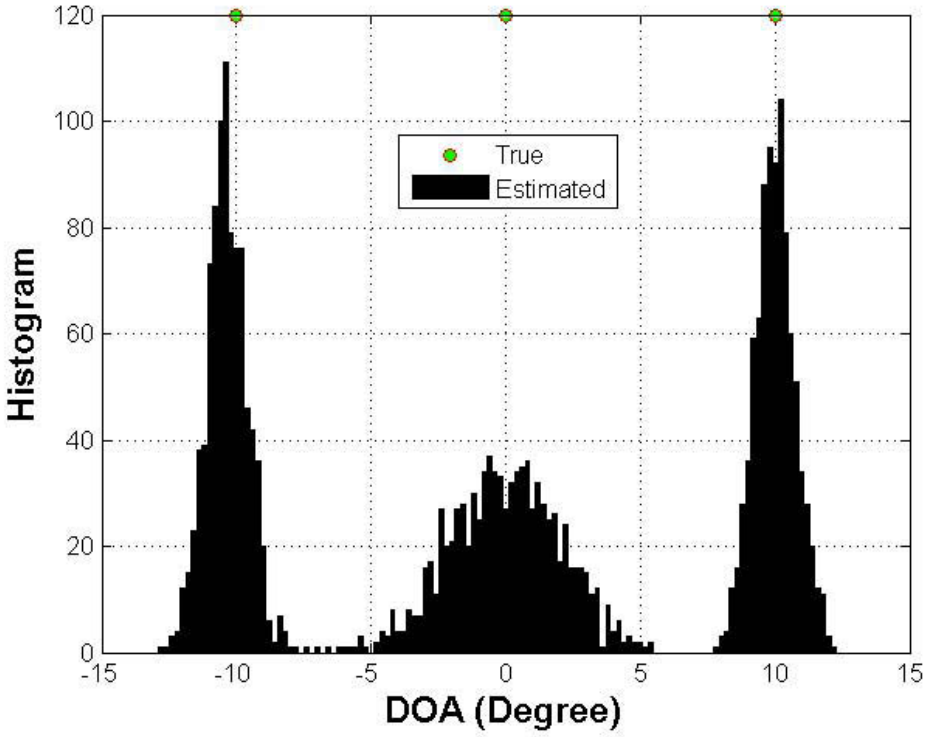


Figure 3.9: Histogram of DOA estimation for three AR modeled sources

In this example, we employ a three-sensor SLA with $d_1 = 1$ and $d_2 = 3$ to estimate the DOA of three AR modeled sources with coefficients $\boldsymbol{a}_1 = [0.872 \quad -0.550]$, $\boldsymbol{a}_2 = [1.096 \quad -0.870]$ and $\boldsymbol{a}_3 = [0.570 \quad -0.222]$ incident from $[\theta_1 \quad \theta_2 \quad \theta_3] = [-10^\circ \quad 0^\circ \quad 10^\circ]$ with identical powers and white Gaussian additive noises. The histogram plot shown in Fig.3.9 indicates that our proposed method can estimate more signal sources than traditional subspace-based methods can do due to the use of the temporal information.

3.4 Conclusion

In this chapter, we have proposed an efficient DOA estimation and tracking method based on Kalman filtering and the QRD-RLS algorithm. Thanks to the state information described by the AR model, we have been able to exploit the elegant Kalman filter to obtain the temporal information and then the efficient QRD-RLS to estimate the DOA and AR coefficients. A unitary transformation is also used to implement DOA estimation in the real-valued domain, where M real-valued vector groups are calculated for estimating the DOAs. In addition, we have also proposed a DOA estimation method based on Kalman filter and TLS. Differing from the first method, herein the symmetric array has been generalized to SLA. In this way, it is more efficient to exploit limited array elements to improve the DOA estimation performance. It is worth noting that since each sensor can be considered as a linear subsystem, the proposed estimator can be used in any array configurations proposed in Chapter 2.

Chapter 4

DOA Estimation and Tracking for Signals with Known Waveform

4.1 Introduction

In the previous chapters, we have developed DOA techniques by exploiting sparse arrays and/or AR modeled sources to significantly improve the estimation accuracy and angular resolution without using the source waveforms. In some applications such as active radar, active sonar, communication systems, and many other multisensor applications, the source waveforms of the signals of interest are possibly *a priori* known. Therefore, in this chapter we develop a novel DOA estimation method for signals with known waveforms based on the SLA. Unlike some previous methods, which estimate the DOA based on spatial signatures of the signals with known waveforms, the proposed method makes use of known waveforms to transform the maximum likelihood problem into multiple linear regression models, each of which contains a pair of DOA

and gain information. In this way, the proposed method can be implemented with a more general noise model than the commonly used uniform one. Here, regression analysis is performed to estimate the coefficients of each linear regression model, and the well-known generalized least squares is used to perform asymptotically optimal estimation of the angles and gains of targets without requiring a search over a large region of the parameter space. The effect of correlated sources on the performance of parameter estimation is also studied. In addition, a new phase unwrapped technique is proposed to deal with the problem of ambiguity. As the targets move, we propose two block RLS-based techniques to update the estimated target angles recursively, one is based on a block QR decomposition recursive least squares (QRD-RLS) technique and the other on a block regularized LS technique. Unlike the usual RLS-based techniques, where the linear regression coefficients are updated at every snapshot, we perform the update on the basis of time intervals. We use the block QRD-based technique to obtain current Cholesky factors from the previous Cholesky factors along with the received data of the time interval. Subsequently, we derive the second tracking method based on the regularized LS technique in view of the fact that the optimal weighting factor or regularization term can be derived by regression analysis [97, 98], and the previously estimated coefficients can be considered as a constrained vector to deal with some "bad" cases such as small snapshots in a time interval or highly correlated sources. Finally, simulation results that demonstrate the estimation performance of the proposed method are given, showing that the proposed DOA tracking techniques can be efficiently applied to a sparse antenna array and can provide a better tracking performance than some of the existing methods do.

4.2 DOA Estimation based on ML and GLS Technique

4.2.1 System Model

Suppose that there are K narrowband signal sources from directions $\theta_1, \theta_2, \cdots, \theta_K$ with the same wavelength, which impinge onto a linear array with inter-element spacings $d_1, d_2, \cdots, d_M$ normalized in terms of the half wavelength, as shown in Fig.2.9, where the sensor element at the origin is used as the reference. Notice that the DOA is now defined as the direction relative to the broadside of array for convenience of description. At time t, $x_m(t), (m = 0, 1, \cdots, M)$ denotes the complex signal at the m^{th} sensor of the linear array with $x_0(t)$ being the data received by the reference element. Then, the complex signals observed at the outputs of the $M + 1$ sensors $\boldsymbol{x}(t) \in \mathbb{C}^{(M+1)}$ can be written as

$$\boldsymbol{x}(t) = [x_0(t), x_1(t), \cdots, x_M(t)]^T \triangleq \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{e}(t), \tag{4.1}$$

where $\boldsymbol{s}(t) = \boldsymbol{\Xi}\boldsymbol{p}(t) \in \mathbb{C}^K$ is the source signals with $\boldsymbol{p}(t) = [p_1(t), p_2(t), \cdots, p_K(t)]^T$ denoting K known signal waveforms and $\boldsymbol{\Xi} = diag([\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_K])$ denoting the unknown complex gains of the K signals, and $\boldsymbol{e}$ is an $(M + 1)$-dimensional vector representing the complex white Gaussian measurement noise with zero mean and unknown covariance matrix $\boldsymbol{C_e} \triangleq E[\boldsymbol{e}(t)\boldsymbol{e}^H(t)] = diag([\sigma_0^2, \sigma_1^2, \cdots, \sigma_M^2])$ [99]. Note that the real part $\mathrm{Re}(\boldsymbol{e}(t))$ and the imaginary part $\mathrm{Im}(\boldsymbol{e}(t))$ are two real Gaussian random vectors of the same distribution $\mathbb{N}(\boldsymbol{0}, \frac{\boldsymbol{C_e}}{2})$. $\boldsymbol{A}$ has the same expression as $\boldsymbol{A}_x$

in (2.36). Since the waveforms of the signals are known, we can describe the output of each element of the antenna array $x_m(t)$ by the following linear regression model

$$x_m(t) = \sum_{k=1}^{K} p_k(t)\varepsilon_k \; e^{-j\varphi_k^{(m)}} + e_m(t) = (\boldsymbol{\Xi}\boldsymbol{p}(t))^T \boldsymbol{b}_m + e_m(t) \triangleq \boldsymbol{p}^T(t)\boldsymbol{g}_m + e_m(t), \quad (4.2)$$

where $\boldsymbol{g}_m \triangleq \boldsymbol{\Xi}\boldsymbol{b}_m$ contains the entire information of directions and complex gains of all the K signals. Thus, using (4.2), (4.1) can be rewritten as

$$\boldsymbol{x}(t) = (\boldsymbol{I}_{M+1} \otimes \boldsymbol{p}^T(t))\mathrm{vec}(\boldsymbol{G}) + \boldsymbol{e}(t), \quad (4.3)$$

where $\boldsymbol{G} \triangleq [\boldsymbol{g}_0, \boldsymbol{g}_1, \cdots, \boldsymbol{g}_M]$. It is important to stress that in the conventional methods, the spatial signatures are exploited to estimate the DOAs, while herein we pay special attention to the coefficients $\boldsymbol{g}_m, (m = 0, 1, \cdots, M)$ of the linear regression model. In the following subsection, a maximum likelihood estimator for $\boldsymbol{g}_m$ is derived based on the signal model obtained above.

4.2.2 Maximum Likelihood Estimator

Suppose the received array signals during the time period T are sampled as

$$\boldsymbol{x}(nT_s) = (\boldsymbol{I}_{M+1} \otimes \boldsymbol{p}^T(nT_s))\mathrm{vec}(\boldsymbol{G}) + \boldsymbol{e}(nT_s), 1 \leqslant n \leqslant N \quad (4.4)$$

where $T_s = T/N$ is the sampling interval and $\boldsymbol{e}(n)$ represents the samples of the noise $\boldsymbol{e}(t)$. The sampled version of the received data is often called "snapshots". Note that for moving targets in Section 4.3, each time period can be considered as a time

interval, that is, each time interval has N snapshots. By omitting T_s for convenience in the sequel, (4.5) can be rewritten as

$$\boldsymbol{x}(n) = (\boldsymbol{I}_{M+1} \otimes \boldsymbol{p}^T(n))\text{vec}(\boldsymbol{G}) + \boldsymbol{e}(n), 1 \leqslant n \leqslant N \tag{4.5}$$

Since the signals are assumed to have known waveforms with unknown complex gains and $\boldsymbol{e}(n)$ is zero-mean i.i.d Gaussian, each snapshot has a complex Gaussian probability density function (pdf) with a different mean but the same covariance, i.e., $\boldsymbol{x}(n) \sim \mathbb{CN}((\boldsymbol{I}_{M+1} \otimes \boldsymbol{p}^T(n))\text{vec}(\boldsymbol{G}), \boldsymbol{C}_n)$. The joint pdf of the independent snapshots $\boldsymbol{x}(n), n = 1, 2, \cdots, N$, can be written as

$$f(\boldsymbol{X}; \boldsymbol{G}, \boldsymbol{C_e}) = \prod_{n=1}^{N} \frac{e^{(-\boldsymbol{e}^H(n)\boldsymbol{C_e}^{-1}\boldsymbol{e}(n))}}{\pi^{M+1}\det(\boldsymbol{C_e})}, \tag{4.6}$$

where $\boldsymbol{X} = [\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(N)]$ is an $(M+1) \times N$ measurement matrix of the whole array. By taking the logarithm of (4.6), noting that $\boldsymbol{C_e} = diag([\sigma_0^2, \sigma_1^2, \cdots, \sigma_M^2])$ and neglecting the constant terms, the log-likelihood function $L(\boldsymbol{G}, \boldsymbol{C_e})$ can be expressed as

$$\begin{aligned} L(\boldsymbol{G}, \boldsymbol{C_e}) &= -N \log det(\boldsymbol{C_e}) + \sum_{n=1}^{N} (-\boldsymbol{e}_H(n)\boldsymbol{C_e}^{-1}\boldsymbol{e}(n)) \\ &= -N \sum_{m=0}^{M} \log \sigma_m^2 - \sum_{n=1}^{N}\sum_{m=0}^{M} \sigma_m^{-2}|e_m(n)|^2, \end{aligned} \tag{4.7}$$

where $e_m(n) = x_m(n) - \boldsymbol{p}^T(t)\boldsymbol{g}_m$ is the m^{th} element of the $(M+1)$ dimensional vector $\boldsymbol{e}(n)$. The ML methodology is to maximize the likelihood or log-likelihood criterion with respect to the unknown parameters by using the given data. Thus, the problem

here is to maximize (4.7) with respect to $\boldsymbol{G}$ and $\boldsymbol{C_e}$. For a fixed $\boldsymbol{B}$, (4.7) can be maximized with respect to σ_m^2 $(m = 0, 1, \cdots, M)$ by setting to zero the partial derivative,

$$\frac{\partial L(\boldsymbol{G}, \boldsymbol{C_e})}{\partial \sigma_m^2} = -N\sigma_m^{-2} + \sum_{n=1}^{N} (\sigma_m^2)^{-2} |e_m(n)|^2 = \sigma_m^{-2}(\sigma_m^{-2} \sum_{n=1}^{N} |e_m(n)|^2 - N), \tag{4.8}$$

leading to an estimate of σ_m^2

$$\hat{\sigma}_m^2 = \frac{\sum\limits_{n=1}^{N} |e_m(n)|^2}{N} = \frac{\boldsymbol{x}_m - \boldsymbol{P}^T \boldsymbol{g}_m}{N}, \tag{4.9}$$

where the received data of the m^{th} sensor $\boldsymbol{x}_m = [x_m(1), x_m(2), \cdots, x_m(N)]$ is the $N \times 1$ vector corresponding to the m^{th} row of the matrix $\boldsymbol{Y}$ and $\boldsymbol{P} = [\boldsymbol{p}(1), \boldsymbol{p}(2), \cdots, \boldsymbol{p}(N)]$ is the $K \times N$ matrix according to the known waveform. Using (4.9) into (4.7), we have

$$\begin{aligned} \bar{L}(\boldsymbol{G}) &= -N \sum \log \frac{\sum\limits_{n=1}^{N} |e_m(n)|^2}{N} - \sum_{n=1}^{N} \sum_{m=0}^{M} \frac{N|e_m(n)|^2}{\sum\limits_{n=1}^{N} |e_m(n)|^2} \\ &= -N \sum_{m=0}^{M} \log \frac{\boldsymbol{x}_m - \boldsymbol{P}^T \boldsymbol{g}_m}{N} - (M+1)N, \end{aligned} \tag{4.10}$$

By ignoring the constant term in the above equation, the ML estimate of $\boldsymbol{G}$ can be described as

$$\hat{\boldsymbol{G}} = \arg\max_{\boldsymbol{G}} (-N \sum_{m=0}^{M} \log \frac{\|\boldsymbol{x}_m - \boldsymbol{P}^T \boldsymbol{g}_m\|^2}{N}). \tag{4.11}$$

Since the logarithm is a monotonic function, the above maximization problem is equivalent to the following minimization problem:

$$\hat{\boldsymbol{G}} = \arg\min_{\boldsymbol{G}}(\sum_{m=0}^{M} \|\boldsymbol{x}_m - \boldsymbol{P}^T\boldsymbol{g}_m\|^2). \tag{4.12}$$

It is of interest to note that with known waveforms of the signals, the ML estimate for $\boldsymbol{G}$ in (4.13) can be represented by $M+1$ LS individual solutions for the $(M+1)$ linear regression models of order K, namely,

$$\hat{\boldsymbol{g}}_m = \arg\min_{\boldsymbol{g}_m} \|\boldsymbol{x}_m - \boldsymbol{P}^T\boldsymbol{g}_m\|^2, (m = 0, 1, \cdots, M). \tag{4.13}$$

If the signals are not coherent and there are sufficient snapshots for each interval such that matrix $\boldsymbol{P}^*\boldsymbol{P}^T$ is nonsingular, the solution to (4.13) is then given by

$$\hat{\boldsymbol{g}}_m = \boldsymbol{R}_{pp}^{-1}\boldsymbol{R}_{px}^{(m)}, \tag{4.14}$$

where

$$\boldsymbol{R}_{pp} \triangleq \frac{\sum_{n=1}^{N}[\boldsymbol{p}^*(n)\boldsymbol{p}^T(n)]}{N} = \frac{\boldsymbol{P}^*\boldsymbol{P}^T}{N} \tag{4.15}$$

and

$$\boldsymbol{R}_{px}^{(m)} \triangleq \frac{\sum_{n=1}^{N}[\boldsymbol{p}^*(n)x(n)]}{N} = \frac{\boldsymbol{P}^*\boldsymbol{x}_m}{N}. \tag{4.16}$$

Note that when the signals are highly correlated with one another or fewer snapshots are available, $\boldsymbol{R}_{pp}$ would be singular or nearly singular. Then, some modified methods

such as in [100] can be adopted to deal with this problem. It is also noted that the previous methods [53–57, 59] have made use of the known waveforms to obtain the spatial signatures for the estimation of parameters such as DOAs and complex gains with the traditional subspace techniques such as the ML, MUSIC and ESPRIT. In this chapter, however, we make use of the LS-based technique, which employs the orthogonal projection matrix of the known waveforms to obtain the coefficients in each sensor, which contains the entire information of DOAs and complex gains. Therefore, the next subsection is devoted to the regression analysis of these coefficients.

4.2.3 Regression Analysis

Using (4.2) and setting $\boldsymbol{g}_m \triangleq \hat{\boldsymbol{g}}_m - \triangle\boldsymbol{g}_m$, where $\triangle\boldsymbol{g}_m = [\triangle b_1^m, \triangle b_2^m, \cdots, \triangle b_K^m]^T$ is the estimation error vector of the coefficients $\boldsymbol{g}_m$, we get

$$\boldsymbol{x}_m = \boldsymbol{P}^T\hat{\boldsymbol{g}}_m - \boldsymbol{P}^T\triangle\boldsymbol{g}_m + \boldsymbol{e}_m = \boldsymbol{P}^T\hat{\boldsymbol{g}}_m + \boldsymbol{\delta}_m = \hat{\boldsymbol{x}}_m + \boldsymbol{\delta}_m, \tag{4.17}$$

where $\boldsymbol{e}_m = [e_m(1), e_m(2), \cdots, e_m(N)]^T$ is the N-dimensional white Gaussian noise vector of the m^{th} sensor, $\boldsymbol{\delta}_m = -\boldsymbol{P}^T\triangle\boldsymbol{g}_m + \boldsymbol{e}_m = [\delta_m(1), \delta_m(2), \cdots, \delta_m(N)]^T$ is defined as the LS residuals, and $\hat{\boldsymbol{x}} = \boldsymbol{P}^T\hat{\boldsymbol{g}}_m$ is the estimate of $\boldsymbol{x}_m$. From (4.17), we obtain

$$\triangle\boldsymbol{g}_m = (\boldsymbol{P}^*\boldsymbol{P}^T)^{-1}\boldsymbol{P}^*(\boldsymbol{e}_m - \boldsymbol{\delta}_m). \tag{4.18}$$

On the other hand, from (4.14) and (4.17), we obtain

$$\boldsymbol{\delta}_m = \boldsymbol{x}_m - \boldsymbol{P}^T\boldsymbol{R}_{pp}^{-1}\boldsymbol{R}_{px}^{(m)} = (\boldsymbol{I}_N - \boldsymbol{P}^T(\boldsymbol{P}^*\boldsymbol{P}^T)^{-1}\boldsymbol{P}^*)\boldsymbol{x}_m \triangleq (\boldsymbol{I}_N - P_{\boldsymbol{P}^T})\boldsymbol{e}_m. \tag{4.19}$$

Then, substituting (4.19) into (4.18) and recalling that $\boldsymbol{e}(n) = [e_0(n), e_1(n), \cdots, e_M(n)]^T$ has zero mean, the mean of the estimation error $\triangle \boldsymbol{g}_m$ is easily obtained as

$$E[\triangle \boldsymbol{g}_m] = (\boldsymbol{P}^* \boldsymbol{P}^T)^{-1} \boldsymbol{P}^* E[\boldsymbol{e}_m - (\boldsymbol{I}_N - P_{\boldsymbol{P}^T})\boldsymbol{e}_m] = (\boldsymbol{P}^* \boldsymbol{P}^T)^{-1} \boldsymbol{P}^* E[\boldsymbol{e}_m] = \boldsymbol{0}, \quad (4.20)$$

implying that the estimate $\hat{\boldsymbol{g}}_m$ is unbiased. It is easy to show that the covariance matrix of the estimate $\hat{\boldsymbol{g}}_m$ is given by

$$\boldsymbol{C}_{\hat{\boldsymbol{g}}_m} = E[\triangle \boldsymbol{g}_m \triangle \boldsymbol{g}_m^H] = (\boldsymbol{P}^* \boldsymbol{P}^T)^{-1} \boldsymbol{P}^* E[\boldsymbol{e}_m \boldsymbol{e}_m^H] \boldsymbol{P}^T (\boldsymbol{P}^* \boldsymbol{P}^T)^{-1} = \sigma_m^2 (\boldsymbol{P}^* \boldsymbol{P}^T)^{-1}. \quad (4.21)$$

In obtaining (4.22), we have used the fact that $E[\boldsymbol{e}_m \boldsymbol{e}_m^H] = \sigma_m^2 \boldsymbol{I}_N$. By using (4.14), the above equation can be rewritten as

$$\boldsymbol{C}_{\hat{\boldsymbol{g}}_m} = \frac{\sigma_m^2 \boldsymbol{R}_{pp}^{-1}}{N}. \quad (4.22)$$

Clearly, from (4.22), the covariance matrix $\boldsymbol{C}_{\hat{\boldsymbol{g}}_m}$ is inversely proportional to the number of snapshots. When N approaches infinity, $\boldsymbol{C}_{\hat{\boldsymbol{g}}_m}$ approaches zero. Note that when the incident signals are uncorrelated with one another in the time period, i.e., the elements in (4.15) satisfy $\sum_{n=1}^{N} p_i^*(n) p_j(n) = 0$, for $i \neq j$, and $i, j \in [1, 2, \cdots, K]$, we have $\boldsymbol{R}_{pp} = diag([\rho_1, \rho_2, \cdots, \rho_K])$, where $\rho_k, (k = 1, 2, \cdots, K)$ is the power of the k^{th} known waveform, since the estimation error $\triangle b_k^{(m)}$ of $b_k^{(m)}$ has a zero-mean white Gaussian distribution with variance $Var(\triangle g_k^{(m)}) = \frac{1}{N \cdot WNR_k^{(m)}}, WNR_k^{(m)} = \frac{\rho_k}{\sigma_m^2}$ being the waveform to noise ratio (WNR) of the k^{th} signal on the m^{th} sensor. It is also clear from (4.22) that when the incident signals are coherent, the covariance matrix

$\boldsymbol{C}_{\hat{\boldsymbol{g}}_m}$ becomes singular and thus, we cannot obtain an accurate estimate of $\boldsymbol{g}_m$ and need some modifications to recover the full rank of $\boldsymbol{R}_{pp}$. When the signals are partly correlated with one another, the estimation error $\triangle g_k^{(m)}$ of $g_k^{(m)}$ has a zero-mean white Gaussian distribution with variance $Var(\triangle g_k^{(m)}) = \frac{\sigma_m^2[\boldsymbol{R}_{pp}^{-1}]_{kk}}{N}$. As the variance of $\hat{b}_k^{(m)}$ depends on the unknown noise variance σ_m^2 of the m^{th} sensor, we show below how to get an estimate of σ_m^2. Using (4.19), the square sum of the LS residuals $\boldsymbol{\delta}_m$ is given by

$$\boldsymbol{\delta}_m^H \boldsymbol{\delta}_m = \boldsymbol{e}_m^H (\boldsymbol{I}_N - P_{\boldsymbol{P}^T})(e)_m = Tr(P_{\boldsymbol{P}^T}^{\perp} \boldsymbol{e}_m \boldsymbol{e}_m^H). \tag{4.23}$$

Thus, the expectation of (4.23) is derived as

$$E[\boldsymbol{\delta}_m^H \boldsymbol{\delta}_m] = E[Tr(P_{\boldsymbol{P}^T}^{\perp} \boldsymbol{e}_m \boldsymbol{e}_m^H)] = (N-K)\sigma_m^2. \tag{4.24}$$

where we have used the property of an idempotent matrix [101] that the trace equal to the rank of the matrix, namely, $Tr(Tr(P_{\boldsymbol{P}^T}^{\perp}) = Rank(\boldsymbol{I}_N) - Rank(P_{\boldsymbol{P}^T})$. Finally, we can obtain an unbiased ML estimate of σ_m^2 from

$$\hat{\sigma}_m^2 = \frac{\|\boldsymbol{x}_m - \boldsymbol{P}^T \hat{g}_m + \boldsymbol{P}^T \triangle \boldsymbol{g}_m\|^2}{N} = \frac{\sum_{n=1}^{N} \delta_m(n) \delta_m * (n)}{N-K}, \tag{4.25}$$

where the number $N-K$ is called the degree of freedom (DOF), which is the number of snapshots minus the number of source signals. Note that when the number of snapshots is much larger than the number of source signals, which is quite common in array signal processing, the ML estimate of σ_m^2 approaches the average of the squared sum of the residuals (errors) $\delta_m(n), n = 1, 2, \cdots, N$. Thus far, we have employed the

regression analysis to obtain the distribution of the coefficients $\boldsymbol{g}_m$, which contain the whole information of the DOA and complex gain. In the next subsection, we will present an optimal estimator for the DOA and complex gain.

4.2.4 DOA and Complex Gain Estimation

Recalling that $\boldsymbol{g}_m \triangleq \boldsymbol{\Xi}\boldsymbol{b}_m = \hat{\boldsymbol{g}}_m - \triangle\boldsymbol{b}_m$, we have

$$\hat{g}_k^{(m)} = \varepsilon_k e^{-j\varphi_k^{(m)}} + \triangle g_k^{(m)}, (m = 1, 2, \cdots, M), \tag{4.26}$$

$$\hat{g}_k^{(0)} = \varepsilon_k + \triangle g_k^{(0)}, \tag{4.27}$$

where $\triangle g_k^{(m)} \triangleq \mathrm{Re}(\triangle g_k^{(m)}) + j\mathrm{Im}(\triangle g_k^{(m)})$ is the estimation error, which is described by a complex Gaussian process with zero mean and variance of $\frac{\sigma_m^2[\boldsymbol{R}_{pp}^{(-1)}]_{k,k}}{N}$. Note that $\mathrm{Re}(\triangle g_k^{(m)})$ and $\mathrm{Im}(\triangle g_k^{(m)})$ are uncorrelated with each other, each having zero mean and a variance $\frac{\sigma_m^2[\boldsymbol{R}_{pp}^{(-1)}]_{k,k}}{2N}$(see Appendix D for proof). When both sides of (4.26) are multiplied by the complex conjugate of (4.27) and then divided by the complex gain power $|\varepsilon_k|^2$, we get

$$\hat{a}_k^{(m)} \triangleq \frac{\hat{g}_k^{(m)}\widehat{g}_k^{(0)*}}{|\varepsilon_k|^2} = e^{-j\varphi_k^{(m)}} + \triangle a_k^{(m)}, \tag{4.28}$$

which can be considered as the estimated value of the m^{th} element of the k^{th} steering vector $\boldsymbol{a}(\theta_k)$, where $\triangle a_k^{(m)}$ is the estimation error with zero mean, variance $\mathrm{Var}\left(\triangle a_k^{(m)}\right) = \frac{(\sigma_m^2+\sigma_0^2)\left[\mathbf{R}_{pp}^{-1}\right]_{k,k}}{N|\varepsilon_k|^2}$ and covariance $\mathrm{Cov}(\triangle a_k^{(p)}, \triangle a_k^{(q)}) = \frac{\sigma_0^2\left[\mathbf{R}_{pp}^{-1}\right]_{k,k}}{N|\varepsilon_k|^2}, p \neq q \in [1, \cdots, M]$ (see Appendix E for detail). Note that the unknown power gain $|\varepsilon_k|^2$ has no effect on the estimation of DOA due to the fact that only the phase

of $\hat{a}_k^{(m)}$ is used to obtain the DOA. As a result, the estimated error vector $\triangle\hat{\boldsymbol{a}}_k = \left[\triangle\hat{a}_k^{(1)}, \triangle\hat{a}_k^{(2)}, \cdots, \triangle\hat{a}_k^{(M)}\right]^T, (k = 1, \cdots, K)$, has a distribution $\mathbb{CN}(\mathbf{0}, \boldsymbol{\Omega}_k)$ with

$$\boldsymbol{\Omega}_k = \eta_k^{(0)} \begin{bmatrix} \eta_k^{(1)} & & \mathbf{1}_{M-1}^T \\ \mathbf{1}_1 & \eta_k^{(2)} & \mathbf{1}_{M-2}^T \\ \cdots & \cdots & \cdots \\ \mathbf{1}_{M-1}^T & & \eta_k^{(M)} \end{bmatrix}, \tag{4.29}$$

and $\eta_k^{(0)} \triangleq \frac{\sigma_0^2 \left[\boldsymbol{R}_{pp}^{-1}\right]_{k,k}}{N|\varepsilon_k|^2}$ and $\eta_k^{(m)} = \frac{\sigma_m^2}{\sigma_0^2} + 1, (m = 1, 2, \cdots, M)$, and $\mathbf{1}_m$ denotes a m-dimensional column vector whose elements are all unity. If each sensor has the uniform additive noise distribution with the same variance, which is often assumed in conventional methods, (4.29) can be simplified as

$$\boldsymbol{\Omega}_k = \eta_k^{(0)} \begin{bmatrix} 2 & & \mathbf{1}_{M-1}^T \\ \mathbf{1}_1 & 2 & \mathbf{1}_{M-2}^T \\ \cdots & \cdots & \cdots \\ \mathbf{1}_{M-1}^T & & 2 \end{bmatrix}. \tag{4.30}$$

Next, let us introduce M-dimensional vectors,

$$\boldsymbol{\Phi}_k = \left[\varphi_k^{(1)}, \varphi_k^{(2)}, \ldots, \varphi_k^{(M)}\right]^T. \tag{4.31}$$

Although each element in (4.31) can be estimated from (4.28) by calculating the phase of $a_k^{(m)}$, when $d_1 > 1$, ambiguity may arise due to its inability to deal with the

2π ambiguity, i.e., $\varphi_k^{(m)} = -\angle\left(a_k^{(m)}\right)$, $\left|\varphi_k^{(m)}\right| > \pi$, where the sign $\angle x$ denotes the real phase of x . If the assumption $d_1 \leq 1$ is made, there is no ambiguity for $\varphi_k^{(m)}$, i.e., $\left|\varphi_k^{(1)}\right| = d_1\pi\,|\sin\theta_k| \leq \pi$. Therefore, we can use $\varphi_k^{(m)}$ as a reference to deal with potential ambiguities for $\sum_{i=2}^{m} d_i \geq 1, m = 2, \cdots, M$ by exploiting the disambiguity method proposed in Chapter 2 of this thesis. Herein, we introduce a new method to solve the problem of ambiguity for the case $d_1 > 1$. Let us write the measured value of the k^{th} element of $\mathbf{\Phi}_k$ as

$$\mod\,(\varphi_k^{(m)}) = \varphi_k^{(m)} - 2\pi l_k^{(m)}, \tag{4.32}$$

where $l_k^{(m)} (m = 1, 2, \cdots, M)$ are unknown integers. Now from (4.32), it can be readily shown that

$$\begin{aligned} \mod(\nabla(\,\mod(\varphi_k^{(m)}))) &= \mod((\varphi_k^{(m+1)} - 2\pi l_k^{(m+1)}) - (\varphi_k^{(m)} - 2\pi l_k^{(m)})) \\ &= \nabla(\varphi_k^{(m)}) + 2\pi(\tau_k^{(m)} - \nabla(l_k^{(m)})), \end{aligned} \tag{4.33}$$

where the operator ∇ is defined as $\nabla(x_k^{(m)}) = x_k^{(m+1)} - x_k^{(m)}$. Also, if $-\pi \leq \nabla(\varphi_k^{(m)}) < \pi$, $\tau_k^{(m)}$ satisfies the relationship $\tau_k^{(m)} = \nabla(l_k^{(m)})$. Hence, from (4.33) we get

$$\varphi_k^{(m+1)} = \varphi_k^{(m)} + \mod(\nabla(\mod(\varphi_k^{(m)}))) \tag{4.34}$$

or

$$\varphi_k^{(m+1)} = \varphi_k^{(1)} + \sum_{i=2}^{m} \mod(\nabla(\mod(\varphi_k^{(m)}))) \tag{4.35}$$

Equation (4.35) states that the measured value of $\varphi_k^{(m+1)}$ can be obtained using the first phase difference $\varphi_k^{(1)}$ and some simple operations. If d_1>1, however, we cannot get an unique estimate for $\varphi_k^{(1)}$. Now we show how to get the actual value of $\varphi_k^{(1)}$. From (4.32), we obtain

$$\varphi_k^{(1)}=2\pi l_k^{(1)}+\text{mod}(\varphi_k^{(1)}) =(2\pi l_k^{(2)}+\text{mod}(\varphi_k^{(2)}))d_1\Big/d_2, \left|l_k^{(1)}\right| \leq \text{fix}(d_1/2)\text{and} \left|l_k^{(2)}\right| \leq \text{fix}(d_2/2) \tag{4.36}$$

where fix(x) rounds x to the nearest integer towards zero. To choose the correct $l_k^{(1)}$ and $l_k^{(2)}$, we use the principle suggested in [102] to minimize the following cost function when d_1 and d_2 are co-prime numbers.

$$F= \min_{l_k^{(1)},l_k^{(2)}} \left|2\pi(d_2 l_k^{(1)}-d_1 l_k^{(2)}) + (d_2\text{mod}(\varphi_k^{(1)}) - d_1\text{mod}(\varphi_k^{(2)}))\right| \tag{4.37}$$

After obtaining the actual $l_k^{(1)}$, it is easy to get $\varphi_k^{(1)}$ and $\varphi_k^{(2)}$ using (4.36), then using (4.35) to obtain $\varphi_k^{(m)}, m = 3, \cdots, M$. From the above discussions, the unbiased estimate $\hat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$ is a Gaussian random vector with the covariance matrix as given by (See the proof in Appendix F)

$$\mathbf{\Sigma}_k \triangleq \text{Cov}\left(\hat{\mathbf{\Phi}}_k\right) \approx \mathbf{\Omega}_k/2. \tag{4.38}$$

Therefore, the MLE of the parameter $\psi_k = \pi\cos\theta$ is equivalent to the minimum variance unbiased estimator (MVUE) [103] of the vector $\mathbf{\Phi}_k$ in (4.31) when the estimated error has a Gaussian distribution, which can be estimated by the method of generalized LS [104]. The parameter ψ_k can be obtained by minimizing the square

error

$$\hat{\psi}_k = \arg\min_{\psi} \Upsilon_k(\psi), \Upsilon_k(\psi) = \left(\hat{\mathbf{\Phi}}_k - \psi \boldsymbol{d}\right)^T \mathbf{\Sigma}_k^{-1} \left(\hat{\mathbf{\Phi}}_k - \psi \boldsymbol{d}\right), \tag{4.39}$$

where $\boldsymbol{d} = \left[d_1, \sum_{i=1}^{2} d_i, \cdots, \sum_{i=1}^{M} d_i\right]^T$. The solution to the above problem is given by

$$\hat{\psi}_k = \frac{\boldsymbol{d}^T \mathbf{\Sigma}_k^{-1} \hat{\mathbf{\Phi}}_k}{\boldsymbol{d}^T \mathbf{\Sigma}_k^{-1} \boldsymbol{d}}. \tag{4.40}$$

Also, it is easy to get the variance of this estimate as

$$\mathrm{Var}\left(\hat{\psi}_k\right) = \frac{\boldsymbol{d}^T \mathbf{\Sigma}_k^{-1} \mathrm{Cov}\left(\hat{\mathbf{\Phi}}_k\right) \mathbf{\Sigma}_k^{-1} \boldsymbol{d}}{\boldsymbol{d}^T \mathbf{\Sigma}_k^{-1} \boldsymbol{d}} = \frac{1}{\boldsymbol{d}^T \mathbf{\Sigma}_k^{-1} \boldsymbol{d}}. \tag{4.41}$$

Note that $\hat{\psi}_k$ is an unbiased estimate with a Gaussian distribution, since $\triangle \mathbf{\Phi}_k \triangleq \mathbf{\Phi}_k - \hat{\mathbf{\Phi}}_k$ is a Gaussian random vector with zero mean. From the expression $\psi_k = \pi \sin\theta_k$, we can get the estimated DOA $\hat{\theta}_k = \arcsin(\hat{\psi}_k/\pi)$. Also the variance of $\hat{\theta}_k$ can be obtained as

$$\mathrm{Var}\left(\hat{\theta}_k\right) = \mathrm{Var}\left(\hat{\psi}_k\right) \Big/ (\pi \cos\theta_k)^2 \tag{4.42}$$

We now present an estimator for the complex gain. Obviously, we could estimate the complex gain using (4.27) directly, but it is not optimal estimator. Here, we make use of the estimated DOA to obtain the optimal estimator for the complex gain. We define

$$\hat{\mathbf{\Xi}}_k = \left[\hat{\varepsilon}_k^{(1)}, \hat{\varepsilon}_k^{(2)}, \cdots, \hat{\varepsilon}_k^{(M)}\right]^T \tag{4.43}$$

where $\hat{\varepsilon}_k^{(m)} = \hat{g}_k^{(m)} e^{j\hat{\varphi}_k^{(m)}} \triangleq \varepsilon_k^{(m)} + \triangle\varepsilon_k^{(m)}, (m= 1, 2, \cdots, M)$ and $\triangle\varepsilon_k^{(m)}$ are zero mean Gaussian random variables with the covariance matrix $\boldsymbol{H}_k$ (see Appendix G for derivation). In a manner similar to deriving (4.39)-(4.41), the optimal estimator for ε_k can be obtained as

$$\hat{\varepsilon}_k = \mathbf{1}_M^T \mathbf{H}_k^{-1} \hat{\boldsymbol{\Xi}}_k \Big/ \mathbf{1}_M^T \mathbf{H}_k^{-1} \mathbf{1}_M. \tag{4.44}$$

Also, it is easy to get the variance of this estimator as

$$\mathrm{Var}\left(\hat{\varepsilon}_k\right) = 1 \Big/ \mathbf{1}_M^T \mathbf{H}_k^{-1} \mathbf{1}_M. \tag{4.45}$$

Before the end of this section, we would like to consider a special case where the source signals are uncorrelated with equal-power, and the uniform linear array (ULA) is of half wavelength inter-element spacing, i.e., $\boldsymbol{d} = [1, 2, \ldots, M]^T$. Note that the additive noises at the ULA are i.i.d Gaussian random processes with $SNR = \frac{\rho}{\sigma^2}|\varepsilon_k|^2 = WNR|\varepsilon_k|^2$, the covariance matrix is then given by

$$[\boldsymbol{\Sigma}_k]_{p,q} = \begin{cases} \frac{1}{N \cdot SNR} & 1 \leq p = q \leq M \\ \frac{1}{2N \cdot SNR} & \text{otherwises,} \end{cases} \tag{4.46}$$

which can be rewritten in the matrix form as

$$\boldsymbol{\Sigma}_k = \frac{\boldsymbol{\Lambda}}{2N \cdot SNR}, \tag{4.47}$$

where $\mathbf{\Lambda} = \boldsymbol{I}_M + \mathbf{1}_M \mathbf{1}_M^T$. By using the Sherman-Morrison formula [105], we have

$$\mathbf{\Lambda}^{-1} = \boldsymbol{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{1+M}. \tag{4.48}$$

Thus

$$\mathbf{\Sigma}_k^{-1} = 2N \cdot SNR \left(\boldsymbol{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{1+M} \right) \tag{4.49}$$

and

$$\boldsymbol{d}^T \mathbf{\Sigma}_k^{-1} \boldsymbol{d} = \frac{(M^2+M)(M+2)}{6} N \cdot SNR \tag{4.50}$$

Hence, the variance of $\hat{\theta}_k$ with known noise covariance matrix can be calculated from (4.41), (4.51) and (4.50) as

$$\mathrm{Var}\Big(\hat{\theta}_k\Big)_{KNOWN} = \frac{6}{(M^2+M)(M+2)N \cdot SNR} \left[\frac{1}{\pi \cos\theta_k} \right]^2. \tag{4.51}$$

Interestingly, using the Cramer-Rao bound (CRB) expression in [53, 54], we can find that (4.51) is identical to CRB, i.e., $\mathrm{CRB}(\theta_k) = \mathrm{Var}\Big(\hat{\theta}_k\Big)_{KNOWN}$ .Therefore, our estimator is a MVUE estimator when the additive noise has i.i.d Gaussian distribution. On the other hand, when the noise covariance matrix is unknown, we can also get the variance by the result of (4.25), namely,

$$\mathrm{Var}\Big(\hat{\theta}_k\Big)_{UNKNOWN} = \frac{6}{(M^2+M)(M+2)(N-K)SNR} \left[\frac{1}{\pi \cos\theta_k} \right]^2. \tag{4.52}$$

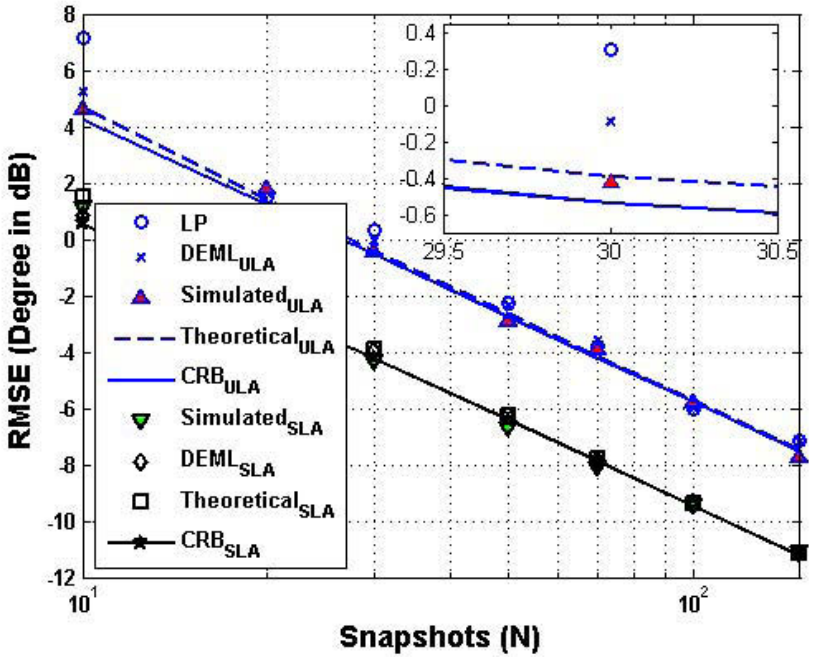
The CRB to variance ratio is then found from (4.51) and (4.52) as

$$\frac{\mathrm{Var}\Big(\hat{\theta}_k\Big)_{KNOWN}}{\mathrm{Var}\Big(\hat{\theta}_k\Big)_{UNKNOWN}} = 1 - \frac{K}{N}. \tag{4.53}$$

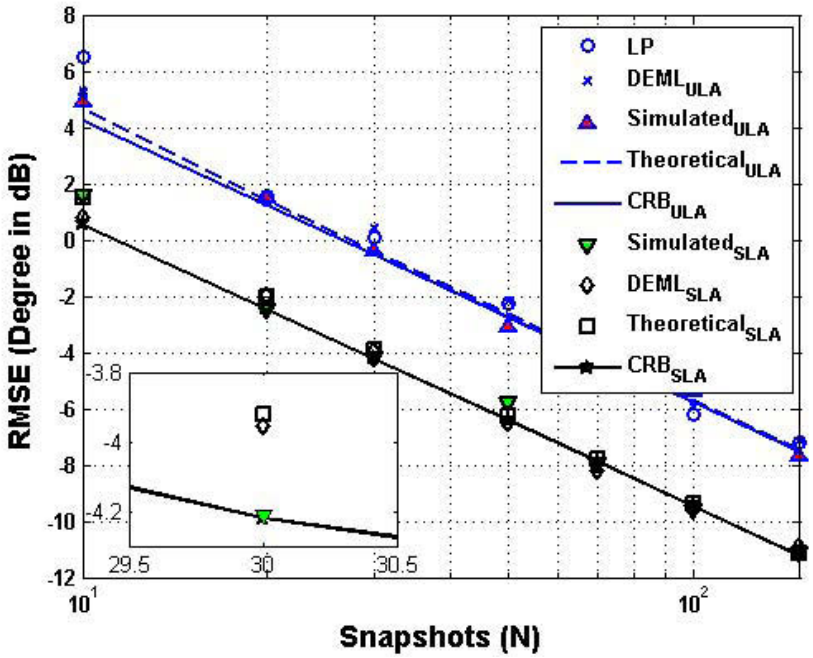
Usually, the number of snapshots is much larger than the number of signals, i.e., $N \gg K$, and therefore (4.53) implies that the performance of our method is almost identical to CRB.

4.2.5 Simulation Results for DOA Estimation

In this subsection, simulation results are presented to show the performance of the proposed DOA estimation techniques as compared to some of the existing methods. In the first two examples, two signals of equal power with angles $\theta_1 = -5^\circ$ and $\theta_2 = 5^\circ$ impinge onto the sparse linear array (SLA) with 3 sensors separated by $d_1 = 1$ and $d_2 = 2$. A 3-sensor ULA with a half wavelength antenna spacing is also considered for the performance study of the proposed method. The additive noise is uniform white noise, i.e., $\boldsymbol{R} = \sigma_0^2 \boldsymbol{I}_{M+1}$, and the complex gains are set to $\boldsymbol{\Xi} = diag(e^{j\pi/4}, e^{j\pi/4})$ and the SNR is defined as the ratio of the power of the source signal to that of the additive noise at each sensor, i.e., $SNR = \rho/\sigma_0^2$. Each example contains 1000 independent trials to obtain the root mean square error (RMSE). The LP method [59] and the DEML [54] along with the theoretical RMSE and CRB [53, 54] are plotted for performance comparison. In the last three examples, we evaluate the performance of our method for various angles of the source signals as well as for different numbers of sources and sensors.

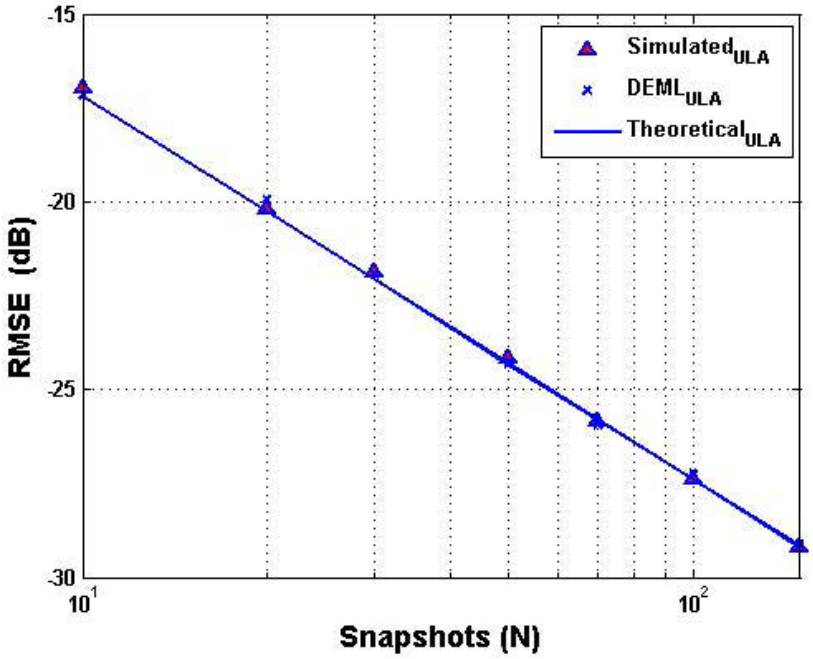


(a)

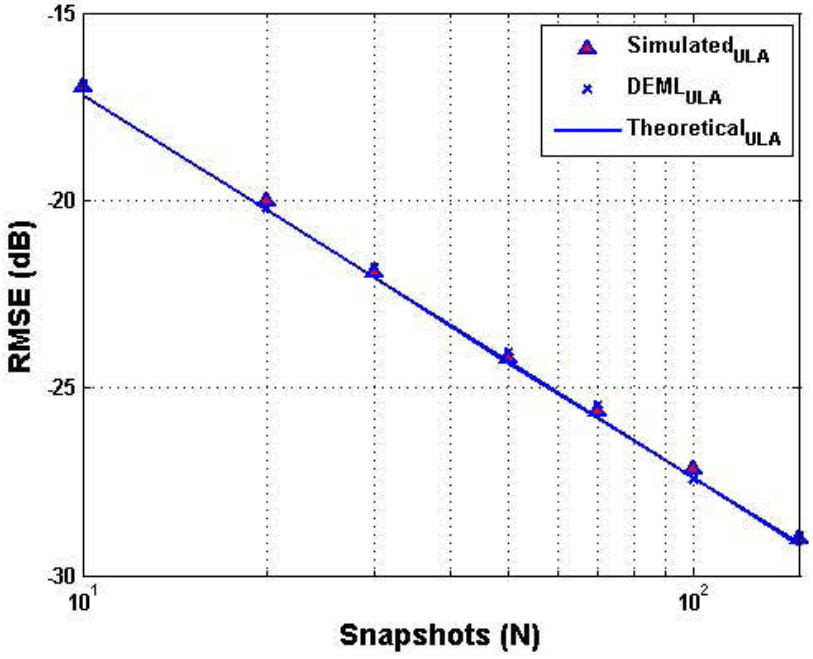


(b)

Figure 4.1: RMSE of DOA estimation for two uncorrelated sources versus snapshots from 10 to 150 at (a)$\theta_1 = -5°$ and (b)$\theta_2 = 5°$ with SNR=5dB.



(a)



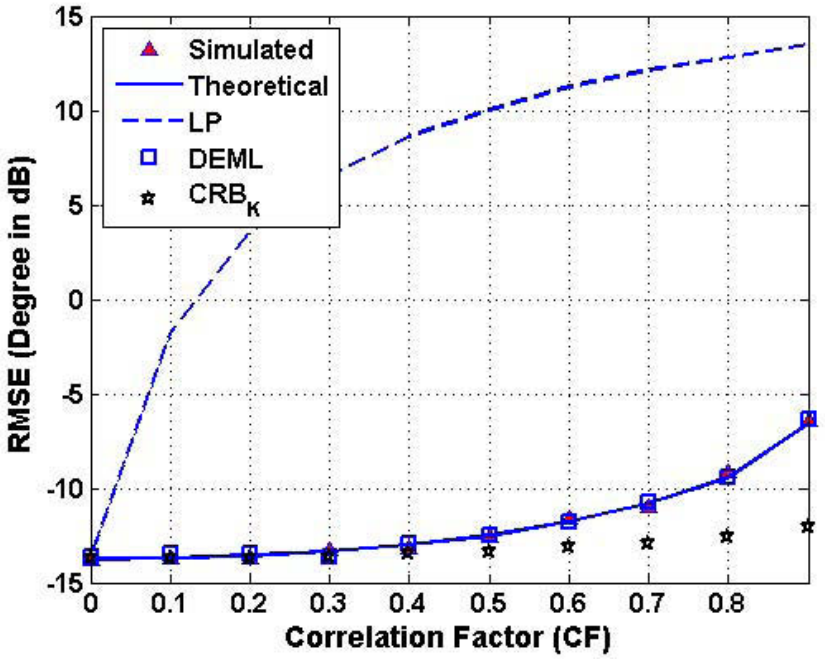
(b)

Figure 4.2: RMSE of complex gain estimation for two uncorrelated sources versus snapshots from 10 to 150 at (a)$\theta_1 = -5^\circ$ and (b)$\theta_2 = 5^\circ$ with SNR=5dB.

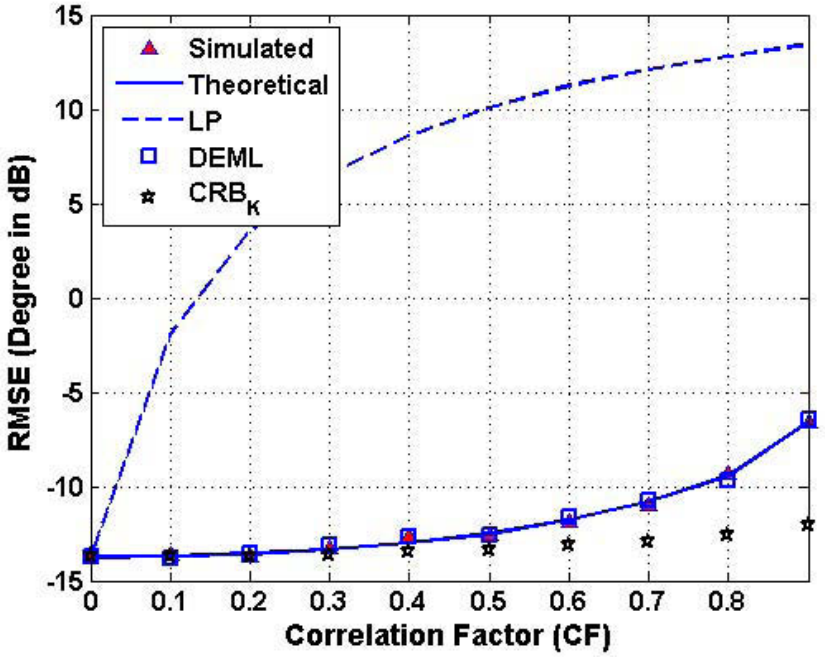
Example 1: *Performance of DOA and complex gain estimation with respect to number of snapshots*

The performance of the proposed DOA estimation method versus the number of snapshots is now assessed for two uncorrelated signals. The number of snapshots is varied from $N = 10$ to $N = 150$, and the SNR is set at 5dB. Both theoretical and simulated RMSEs of the estimated angles are plotted in Fig.4.1, along with the CRBs for ULA and SLA. We see that when the ULA is used, the proposed method yields an estimation performance similar to that of the LP and DEML methods, and the three methods are consistent with the theoretical RMSE and the CRBs, especially at a large number of snapshots. More importantly, the performance of the proposd method with SLA is much better than that of ULA due to the larger aperture of the array of SLA. Furthermore, we note that the simulated RMSEs of our method agree very well with the theoretical RMSEs in (4.51) and CRBs, which decrease almost linearly with the number of snapshots for both ULA and SLA. Note that in these figures, "DEML$_{SLA}$","Simulated$_{SLA}$", "Theoretical$_{SLA}$", and "CRB$_{SLA}$" denote the RMSE of the DEML, the simulated RMSE, and the theoretical RMSE of the proposed method and the CRB for the sparse linear array, while "DEML$_{ULA}$", "Simulated$_{ULA}$", "Theoretical$_{ULA}$", and "CRB$_{ULA}$" denote the corresponding results for the uniform linear array. In addition, the estimation performance of our method for complex gain is also compared with the DEML method shown in Fig.4.2, along with the theoretical performance for ULA. From the figures, we see that the RMSE curve of our method is identical to that of the DEML method and the theoretical results.

Example 2: *DOA estimation performance for correlated signals*



(a)



(b)

Figure 4.3: RMSE of DOA estimation for two sources versus correlation factor (CF) from 0 to 0.9 at (a)$\theta_1 = -5°$ and (b)$\theta_2 = 5°$ with 10dB SNR and 200 snapshots.

We study the performance of the proposed method against the correlation between two incident signals. The correlation factor (CF) between the two signals is varied from 0 to 0.9 with SNR=10dB and the number of snapshots is 200. Both simulated and theoretical RMSEs of the estimated angles against the CF are shown in Fig.4.3, together with the CRB. The performances of the LP method degrade severely at medium and strong correlation of the signals, while the RMSEs of the proposed method and the DEML is very close to the theoretical value and that of CRB, especially for weak and moderately correlated signals. The reason for this phenomenon lies in the fact that the spatial signatures of the LP method are obtained by the correlation matrix between the received data and the known waveforms, the orthogonal property among the different waveforms is destroyed and hence we cannot obtain the pure spatial signature corresponding to the waveform when the signals have medium to high correlation; while the regression coefficients of our method is obtained by making use of the LS-based technique in which the orthogonal projection matrix of the known waveforms and the spatial signatures of the DEML is based on the ML principle. Therefore, like the DEML method, our method is also very suitable for source signals with medium or high correlation with very less computational cost.

Example 3: *Performance of DOA estimation with respect to various angles*

In this example, we examine the performance of the proposed method against angles varying from -6° to 10° for different CFs, namely, CF=0, 0.5 and 0.9. The incident directions of the two sources are $\theta_1 = -6^\circ$ and θ_2 changing from -6° to 10° with SNR=5dB and snapshots=100. As shown in Fig.4.4, when CF is equal to zero, the proposed method and DEML have the same performance, being identical to the

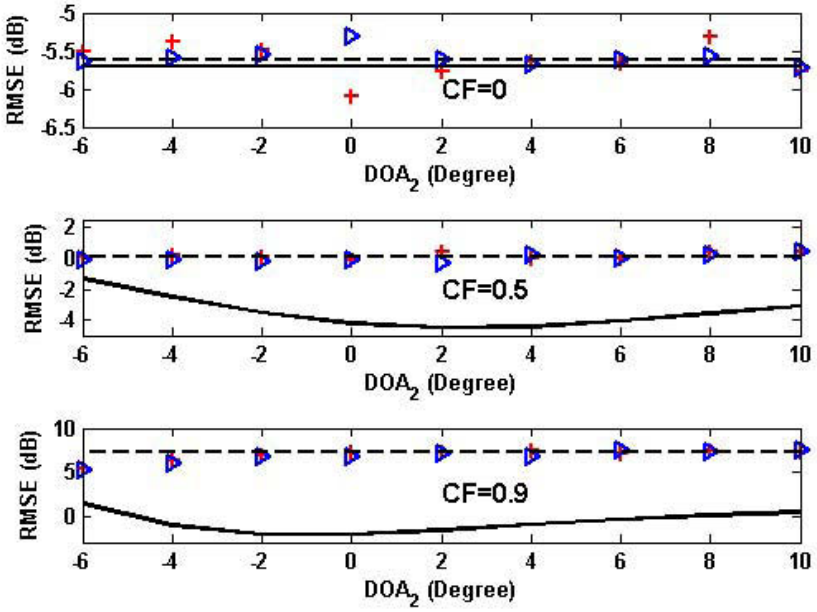
CRB. In addition, the performance curves of both the DEML and proposed method are the same as that of the theoretical results.

Example 4: *Performance of DOA estimation with respect to the number of sensors*

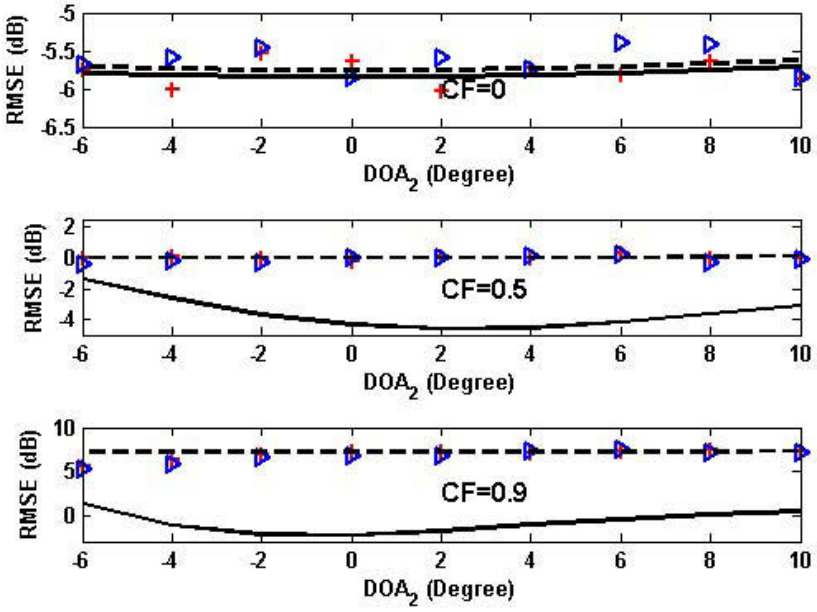
Now, the performance of the proposed method versus the number of sensors with the same array aperture is assessed. The assumptions for source signals and DOAs are similar to those in Example 1, and the number of snapshots is fixed at 100 and the SNR ranges from 0dB to 12dB. The configurations of SLA are set at 3 ($d_1 = 1$, $d_2 = 4$), 4 ($d_1 = 1$, $d_2 = 3$, $d_3 = 1$), 5 ($d_1 = 1$, $d_2 = 2$, $d_3 = 1$, $d_4 = 1$), and 6 ($d_1 = 1$, $d_2 = 1$, $d_3 = 1$, $d_4 = 1$, $d_5 = 1$). The simulated RMSEs of the estimated DOAs using the proposed method are plotted in Fig.4.5. It is clear that for the same array aperture, the more sensors used, the better performance can be obtained by our method. Note that here we also provide the theoretical result and CRB for the 6-sensor ULA.

Example 5: *Performance of DOA estimation with respect to the number of sources*

In the previous examples, the performance of the proposed method in estimating the directions of two sources based on three sensors ($d_1 = 1, d_2 = 5$) is tested. Here, we evaluate its estimation performance in terms of multiple uncorrelated sources: two sources (the same as in the previous examples), three sources $\big(\text{DOAs} = [-5^\circ \quad 5^\circ \quad 10^\circ], \boldsymbol{\Xi} = diag(e^{-j\pi/4}, e^{j\pi/4}, 1)\big)$, four sources (DOAs $= [-5^\circ \quad 5^\circ \quad 10^\circ \quad 15^\circ]$, $\boldsymbol{\Xi} = diag(e^{-j\pi/4}, e^{j\pi/4}, 1, -1)$) and five sources (DOAs $= [-5^\circ \quad 5^\circ \quad 10^\circ \quad 15^\circ \quad 20^\circ]$, $\boldsymbol{\Xi} = diag(e^{j\pi/4}, e^{-j\pi/4}, 1, -1, e^{j\pi/6})$). Fig.4.6 shows the performance of the proposed method as a function of the SNR for multiple sources at DOA $= -5^\circ$ and DOA $= 5^\circ$, respectively. We note that the proposed method is robust to the number

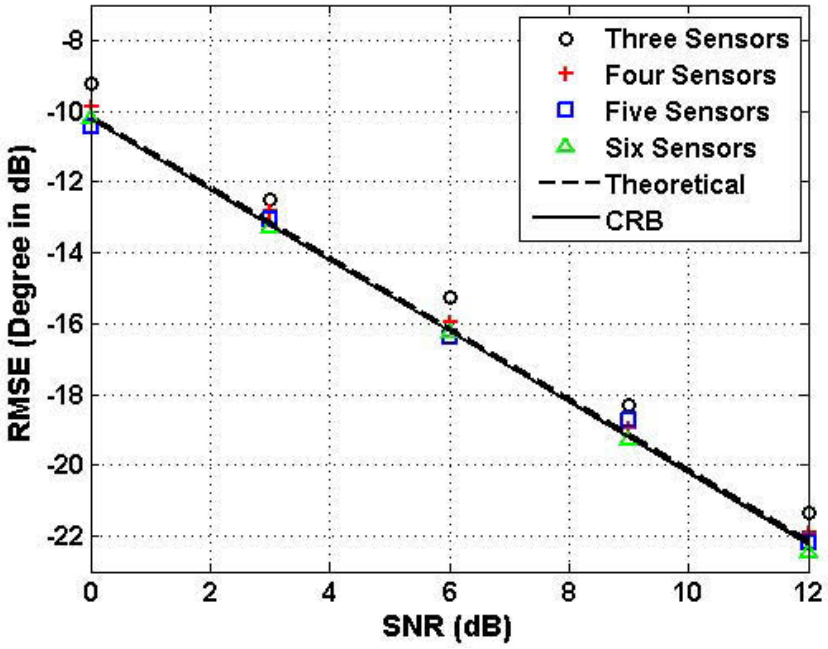


(a)

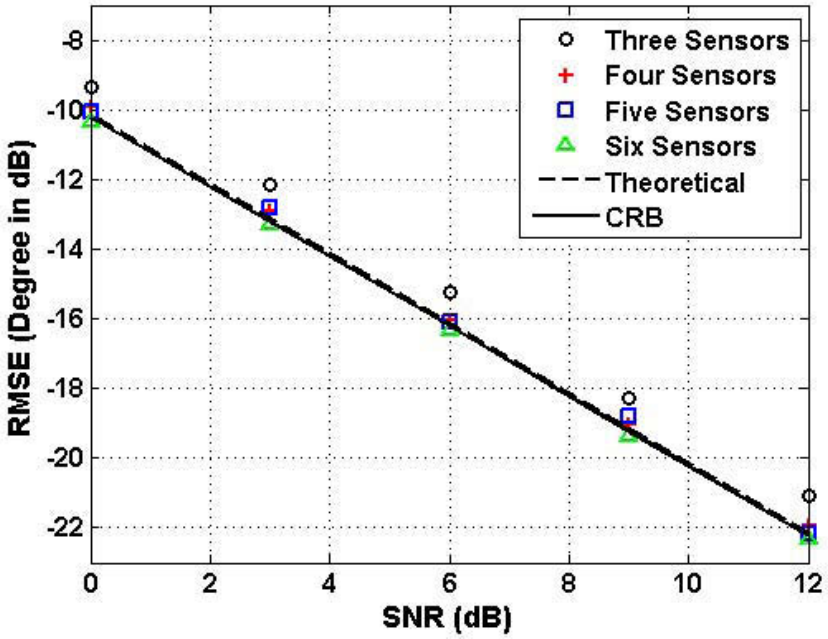


(b)

Figure 4.4: RMSE of DOA estimation for two sources with one DOA fixed at $-10°$ and the other DOA varying from $-6°$ to $-10°$ at (a) $-10°$ and (b) various angles with 5dB SNR and 100 snapshots, and CF =0, 0.5 and 0.9. Note that $+$, $>$,$--$ and $-$ stand for the DEML method, the proposed method, the theoretical result and the CRB, respectively.

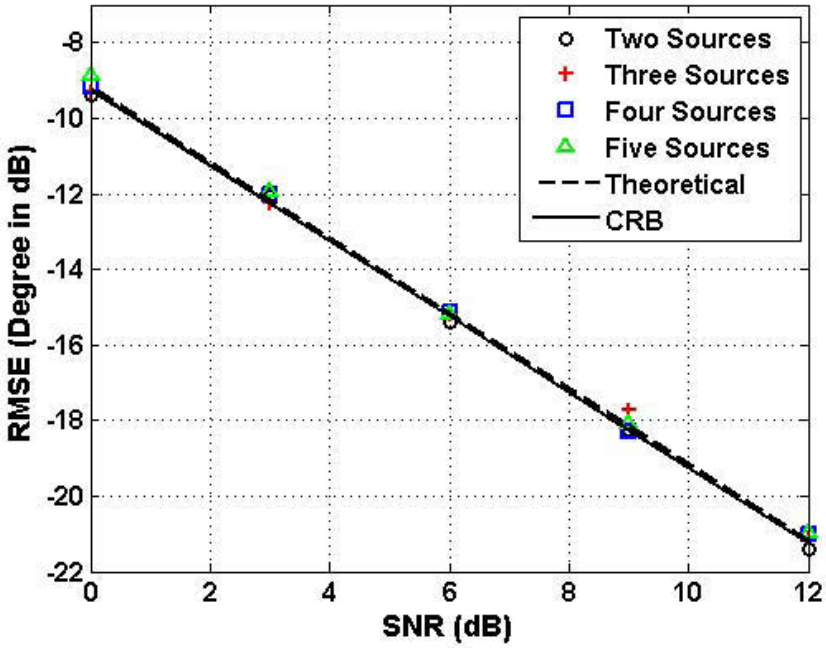


(a)

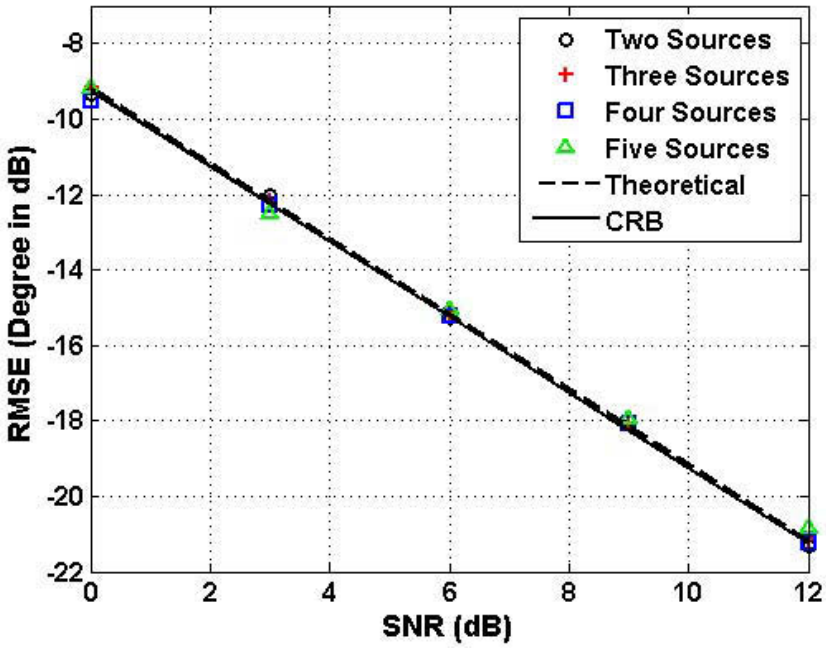


(b)

Figure 4.5: RMSE of DOA estimation for two sources using multiple sensors versus SNR from 0 to 12dB at (a)$\theta_1 = -5°$ and (b)$\theta_2 = 5°$ with 100 snapshots.



(a)



(b)

Figure 4.6: RMSE of DOA estimation for multiple sources versus SNR from 0 to 12dB at (a)$\theta_1 = -5^\circ$ and (b)$\theta_2 = 5^\circ$ with 100 snapshots.

of sources, even when the number of sources is more than the number of sensors, where most of the subspace-based methods with unknown waveforms do not work. Note that we also plot the curves of the theoretical results and CRB of the two-source situation for comparison.

4.3 DOA Tracking based on block RLS Techniques

In the previous section, we discussed the problem of DOA estimation for steady targets. As the targets move, their motion is tracked through a recursive algorithm which updates the estimate of target angles by using the data provided by the most recent output of the sensor array. We know that LS method is often used as a statistical procedure to fit the linear regression model. This is because when the noise is a Gaussian random process and the target is time-invariant, the LS method gives an estimation performance similar to that of the ML method, which is the best linear unbiased estimator (BLUE) for linear regression model. Herein, we derive two block-based LS algorithms to update the estimate over the time intervals based on the first time interval angles estimated by the method discussed in the previous section. In particular, one tracking method will be derived by using the block QRD-RLS technique, which operates directly on the data matrix through unitary transformations rather than on the corresponding covariance matrix, thus making the method superior to the traditional method with regard to numerical stability and computational complexity [91]. Unlike the usual RLS-based techniques, where the linear regression coefficients are updated at every snapshot, we perform the update on the basis of time

intervals. We use the block QRD-based technique to obtain current Cholesky factors from the previous Cholesky factors along with the received data of the time interval. Subsequently, we derive the second tracking method based on the regularized LS technique in view of the fact that the optimal weighting factor or regularization term can be derived by regression analysis [97,98], and the previously estimated coefficients can be considered as a constrained vector to deal with some "bad" cases such as small snapshots in a time interval or highly correlated sources.

4.3.1 Data Model

In this section, we still use the array system suggested in Section 4.2 except that we assume both $\boldsymbol{\theta}(t)$ and the corresponding gain $\boldsymbol{\Xi}(t)$ are varying with time. We also assume that over each time interval the changes for $\boldsymbol{\theta}(t)$ and $\boldsymbol{\Xi}(t)$ are small enough [71] so that

$$\boldsymbol{\theta}(t) \approx \boldsymbol{\theta}(nT), \quad \boldsymbol{\Xi}(t) \approx \boldsymbol{\Xi}(nT), t \in ((n-1)T, nT], (n = 1, 2, \cdots) \tag{4.54}$$

where each time interval T has N snapshots. Consequently, using the knowledge of the previous section, the $m^{th}, (m = 0, 1, 2, \cdots, M)$ element of (4.5) for the n^{th} time interval can be expressed as

$$\boldsymbol{x}_m(n) = \boldsymbol{P}(n)\boldsymbol{g}_m(n) + \boldsymbol{e}_m(n), \quad n = 1, 2, \cdots, \tag{4.55}$$

where $\boldsymbol{x}_m(n) = [x_m(n,1), x_m(n,2), \cdots, x_m(n,N)]^T$ is a $N \times 1$ vector, $\boldsymbol{P}(n) = [\boldsymbol{p}(n,1), \boldsymbol{p}(n,2), \cdots, \boldsymbol{p}(n,N)]^T$ is a $N \times K$ matrix according to the known waveforms of the n^{th} time interval, and $\boldsymbol{e}_m(n) = [e_m(n,1), e_m(n,2), \cdots, e_m(n,N)]^T$ is a $N \times 1$ vector of multivariate normal distribution with zero mean and covariance matrix $\sigma_m^2(n)\boldsymbol{I}_N$.

4.3.2 Block QRD-RLS technique

We now derive a block QRD-RLS technique to track the moving targets over time intervals. Let us denote the $nN \times K$ known waveform matrix as

$$\boldsymbol{P}_n \triangleq \left[\boldsymbol{P}^T(1), \boldsymbol{P}^T(2), \cdots, \boldsymbol{P}^T(n)\right]^T = \begin{bmatrix} \boldsymbol{P}_{n-1} \\ \boldsymbol{P}(n) \end{bmatrix}, \tag{4.56}$$

and the $nN \times 1$ received data vector of the m^{th} sensor as

$$\boldsymbol{x}_n^{(m)} \triangleq \left[\boldsymbol{x}_m^T(1), \boldsymbol{x}_m^T(2), \ldots, \boldsymbol{x}_m^T(n)\right]^T = \begin{bmatrix} \boldsymbol{x}_{n-1}^{(m)} \\ \boldsymbol{x}_m(n) \end{bmatrix}. \tag{4.57}$$

Further, we define an $nN \times 1$ LS residual vector as

$$\boldsymbol{e}_n^{(m)} \triangleq \left[\boldsymbol{e}_m^T(1), \boldsymbol{e}_m^T(2), \ldots, \boldsymbol{e}_m^T(n)\right]^T = \begin{bmatrix} \boldsymbol{e}_{n-1}^{(m)} \\ \boldsymbol{e}_m(n) \end{bmatrix}, \tag{4.58}$$

and an $nN \times nN$ block-diagonal weighted matrix as

$$\mathbf{M}_n^2 \triangleq \text{diag}\left[\mu^{N-1}\boldsymbol{\beta}, \mu^{N-2}\boldsymbol{\beta}, \cdots, \boldsymbol{\beta}\right], \tag{4.59}$$

where $\mu < 1$ is the forgetting factor for the time intervals and $\boldsymbol{\beta} \triangleq \text{diag}[\beta^{N-1}, \beta^{N-2}, \cdots, 1]$ with $\beta > \mu$ being the forgetting factor for snapshots. Since the targets are steady or slowly changing in the same time interval, β is very close to unity. Thus, the optimal estimator for the coefficients $\boldsymbol{g}_n^{(m)}, (m= 0, 1, \cdots, M)$ is found by minimizing the following function

$$\min_{\boldsymbol{g}_n^{(m)}} \overline{\Upsilon}_n^{(m)} = \overline{\Upsilon}_m(n) + \overline{\Upsilon}_{n-1}^{(m)} = \|\boldsymbol{\beta} \boldsymbol{e}_m(n)\|^2 + \sum_{i=1}^{n-1} \mu^{n-i} \overline{\Upsilon}_m(i) = \sum_{i=1}^{n} \|\mathbf{M}(i) \boldsymbol{e}_m(i)\|^2, \quad (4.60)$$

where $\mathbf{M}(i) = \mu^{\frac{n-i}{2}} \boldsymbol{\beta}^{\frac{1}{2}}$, and $\overline{\Upsilon}_m(i)$ is the weighted LS of $\Upsilon_m(i)$, e.g., $\overline{\Upsilon}_m(i) = \left\| \beta^{\frac{1}{2}} \mathbf{n}_m(i) \right\|^2 = \Upsilon_m(i)$ for $\boldsymbol{\beta} = \boldsymbol{I}_N$. Obviously, the general RLS-based techniques [106, 107] can be used to estimate and update the coefficients from snapshot to snapshot, but here we consider the case where the update takes place for each time interval, i.e., every time interval with multiple snapshots. Different from the idea suggested in [108], wherein the estimates of the coefficients are updated for block processing based on the QRD algorithm by implementing the so-called block Householder transformation, we, herein, propose a new block QRD RLS technique to update the estimates of the coefficients for each interval. First of all, let us consider the QRD for the known waveform data of the n^{th} time interval as

$$\boldsymbol{\beta}^{\frac{1}{2}} \boldsymbol{P}(n) = \boldsymbol{Q}(n) \begin{bmatrix} \overline{\boldsymbol{R}}(n) \\ \\ \mathbf{0}_{(N-K)\times K} \end{bmatrix}, n = 1, 2, \cdots \quad (4.61)$$

and the QRD for the known waveform data of the first $n-1$ time intervals as

$$\mathbf{M}_{n-1}\boldsymbol{P}_{n-1} = \boldsymbol{Q}_{n-1}\begin{bmatrix} \overline{\boldsymbol{R}}_{n-1} \\ \\ \mathbf{0}_{(nN-K-N)\times K} \end{bmatrix}, \tag{4.62}$$

where both $\overline{\mathbf{R}}(n)$ and $\overline{\boldsymbol{R}}_{n-1}$ are unique upper triangular matrices with positive diagonal elements and $\boldsymbol{Q}(n)$ and $\boldsymbol{Q}_{n-1}$ are unitary matrices. As such, we can get $\mathbf{M}_{n-1}\boldsymbol{x}_{n-1}^{(m)} = \boldsymbol{Q}_{n-1}\begin{bmatrix}\overline{\boldsymbol{p}}_{n-1}^T & \overline{\boldsymbol{\delta}}_{n-1}^{(m)T}\end{bmatrix}^T$ where $\overline{\mathbf{p}}_{n-1}$ is the Cholesky factor of $\mathbf{M}_{n-1}\boldsymbol{x}_{n-1}^{(m)}$ and $\overline{\boldsymbol{\delta}}_m(n-1)$ is the LS residual from which the estimated noise power $\hat{\sigma}_m^2(n-1)$ can easily be computed.

Next, we give two lemmas to derive the proposed block QRD RLS technique.

Lemma 1. *Let* $\boldsymbol{Z} \triangleq \begin{bmatrix}\boldsymbol{X}^H & \boldsymbol{Y}^H\end{bmatrix}^H$ *with both* $\boldsymbol{X} = \boldsymbol{Q}_X\begin{bmatrix}\boldsymbol{A}^T & \mathbf{0}_{K\times(m-K)}\end{bmatrix}^T$ *and* $\boldsymbol{Y} = \boldsymbol{Q}_Y\begin{bmatrix}\boldsymbol{B}^T & \mathbf{0}_{K\times(n-K)}\end{bmatrix}^T$ *being full column rank matrices, where* $\boldsymbol{Q}_X$ *and* $\boldsymbol{Q}_Y$ *are unitary matrices. Then computing the QRD of the* $(m+n)\times K$ *block data matrix* $\boldsymbol{Z} = \begin{bmatrix}\boldsymbol{X}^T & \boldsymbol{Y}^T\end{bmatrix}^T$ *is equivalent to implementing the QRD of the* $2K\times K$ *matrix* $\begin{bmatrix}\boldsymbol{A}^T & \boldsymbol{B}^T\end{bmatrix}^T$. *(See Appendix H for the proof)*

Lemma 2. *Let* $\boldsymbol{T} \triangleq \begin{bmatrix}\boldsymbol{A}^T & \boldsymbol{B}^T\end{bmatrix}^T$, *where* $\boldsymbol{A}$ *is a* $K\times K$ *nonsingular matrix and* $\boldsymbol{B}$ *is an arbitrary matrix of dimension* $m\times K$. *Then, there exists an* $(K+m)\times(K+m)$ *unitary matrix* $\boldsymbol{Q}\left(\boldsymbol{Q}^H\boldsymbol{Q} = \boldsymbol{I}_{(K+m)}\right)$ *such that* $\boldsymbol{Q}^H\boldsymbol{T} = \begin{bmatrix}(\boldsymbol{D}\boldsymbol{A})^T & \mathbf{0}_{K\times m}\end{bmatrix}^T$, *where* $\boldsymbol{D}$ *is the unique upper triangular matrix of the Cholesky decomposition of* $\boldsymbol{A}^{-H}\left(\boldsymbol{A}^H\boldsymbol{A}+\boldsymbol{B}^H\boldsymbol{B}\right)\mathbf{A}^{-1}$. *(See Appendix I for the proof).*

By using the Cholesky factors $\overline{\boldsymbol{R}}_{n-1}$ and $\overline{\boldsymbol{p}}_{n-1}$ obtained through the QRD of $\mathbf{M}_{n-1}\boldsymbol{P}_{n-1}$ and $\mathbf{M}_{n-1}\boldsymbol{x}_{n-1}^{(m)}$ along with the new incoming data $\{\boldsymbol{P}(n), \boldsymbol{x}_m(n)\}$ at time

interval n, we can obtain the current Cholesky factors $\overline{\boldsymbol{R}}_n$ and $\overline{\boldsymbol{p}}_n$, as stated in the following theorem.

Theorem 3. *There exists a unitary matrix $\boldsymbol{Q}_n$ such that*

$$\begin{bmatrix} \mu^{\frac{1}{2}}\overline{\boldsymbol{R}}_{n-1} & \mu^{\frac{1}{2}}\overline{\boldsymbol{p}}_{n-1} \\ \boldsymbol{\beta}\boldsymbol{P}(n) & \boldsymbol{\beta}\boldsymbol{x}_m(n) \end{bmatrix} = \boldsymbol{Q}_n \begin{bmatrix} \overline{\boldsymbol{R}}_n & \overline{\boldsymbol{p}}_n \\ \boldsymbol{0}_{N\times K} & \overline{\boldsymbol{\delta}}_m(n) \end{bmatrix} \tag{4.63}$$

where $\overline{\boldsymbol{R}}_n = \mu^{\frac{1}{2}}\boldsymbol{D}_1(n)\overline{\boldsymbol{R}}_{n-1}$, $\overline{\boldsymbol{p}}_n = \mu^{\frac{1}{2}}\boldsymbol{D}_1^{-H}(n)\left[\overline{\boldsymbol{p}}_{n-1} + \mu^4\overline{\boldsymbol{R}}_{n-1}^{-H}\boldsymbol{P}^H(n)\boldsymbol{\beta}\mathbf{x}_m(n)\right]$ *and* $\overline{\boldsymbol{\delta}}_m(n) = \boldsymbol{D}_2^{-H}(n)\boldsymbol{\beta}^{\frac{1}{2}}\left[\boldsymbol{x}_m(n) - \boldsymbol{P}(n)\overline{\boldsymbol{R}}_{n-1}^{-1}\overline{\boldsymbol{p}}_{n-1}\right]$. *Here, the matrices $\boldsymbol{D}_1(n)$ and $\boldsymbol{D}_2(n)$ are the unique upper triangular matrices of the Cholesky decomposition of $\overline{\boldsymbol{R}}_{n-1}^{-H}\left(\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1} + \mu^2\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)\overline{\boldsymbol{R}}_{n-1}^{-1}$ and $\left(\mu^4\overline{\boldsymbol{R}}(n)\left(\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1}\right)^{-1}\overline{\boldsymbol{R}}^H(n) + \boldsymbol{I}_N\right)$, respectively.*

Proof: By constructing a unitary matrix

$$\hat{\boldsymbol{Q}}(n) = \begin{bmatrix} \boldsymbol{I}_{K\times K} & \boldsymbol{0}_{K\times N} \\ \boldsymbol{0}_{N\times K} & \boldsymbol{Q}(n) \end{bmatrix} \tag{4.64}$$

and using (4.61), we obtain

$$\begin{bmatrix} \mu^{\frac{1}{2}}\overline{\boldsymbol{R}}_{n-1} \\ \boldsymbol{\beta}\boldsymbol{P}(n) \end{bmatrix} = \hat{\boldsymbol{Q}}(n)\begin{bmatrix} \mu^{\frac{1}{2}}\overline{\boldsymbol{R}}_{n-1}^H & \overline{\boldsymbol{R}}^H(n) & \boldsymbol{0}_{K\times(N-K)} \end{bmatrix}^H \tag{4.65}$$

According to *Lemma 1*, the QRD of the matrix $\begin{bmatrix} \mu^{\frac{1}{2}}\overline{\boldsymbol{R}}_{n-1}^T & \boldsymbol{\beta}\mathbf{P}^T(n) \end{bmatrix}^T$ is equivalent to the QRD of the matrix $\begin{bmatrix} \mu^{\frac{1}{2}}\overline{\boldsymbol{R}}_{n-1}^T & \overline{\boldsymbol{R}}^T(n) \end{bmatrix}^T$; therefore, using *Lemma 2*, one can

construct a $2K \times 2K$ unitary matrix $\hat{\boldsymbol{Q}}_n$ such that

$$\begin{bmatrix} \mu^{\frac{1}{2}}\overline{\boldsymbol{R}}_{n-1} \\ \overline{\mathbf{R}}(n) \end{bmatrix} = \hat{\boldsymbol{Q}}_n \begin{bmatrix} \overline{\boldsymbol{R}}_n \\ \mathbf{0}_{K\times K} \end{bmatrix} \tag{4.66}$$

where $\hat{\boldsymbol{Q}}_n = \begin{bmatrix} \boldsymbol{D}_1^{-1}(n) - & \mu^2\overline{\boldsymbol{R}}_{n-1}^{-H}\overline{\boldsymbol{R}}^H(n)\boldsymbol{D}_2^{-1}(n) \\ \mu^2\overline{\boldsymbol{R}}(n)\overline{\boldsymbol{R}}_{n-1}^{-1} & \boldsymbol{D}_1^{-1}(n) \quad \boldsymbol{D}_2^{-1}(n) \end{bmatrix}$ and $\overline{\boldsymbol{R}}_n = \mu^{\frac{1}{2}}\boldsymbol{D}_1(n)\overline{\boldsymbol{R}}_{n-1}$, and $\boldsymbol{D}_1(n)$ and $\boldsymbol{D}_2(n)$ are the unique upper triangular matrices of the Cholesky decomposition of $\overline{\boldsymbol{R}}_{n-1}^{-H}\left(\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1} + \mu^2\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)\overline{\boldsymbol{R}}_{n-1}^{-1}$ and $\mu^4\overline{\boldsymbol{R}}(n)\left(\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1}\right)^{-1}\overline{\boldsymbol{R}}^H(n) + \boldsymbol{I}_N$. Herein we have used the properties of the upper triangular matrix. Next, we construct a unitary $\boldsymbol{Q}_n$ to satisfy (4.63) by using (4.64) and (4.66), giving

$$\boldsymbol{Q}_n = \begin{bmatrix} \boldsymbol{I}_{K\times K} & \mathbf{0}_{K\times N} \\ \mathbf{0}_{N\times K} & \boldsymbol{Q}(n) \end{bmatrix} \begin{bmatrix} \boldsymbol{D}_1^{-1}(n) & -\mu^2\overline{\boldsymbol{R}}_{n-1}^{-H}\overline{\boldsymbol{R}}^H(n)\boldsymbol{D}_2^{-1}(n) & \mathbf{0}_{K\times(N-K)} \\ \mu^2\overline{\boldsymbol{R}}(n)\overline{\boldsymbol{R}}_{n-1}^{-1}\boldsymbol{D}_1^{-1}(n) & \boldsymbol{D}_2^{-1}(n) & \mathbf{0}_{K\times(N-K)} \\ \mathbf{0}_{(N-K)\times K} & \mathbf{0}_{(N-K)\times K} & \boldsymbol{I}_{(N-K)\times(N-K)} \end{bmatrix}. \tag{4.67}$$

Therefore,

$$\begin{aligned} \begin{bmatrix} \overline{\boldsymbol{p}}_n \\ \overline{\boldsymbol{\delta}}_m(n) \end{bmatrix} &= \boldsymbol{Q}_n^H \begin{bmatrix} \mu^{\frac{1}{2}}\overline{\boldsymbol{p}}_{n-1} \\ \boldsymbol{\beta}^{\frac{1}{2}}\boldsymbol{x}_m(n) \end{bmatrix} \\ &= \begin{bmatrix} \mu^{\frac{1}{2}}\boldsymbol{D}_1^{-H}(n)\overline{\boldsymbol{p}}_{n-1} + \mu^2\boldsymbol{D}_1^{-H}(n)\overline{\boldsymbol{R}}_{n-1}^{-H}\overline{\boldsymbol{R}}^H(n)\tilde{\boldsymbol{x}}_m^{(n)}(1:K) \\ -\boldsymbol{D}_2^{-H}(n)\overline{\boldsymbol{R}}(n)\overline{\boldsymbol{R}}_{n-1}^{-1}\overline{\boldsymbol{p}}_{n-1} + \boldsymbol{D}_2^{-H}(n)\tilde{\boldsymbol{x}}_m^{(n)}(1:K) \\ \tilde{\boldsymbol{x}}_m^{(n)}(K+1:N) \end{bmatrix}, \end{aligned} \tag{4.68}$$

where $\tilde{\boldsymbol{x}}_m^{(n)} \triangleq \boldsymbol{Q}^H(n)\boldsymbol{\beta}\boldsymbol{x}_m(n)$. Using (4.61), we can get the relationship $\overline{\boldsymbol{R}}^H(n)\tilde{\boldsymbol{x}}_m^{(n)}$

$(1:K) = \boldsymbol{P}^H(n)\boldsymbol{\beta}\mathbf{x}_m(n)$. With the results given by (4.66) and (4.68), the theorem is proven. □

Since our goal is to estimate the updating coefficients from the updated Cholesky factors $\overline{\boldsymbol{R}}_n$ and $\overline{\boldsymbol{p}}_n$, $\boldsymbol{g}_n^{(m)}$ can be estimated as per Theorem 3 as

$$\begin{aligned}
\hat{\boldsymbol{g}}_n^{(m)} &= \overline{\boldsymbol{R}}_n^{-1}\overline{\boldsymbol{p}}_n = \overline{\boldsymbol{R}}_{n-1}^{-1}\boldsymbol{D}_n^{(1)-1}(\boldsymbol{D}_n^{(1)-H}\overline{\boldsymbol{p}}_{n-1} + \mu^4\boldsymbol{D}_n^{(1)-H}\overline{\boldsymbol{R}}_{n-1}^{-H}\boldsymbol{P}^H(n)\boldsymbol{\beta}\boldsymbol{x}_m(n)) \\
&= \overline{\boldsymbol{R}}_{n-1}^{-1}\boldsymbol{D}_n^{(1)-1}\boldsymbol{D}_n^{(1)-H}\boldsymbol{D}_n^{(1)H}(\boldsymbol{D}_n^{(1)-H}\overline{\boldsymbol{p}}_{n-1} + \mu^4\boldsymbol{D}_n^{(1)-H}\overline{\boldsymbol{R}}_{n-1}^{-H}\boldsymbol{P}^H(n)\boldsymbol{\beta}\boldsymbol{x}_m(n)) \\
&= \overline{\boldsymbol{R}}_{n-1}^{-1}\left(\boldsymbol{D}_n^{(1)H}\boldsymbol{D}_n^{(1)}\right)^{-1}(\overline{\boldsymbol{p}}_{n-1} + \mu^4\overline{\boldsymbol{R}}_{n-1}^{-H}\boldsymbol{P}^H(n)\boldsymbol{\beta}\boldsymbol{x}_m(n)) \\
&= \left(\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1} + \mu^4\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1}(\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{p}}_{n-1} + \mu^4\boldsymbol{P}^H(n)\boldsymbol{\beta}\boldsymbol{x}_m(n)).
\end{aligned} \tag{4.69}$$

By paying special attention to the first time interval, i.e.,$n = 1, \overline{\boldsymbol{R}}_0 = \boldsymbol{I}_K, \overline{\boldsymbol{p}}_0 = \mathbf{0}_{K\times 1}$, then $\hat{\boldsymbol{g}}_1^{(m)} = \left(\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n) + \mu^{-4}\boldsymbol{I}_K\right)^{-1}\boldsymbol{P}^H(n)\boldsymbol{\beta}\boldsymbol{x}_m(n)$, where, compared with (4.14), we find that interestingly $\hat{\boldsymbol{g}}_1^{(m)}$ can be considered as a ridge regression estimate [109] with a fixed ridge parameter instead of the LS one. In addition, the mean and covariance matrix of the estimation $\boldsymbol{g}_n^{(m)}$ can be calculated below,

$$\begin{aligned}
E\left[\hat{\boldsymbol{g}}_n^{(m)}\right] &= \left(\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1} + \mu^4\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1} \\
&\qquad (\overline{\boldsymbol{R}}_{n-1}^H E\left[\overline{\boldsymbol{p}}_{n-1}\right] + \mu^4\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\boldsymbol{g}_n^{(m)}) \\
&= \boldsymbol{g}_n^{(m)} + (\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1} + \mu^4\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n))^{-1} \\
&\qquad (\overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1}E[\hat{\boldsymbol{g}}_{n-1}^{(m)}] - \overline{\boldsymbol{R}}_{n-1}^H\overline{\boldsymbol{R}}_{n-1}\boldsymbol{g}_n^{(m)}) \\
&= \boldsymbol{g}_n^{(m)} + \left(\boldsymbol{I}_K + \mu^4\overline{\boldsymbol{R}}_{n-1}^{-1}\overline{\boldsymbol{R}}_{n-1}^{-H}\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1}\left(E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right] - \boldsymbol{g}_n^{(m)}\right)
\end{aligned} \tag{4.70}$$

and

$$
\begin{aligned}
\boldsymbol{C}_{\hat{\boldsymbol{g}}_n^{(m)}} &= E\left[\left(\hat{\boldsymbol{g}}_n^{(m)} - E\left(\hat{\boldsymbol{g}}_n^{(m)}\right)\right)\left(\hat{\boldsymbol{g}}_n^{(m)} - E\left(\hat{\boldsymbol{g}}_n^{(m)}\right)\right)^H\right] \\
&= \left(\overline{\mathbf{R}}_{n-1}^H \overline{\boldsymbol{R}}_{n-1} + \mu^4 \overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1}\left[\mu^8 \boldsymbol{P}^H(n)\boldsymbol{\beta}^{\frac{1}{2}}\boldsymbol{C}_{\boldsymbol{e}_m}(n)\boldsymbol{\beta}^{\frac{1}{2}}\boldsymbol{P}(n)\right. \\
&\quad \left.+\overline{\boldsymbol{R}}_{n-1}^H \overline{\boldsymbol{R}}_{n-1}\boldsymbol{C}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}\overline{\boldsymbol{R}}_{n-1}^{-1}\overline{\boldsymbol{R}}_{n-1}^{-H}\right]\left(\overline{\boldsymbol{R}}_{n-1}^H \overline{\boldsymbol{R}}_{n-1} + \mu^4 \overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1} \qquad (4.71) \\
&= \left(\overline{\boldsymbol{R}}_{n-1}^H \overline{\boldsymbol{R}}_{n-1} + \mu^4 \overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1}\left[\mu^8 \sigma_m^2(n)\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right. \\
&\quad \left.+\overline{\boldsymbol{R}}_{n-1}^H \overline{\boldsymbol{R}}_{n-1}\boldsymbol{C}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}\overline{\boldsymbol{R}}_{n-1}^{-1}\overline{\boldsymbol{R}}_{n-1}^{-H}\right]\left(\overline{\boldsymbol{R}}_{n-1}^H \overline{\boldsymbol{R}}_{n-1} + \mu^4 \overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1}.
\end{aligned}
$$

If we set $\boldsymbol{g}_n^{(m)} \triangleq \hat{\boldsymbol{g}}_n^{(m)} - \Delta\boldsymbol{g}_n^{(m)}$ and use (4.70), the bias of $\Delta\boldsymbol{g}_n^{(m)}$ can be obtained as

$$
E\left[\Delta\boldsymbol{g}_n^{(m)}\right] = E\left[\hat{\boldsymbol{g}}_n^{(m)}\right] - \boldsymbol{g}_n^{(m)} = \left(\boldsymbol{I}_K + \mu^4 \overline{\boldsymbol{R}}_{n-1}^{-1}\overline{\boldsymbol{R}}_{n-1}^{-H}\overline{\boldsymbol{R}}^H(n)\overline{\boldsymbol{R}}(n)\right)^{-1}\left(E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right] - \boldsymbol{g}_n^{(m)}\right), \qquad (4.72)
$$

where we have used the relationship $\boldsymbol{g}_n^{(m)} = \boldsymbol{g}_m(n)$. Note that if the targets are steady over the time intervals, then (4.72) approaches zero, so the method yields an asymptotic unbiased estimator for the steady targets. Otherwise,(4.72) can be used to compensate the bias for moving targets.

4.3.3 Regularized Block Least Squares

In the previous subsection, we proposed a method to update the coefficients based on the block QRD-RLS technique, in which constant forgetting factors are employed. In this subsection we present another method to update the coefficients by using

regularized block least squares. Let us consider the following cost function

$$\begin{aligned}\min_{\boldsymbol{g}_n^{(m)}} \overline{\Upsilon}_n^{(m)} &= \left\|\boldsymbol{\beta}^{\frac{1}{2}}\boldsymbol{e}_m(n)\right\|^2+\mu(n)\left\|\boldsymbol{g}_n^{(m)}-\hat{\boldsymbol{g}}_{n-1}^{(m)}\right\|^2 \\ &= \left[\boldsymbol{x}_m(n)-\boldsymbol{P}(n)\boldsymbol{g}_n^{(m)}\right]^H\boldsymbol{\beta}\left[\boldsymbol{x}_m(n)-\boldsymbol{P}(n)\boldsymbol{g}_n^{(m)}\right] \\ &\quad +\mu(n)\left(\mathbf{g}_n^{(m)}-\hat{\boldsymbol{g}}_{n-1}^{(m)}\right)^H\left(\boldsymbol{g}_n^{(m)}-\hat{\boldsymbol{g}}_{n-1}^{(m)}\right).\end{aligned} \tag{4.73}$$

Compared with (4.60), we have replaced $\overline{\Upsilon}_{n-1}^{(m)}$ with the regularization term $\left\|\boldsymbol{g}_n^{(m)}-\hat{\boldsymbol{g}}_{n-1}^{(m)}\right\|^2$ $\mu(n)$, where $\mu(n)$ is a non-negative time varying factor to obtain the optimal estimate for $\boldsymbol{g}_n^{(m)}$, and $\hat{\boldsymbol{g}}_{n-1}^{(m)}$ is the previous optimal estimate for the coefficients. Now, from complex function theory, we can get the derivative of $\overline{\Upsilon}_n^{(m)}$ with respect with $\boldsymbol{g}_n^{(m)*}$ as

$$\frac{\partial\overline{\Upsilon}_n^{(m)}}{\partial\boldsymbol{g}_m^*(n)} = -\boldsymbol{P}^H(n)\boldsymbol{\beta}(\boldsymbol{x}_m(n)-\boldsymbol{P}(n)\boldsymbol{g}_n^{(m)})+\mu(n)\left(\boldsymbol{g}_n^{(m)}-\hat{\boldsymbol{g}}_{n-1}^{(m)}\right). \tag{4.74}$$

Setting the RHS of (4.74) to zero, the regularized LS estimate for $\boldsymbol{g}_n^{(m)}$ is obtained as

$$\hat{\boldsymbol{g}}_n^{(m)} = \left(\boldsymbol{P}^H(n)\boldsymbol{\beta}\boldsymbol{P}(n)+\mu(n)\boldsymbol{I}_K\right)^{-1}\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}\boldsymbol{x}_m(n)+\mu(n)\hat{\boldsymbol{g}}_{n-1}^{(m)}\right). \tag{4.75}$$

Interestingly, when $\mu(1)=\mu^{-4}$ and $\hat{\boldsymbol{g}}_0^{(m)}=\mathbf{0}$, (4.75) gives the same expression as (4.69). In this method, we would like to obtain the optimal estimate of $\boldsymbol{g}_n^{(m)}$ by choosing the regularization term $\mu(n)$. Herein, our measure of goodness is the mean square error (MSE), i.e., the trace of $\mathrm{MSE}_{\hat{\boldsymbol{g}}_n^{(m)}} \triangleq E\left[\left(\hat{\boldsymbol{g}}_n^{(m)}-\boldsymbol{g}_n^{(m)}\right)\left(\hat{\boldsymbol{g}}_n^{(m)}-\boldsymbol{g}_n^{(m)}\right)^H\right]$.

The mean and covariance matrix of the estimate $\hat{\boldsymbol{g}}_n^{(m)}$ can be found as

$$\begin{aligned} E\left[\hat{\boldsymbol{g}}_n^{(m)}\right] &= \left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n) + \mu(n)\boldsymbol{I}_K\right)^{-1}\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)\boldsymbol{g}_n^{(m)} + \mu(n)E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right]\right) \\ &= \boldsymbol{g}_n^{(m)} - \mu(n)\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n) + \mu(n)\boldsymbol{I}_K\right)^{-1}\left(E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right] - \boldsymbol{g}_n^{(m)}\right), \end{aligned} \tag{4.76}$$

and

$$\begin{aligned} \boldsymbol{C}_{\hat{\boldsymbol{g}}_n^{(m)}} &\triangleq E\left[\left(\hat{\boldsymbol{g}}_n^{(m)} - E\left[\hat{\boldsymbol{g}}_n^{(m)}\right]\right)\left(\hat{\boldsymbol{g}}_n^{(m)} - E\left[\hat{\boldsymbol{g}}_n^{(m)}\right]\right)^H\right] \\ &= \left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n) + \mu(n)\boldsymbol{I}_K\right)^{-1}\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)\boldsymbol{C}_{\hat{\boldsymbol{g}}_m}(n)\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)\right. \\ &\quad \left.+\mu^2(n)\boldsymbol{C}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}\right)\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n) + \mu(n)\boldsymbol{I}_K\right)^{-1} \\ &= \left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n) + \mu(n)\boldsymbol{I}_K\right)^{-1}\left(\sigma_m^2(n)\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)\right. \\ &\quad \left.+\mu^2(n)\boldsymbol{C}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}\right)\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n) + \mu(n)\boldsymbol{I}_K\right)^{-1}. \end{aligned} \tag{4.77}$$

The bias of the estimated $\hat{\boldsymbol{g}}_n^{(m)}$ is

$$\begin{aligned} E\left[\Delta\boldsymbol{g}_n^{(m)}\right] &= E\left[\hat{\boldsymbol{g}}_n^{(m)}\right] - \boldsymbol{g}_n^{(m)} \\ &= \mu(n)\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n) + \mu(n)\boldsymbol{I}_K\right)^{-1}\left(\boldsymbol{g}_n^{(m)} - E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right]\right). \end{aligned} \tag{4.78}$$

Note that when we employ (4.14) to estimate the coefficients for the first time interval as the initial vector and assume that the targets are stationary over the time intervals, i.e., $\boldsymbol{g}_n^{(m)} = E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right]$, the estimator is also an unbiased estimator.

Therefore by (4.76) and (4.77) we can get

$$\begin{aligned}
\mathrm{MSE}_{\hat{\boldsymbol{g}}_n^{(m)}} \triangleq & E\left[\left(\hat{\boldsymbol{g}}_n^{(m)}-\boldsymbol{g}_n^{(m)}\right)\left(\hat{\boldsymbol{g}}_n^{(m)}-\boldsymbol{g}_n^{(m)}\right)^H\right] \\
=&\boldsymbol{C}_{\hat{\boldsymbol{g}}_N^{(m)}}+\left(E\left[\hat{\boldsymbol{g}}_n^{(m)}\right]-\boldsymbol{g}_n^{(m)}\right)\left(E\left[\hat{\boldsymbol{g}}_n^{(m)}\right]-\boldsymbol{g}_n^{(m)}\right)^H \\
=&\left(\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)+\mu(n)\boldsymbol{I}_K\right)^{-1}\left(\sigma_m^2(n)\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)+\mu^2(n)\left(\left(E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right]-\boldsymbol{g}_n^{(m)}\right)\right.\right. \\
&\left.\left.\left(E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right]-\boldsymbol{g}_n^{(m)}\right)^H+\boldsymbol{C}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}\right)\right)\left(\boldsymbol{X}^H(N)\boldsymbol{\beta}^2\boldsymbol{X}(N)+\mu(N)\boldsymbol{I}_K\right)^{-1} \\
=&\left(\frac{\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)}{\mu(n)}+\boldsymbol{I}_K\right)^{-1}\left(\frac{\sigma_m^2(n)\mathbf{P}^H(n)\boldsymbol{\beta}^2\mathbf{P}(n)}{\mu^2(n)}+\overline{\mathrm{MSE}}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}\right) \\
&\left(\frac{\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\mathbf{P}(n)}{\mu(n)}+\boldsymbol{I}_K\right)^{-1},
\end{aligned} \tag{4.79}$$

where $\overline{\mathrm{MSE}}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}} \triangleq \left(E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right]-\boldsymbol{g}_n^{(m)}\right)\left(E\left[\hat{\boldsymbol{g}}_{n-1}^{(m)}\right]-\boldsymbol{g}_n^{(m)}\right)^H+\boldsymbol{C}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}$. Implementing eigendecomposition for the matrices $\overline{\mathrm{MSE}}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}$ and $\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)$ in (4.78) results in

$$\overline{\mathrm{MSE}}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}=\sum_{k=1}^{K}\lambda_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)}\boldsymbol{u}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)}\boldsymbol{u}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)H} \quad \text{and} \quad \boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)=\sum_{k=1}^{K}\lambda_{\boldsymbol{P}(n)}^{(k)}\boldsymbol{u}_{\boldsymbol{P}(n)}^{(k)}\boldsymbol{u}_{\boldsymbol{P}(n)}^{(k)H}, \tag{4.80}$$

where $\boldsymbol{u}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)}$ and $\boldsymbol{u}_{\boldsymbol{P}(n)}^{(k)}$ denote the k^{th} eigenvector of $\overline{\mathrm{MSE}}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}$ and $\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)$, respectively, and $\lambda_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)}$ and $\lambda_{\boldsymbol{P}(n)}^{(k)}$ are the corresponding real eigenvalues. Then, the

trace of $\mathrm{MSE}_{\hat{\boldsymbol{g}}_n^{(m)}}$ can be derived as

$$\begin{aligned}
\mathrm{TMSE}_{\hat{\boldsymbol{g}}_n^{(m)}} &\triangleq \mathrm{Tr}\left\{\mathrm{MSE}_{\hat{\boldsymbol{g}}_n^{(m)}}\right\} \\
&= \mathrm{Tr}\left\{\left(\frac{\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)}{\mu(n)} + \boldsymbol{I}_K\right)^{-2}\left(\frac{\sigma_m^2(n)\boldsymbol{P}^H(n)\boldsymbol{\beta}^2\boldsymbol{P}(n)}{\mu^2(n)} + \overline{\mathrm{MSE}}_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}\right)\right\} \\
&= \sigma_m^2(n)\sum_{k=1}^{K}\frac{\lambda_{\boldsymbol{P}(n)}^{(k)}}{\left(\mu(n)+\lambda_{\boldsymbol{P}(n)}^{(k)}\right)^2} + \mu^2(n)\sum_{k=1}^{K}\frac{\lambda_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)}}{\left(\mu(n)+\lambda_{\boldsymbol{P}(n)}^{(k)}\right)^2} \\
&= \sum_{k=1}^{K}\frac{\mu^2(n)\lambda_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)} + \sigma_m^2(n)\lambda_{\boldsymbol{P}(n)}^{(k)}}{\left(\mu(n)+\lambda_{\boldsymbol{P}(n)}^{(k)}\right)^2}.
\end{aligned} \tag{4.81}$$

Note that $\mathrm{TMSE}_{\hat{\boldsymbol{g}}_n^{(m)}} = \sum_{k=1}^{K}\frac{\sigma_m^2(n)}{\lambda_{\boldsymbol{P}(n)}^{(k)}}$ when $\mu(n)=0$, which means $\hat{\boldsymbol{g}}_n^{(m)} = \hat{\boldsymbol{g}}_m(n)$. Meanwhile $\mathrm{TMSE}_{\hat{\boldsymbol{g}}_n^{(m)}} = \sum_{k=1}^{K}\lambda_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)}$ when $\mu(n)=+\infty$, implying that $\hat{\boldsymbol{g}}_n^{(m)} = \hat{\boldsymbol{g}}_{n-1}^{(m)}$. Other values of $\mu(n)\in(0,+\infty)$ may give an effective compromise between $\hat{\boldsymbol{g}}_m(n)$, an unstable BLUE based on the snapshots of the n^{th} time interval only, and $\hat{\boldsymbol{g}}_{n-1}^{(m)}$ as a stable estimate obtained by using the previous $n-1$ time intervals. Therefore, we now use the first derivative with respect to $\mu(n)$ of (4.81) to obtain the optimal estimator for $\boldsymbol{g}_n^{(m)}$. The derivative is

$$\frac{d\mathrm{TMSE}_{\hat{\boldsymbol{g}}_n^{(m)}}}{d\mu(n)} = \frac{2\lambda_{\boldsymbol{P}(n)}^{(k)}}{\left(\mu(n)+\lambda_{\boldsymbol{P}(n)}^{(k)}\right)^3}\sum_{k=1}^{K}\left(\mu(n)\lambda_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)} - \sigma_m^2(n)\right). \tag{4.82}$$

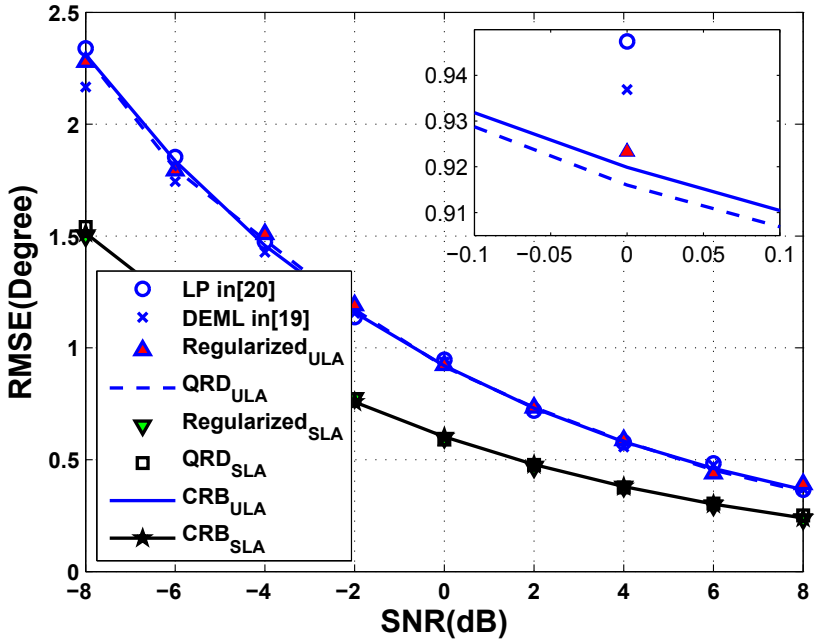
For any $\mu(n)>0$ we have $(\mu(n)+\lambda_{\boldsymbol{P}(n)}^{(k)})^3 > 0$. By setting the RHS of (46) to zero, we obtain the best regularized value, $\mu(n) = K\sigma_m^2(n)\Big/\sum_{k=1}^{K}\lambda_{\hat{\boldsymbol{g}}_{n-1}^{(m)}}^{(k)}$ which gives the optimal estimate $\hat{\boldsymbol{g}}_n^{(m)}$.

4.3.4 Simulation Results for DOA Tracking

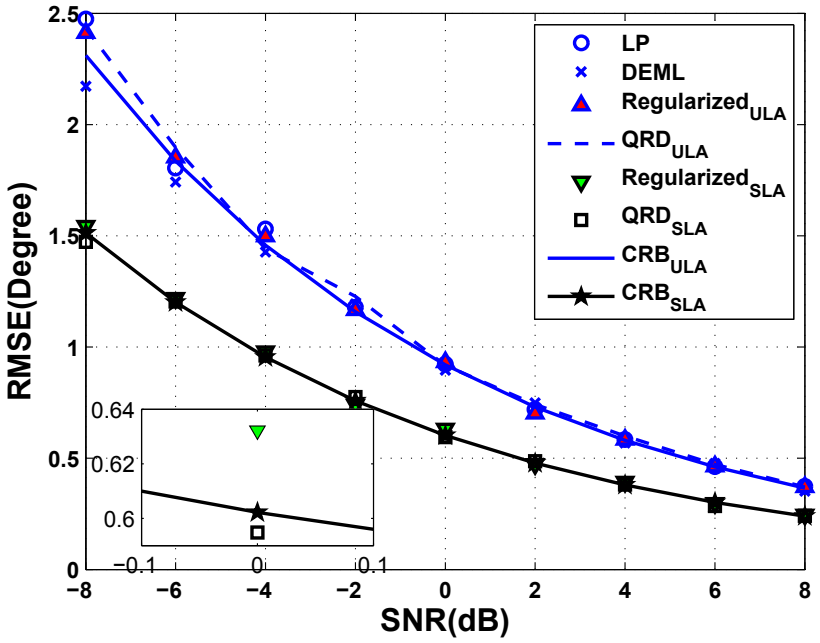
In this subsection, simulation results are presented to validate the proposed DOA estimation and tracking techniques. In the first example, we consider two steady signals from angles $\theta_1 = -5^\circ$ and $\theta_2 = 5^\circ$ with equal power impinging onto an array with 3 sensors. The parameters for the array system are the same in the first example of Section 4.2. Each experiment contains 1000 independent trials to compute the root mean square error (RMSE). The proposed block QRD-RLS and the regularized LS methods are compared with the LP method [59], the DEML method [54] as well as the CRB [53, 54]. The second example shows the tracking performance of the QRD-RLS method for different numbers of snapshots in each time interval, where the coefficients are constant, along with that of the DEML method for one steady target and one moving target using ULA with four sensors. In the third one, we will show the tracking performance of our QRD-RLS method using SLA with different forgetting factors, as well as their bias compensation results. The tracking performance for two crossing moving targets is shown in the fourth example, where both the proposed QRD-RLS and regularized LS methods are evaluated using the same SLA as in the third example. Finally, the tracking performance for multiple moving sources with the regularized LS is presented in the fifth example.

Example 1: *Estimation performance of steady targets with respect to SNR*

We evaluate the DOA estimation performance of the proposed block QRD-RLS and regularized LS methods against the SNR. The SNR of the two uncorrelated signals is varied from -4 to 6 dB. The number of snapshots is P=100. The simulated RMSE



(a)



(b)

Figure 4.7: RMSE of DOA estimation for two uncorrelated sources versus SNR from -4 to 6dB at (a)$\theta_1 = -5°$ and (b)$\theta_2 = 5°$ with 100 snapshots.

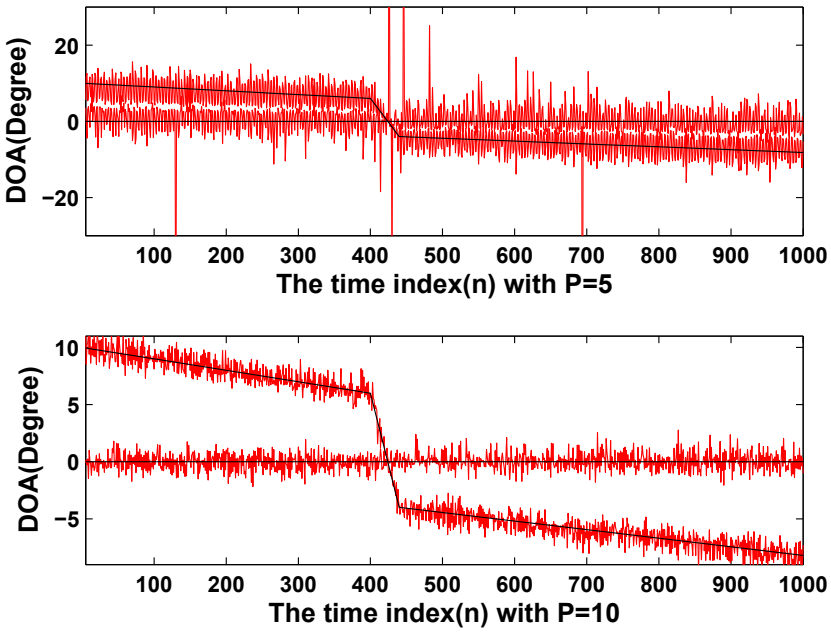
of the DOA estimate is shown in Fig.4.7, together with the CRBs for ULA and SLA. It is seen that the simulated RMSEs of the proposed method for ULA and SLA are almost identical to their respective CRBs. It is also seen that the simulated RMSEs of the proposed method decreases monotonically with increasing SNR. In addition, although the LP and DEML methods exhibit the same performance as our proposed method in the case of using ULA, yet our methods can make use of SLA to improve the estimation performance significantly.

Example 2: *Tracking performance for one steady and one moving targets*

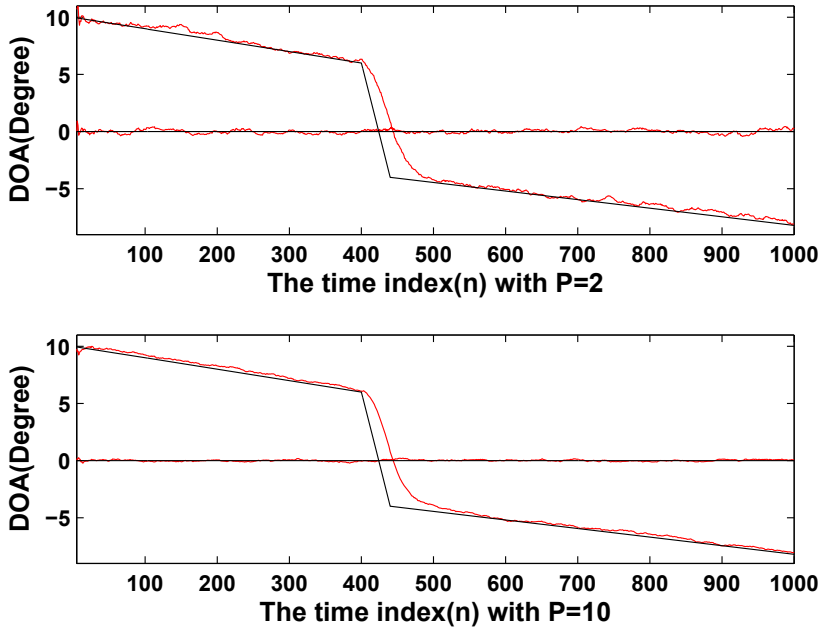
This example studies the tracking performances of the proposed block QRD-RLS method with $\mu = 0.98$ and the DEML method using a four-sensor ULA. Two equal power uncorrelated signals with SNR=10dB impinge on ULA, where the first angle is 0° and the other one is governed by

$$\text{DOA} = \begin{cases} 10 - n/100 & n \in [1 \quad 400] \\ 6 - 2.5(n-400)/10 & n \in [401 \quad 440] \\ -4 - 7.5(n-440)/1000 & n \in [441 \quad 1000] \end{cases} \tag{4.83}$$

Here, we assume that each time interval has N snapshots, and we employ $N = 5$ and $N = 10$ snapshots for DEML and $N = 2$ and $N = 10$ snapshots for our QRD-RLS method, respectively, to update and track the DOAs. The results are shown in Fig.4.8(a) and 4.8(b). Note that the DEML method is unable to track angles when $N = 5$, while the proposed method yields satisfactory tracking results even when $N = 2$. The possible explanation is that the recursion-based technique is employed in our method, while DEML is a batch based technique. From Fig.4.8(b), we also



(a)



(b)

Figure 4.8: Tracking of two sources with a four-sensor ULA: one DOA fixed at 0° and one moving from 10° to −8° using(a) DEML method, and (b) proposed QRD-RLS method.

observe that the more the snapshots used in the same time interval, the better the performance in terms of the variance.

Example 3: *Tracking performance for different forgetting factors with the bias compensation*

The tracking performance of the proposed block QRD-RLS method with $\mu = 0.98$ and $\mu = 0.90$ as well as their bias compensation results is presented in Fig.4.9. The simulation conditions are as follows: $N = 10$, SNR=10dB, $d_1 = 1, d_2 = 5$, one steady target fixed at 0° and the other moving target generated according to

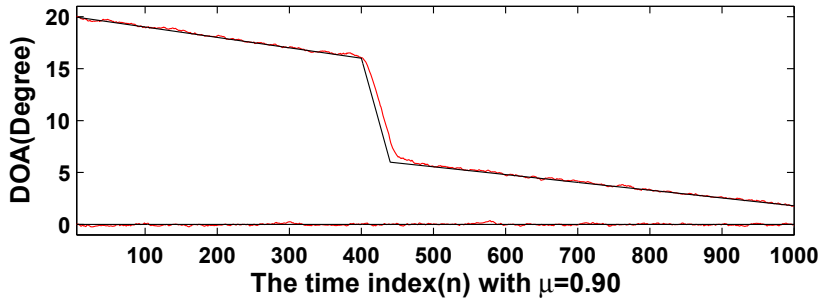
$$\text{DOA} = \begin{cases} 20 - n/100 & n \in [1 \quad 400] \\ 16 - 2.5(n - 400)/10 & n \in [401 \quad 440] \\ 6 - 7.5(n - 440)/1000 & n \in [441 \quad 1000] \end{cases} \tag{4.84}$$

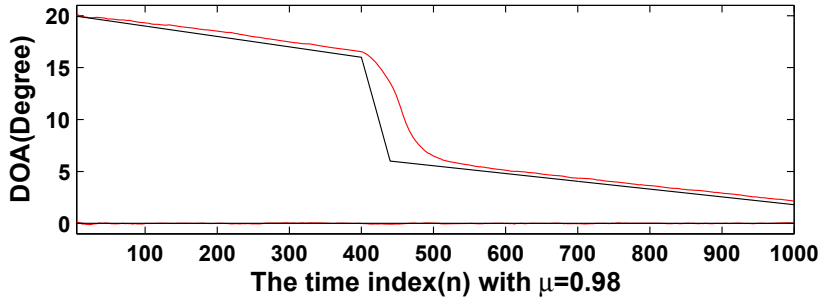
It can be seen from Fig.4.9(a) that the larger the forgetting factor, the better the tracking performance for the fast moving target at the cost of a larger variance. Fig.4.9(a) indicates that we can improve the tracking performance for the fast moving target by exploiting the bias compensation term in (4.78).

Example 4: *Tracking errors for two crossing moving targets*

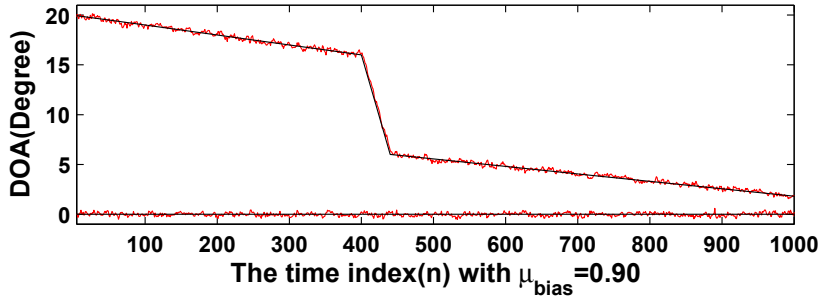
Here, we present the tracking errors between the estimated DOA and the true DOA of two crossing moving targets according to $10 \times \cos(\pi \times n/2000)$ and $10 \times \sin(\pi \times n/2000)$, where the array system model is the same as in the third example. From Fig.4.10, we see that the regularized LS has the best tracking performance for the fast moving targets in terms of the bias; however, it has the largest variance.

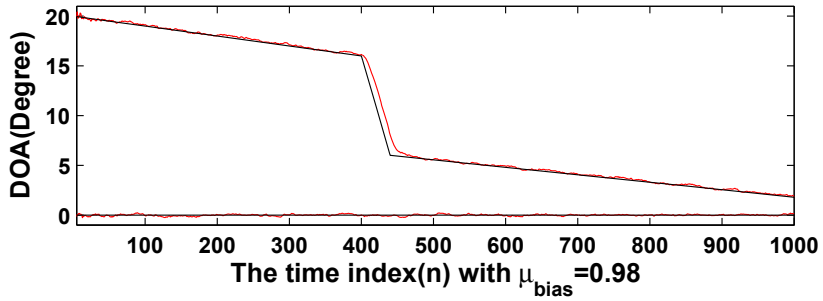
Example 5: *Performance of DOA tracking for multiple moving targets*





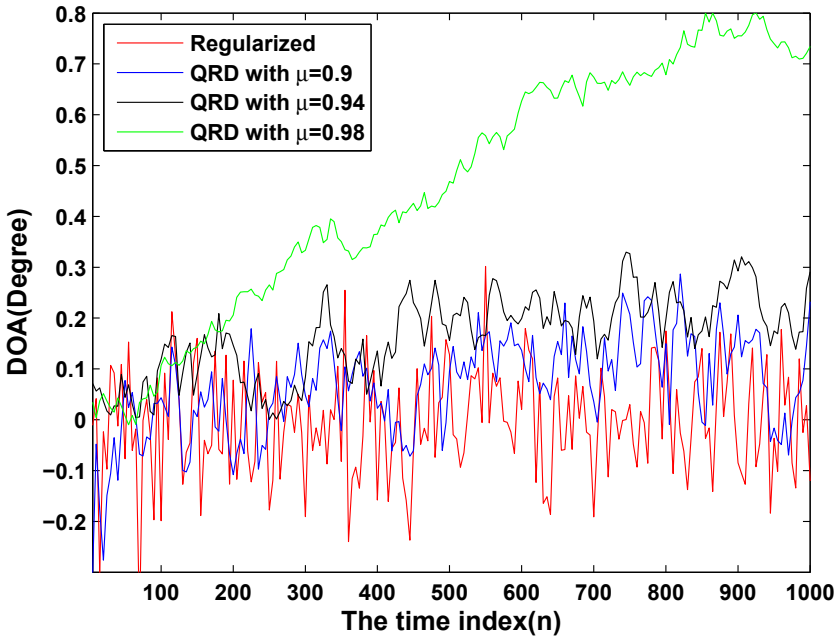
(a)



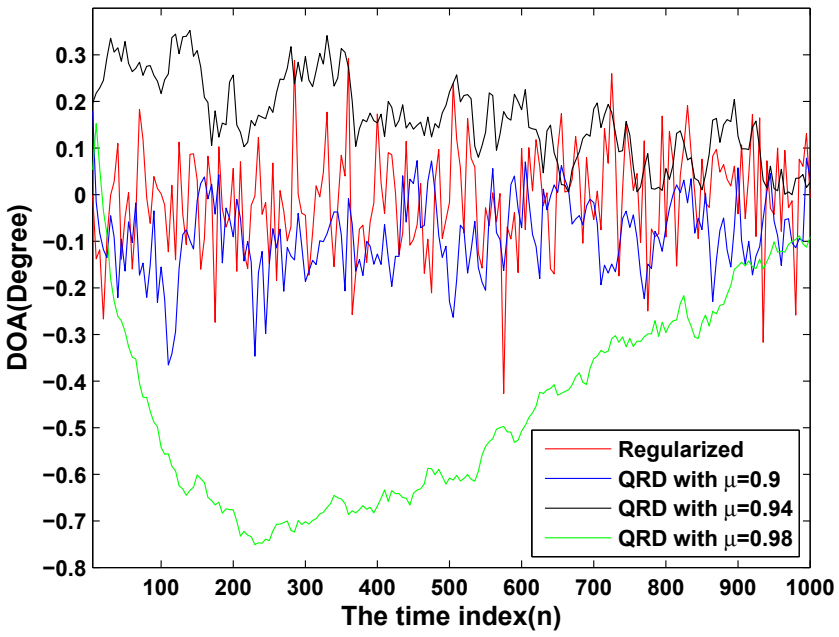


(b)

Figure 4.9: Tracking performance for (a)different forgetting factors and (b) their corrsponding bias compensation results with SNR=10dB and N=10

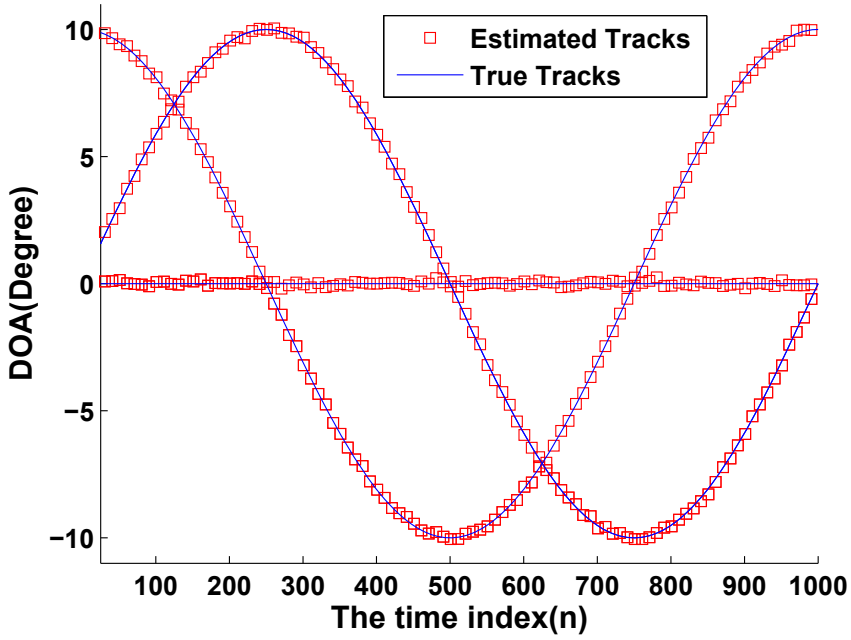


(a)

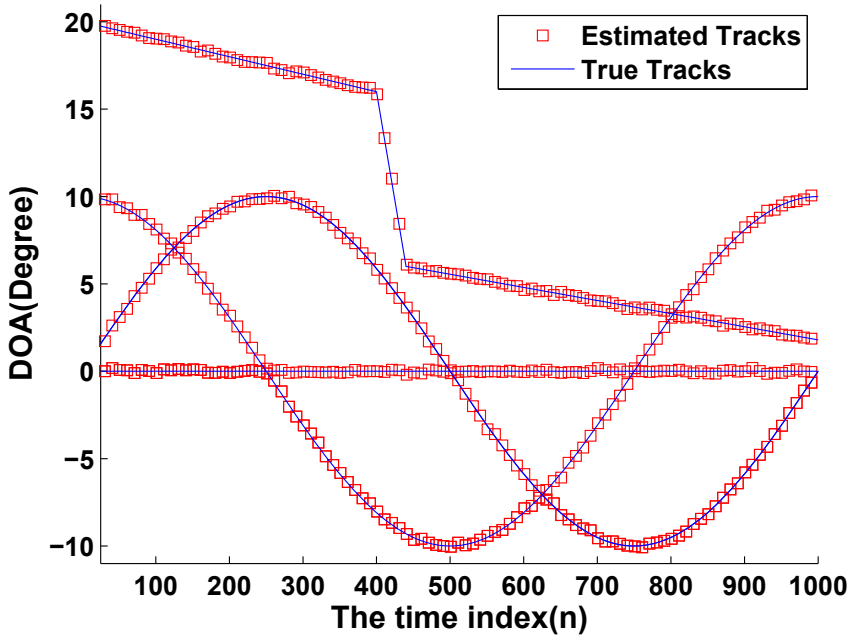


(b)

Figure 4.10: Tracking errors between the estimated DOA and the true DOA of the crossing moving targets changing at (a)$10 \times \cos(\pi \times n/2000)$ and (b) $10 \times \sin(\pi \times n/2000)$ via SLA with $d_1 = 1 \quad$ and $\quad d_2 = 5$, where SNR=10dB and N=10.



(a)



(b)

Figure 4.11: Tracking multiple sources with one steady target plus (a) two moving targets via ULA, or (b) three moving targets via SLA with $d_1 = 1$ and $d_2 = 5$, where SNR=10dB and N=10.

The tracking performance of the proposed regularized LS method for multiple targets is studied. Here we consider both the 3-source and 4-source scenarios. In the case of 2 moving targets with one steady target fixed in 0°, the two moving tracks of the DOA are described by $10 \times \cos(\pi \times n/2000)$ and $10 \times \sin(\pi \times n/2000)$, respectively, with the complex gains given by $\mathbf{\Xi}(n) = diag(e^{j\pi/4}, e^{-j\pi/4}, 1)$. In the case of 4 targets, three targets have the same tracks as in the 3-source case and the other one is expressed by (4.84), and the gains are $\mathbf{\Xi}(n) = diag(e^{j\pi/4}, e^{-j\pi/4}, 1, -1)$. Other simulation conditions are the same as in Example 3. It can be seen from Fig.4.11 that the proposed method provides a good tracking performance for multiple moving targets and the estimated tracks are almost identical to the true values even when the number of targets is larger than or equal to the number of sensors, which is unobtainable from the traditional methods without using the temporal information.

4.4 Conclusion

In this chapter, we have proposed a very efficient DOA estimation method for signals with known waveforms based on sparse linear array. Different from the traditional methods used to estimate DOA with known waveforms, our method splits the LS problem into several linear regression expressions, wherein each coefficient of the linear regression model includes a pair of angle and gain. The optimal estimator for the DOA and gain has been derived along with its statistical performance analysis. When DOAs of the targets are varying with the time interval, we have proposed two efficient DOA estimation and tracking methods based on a linear regression model, termed

the block QRD-RLS technique and the block regularized LS. Thanks to the temporal information of the incident signals, we have been able to employ linear regression analysis to exploit the spatial information of the signals impinging on each sensor, and develop efficient QRD-RLS and regularized LS methods to update the changing regression coefficents. Simulation results have shown that our proposed method can provide a better DOA estimation and tracking performance than the previous methods, such as the LP and the DEML do. In addition, it is confirmed that the larger the array aperture, the better the estimation and tracking performance, making it possible to achieve the desired performance via a sparse array without having to handle the ambiguity. It is also worth noting that the proposed DOA estimation/tracking techniques are based on the assumption that each sensor can be considered as a linear regression model, therefore, these techniques can be used for arbitrary arrays including the special arrays proposed in Chapters 2 and 3 along with some phase unwrapped techniques.

Chapter 5

Summary and Further Research Directions

5.1 Concluding Remarks

In this thesis, the problem of DOA estimation and tracking using sparse array systems has been thoroughly studied. By meticulously designing sparse array geometries, several approaches have been developed to estimate and track the DOAs of different types of source signals including those with known or unknown waveforms or modelled by autoregressive (AR) signals.

First, two nonuniform sparse linear arrays (SLAs) have been designed to improve the accuracy of conventional DOA estimation methods. One is based on the principle of the minimum redundancy linear array (MRLA) and the interferometer sensing for estimating 1-D DOA of uncorrelated sources, and constructs an extended correlation matrix by using the Kronecker steering vectors (KSVs) each containing a pair of

ambiguous and unambiguous angles. The other is constructed by two sparse uniform linear arrays (ULAs) for estimating 1-D DOA of correlated sources, where the inter-element spacing of each sparse ULA is much large than half wavelength, while the minimal inter-element spacing of the whole sparse linear array is less than or equal to half wavelength. By this scheme, a new array design strategy to improve the DOA estimate accuracy significantly is proposed with the addition of the appropriate DOA estimation algorithms. Inspired by the idea of the generalized ESPRIT, we have obtained the rough DOA without ambiguity, and then designed the alternating null-steering technique to select the true fine DOA in the interval around the rough DOA. We have also proposed a computationally efficient 2-D DOA estimation method based on sparse L-shaped array for estimating 2-D DOA of correlated sources, where one ULA is arranged along the z-axis and one SLA on the x-axis. In this method, the elevation angle is estimated by the cross-correlation matrix that is free of the effect of additional noise, and then the source waveform is estimated using the estimated elevation angle. By taking advantage of the estimated source waveform and elevation angle, a total least squares-like technique is exploited to estimate the spatial signatures in terms of the azimuth angle of the incident signals. This method has avoided the pair matching problem due to the one-to-one relationship between the source waveform and the corresponding incident angle.

Second, the DOA estimation of AR-modeled source signals has been investigated. Using the properties of these signals and the symmetrical sparse array structures we have proposed two Kalman fiter-based DOA estimation methods. The first method via symmetrical sparse array consists of two main steps: (i) to obtain the source waveform

of the AR modeled source signal by the celebrated Kalman filter; and (ii) to estimate the DOAs and AR coefficients by exploiting the efficient QR-decomposition-based recursive least square (QRD-RLS) technique. There are two advantages with this method over the traditional methods. One is that the AR-modeled sources can provide useful temporal information to handle cases such as the number of sources being larger than the number of antennas. The other is that the QRD-RLS is exploited to estimate the slowly changing targets during the same time interval. In addition, the symmetric array enables one to transfer a complex-valued nonlinear problem to a real-valued linear one, thus reducing the computational complexity. Although this technique can provide a good DOA estimation and tracking performance with small computational burden, there are still some restrictions such as symmetric sparse array and real-valued AR coefficients. Therefore, to overcome these weaknesses, we have proposed another DOA estimation technique for AR-modeled sources based on SLA. Each sensor of the array is considered as a subsystem to obtain the angle information for the DOA estimation, and then an asymptotic optimal unbiased estimator is developed to obtain the final DOAs based on total least squares-like technique. In addition, our proposed estimator can be used in the array configurations proposed in Chapter 2.

Third, we have considered the DOA estimation of signals with known waveform. In this regard, we have introduced a very efficient DOA estimation method based on sparse linear array. Different from the traditional methods used to estimate DOA with known waveforms, our method splits the LS problem into several linear regression expressions, wherein each coefficient of the linear regression model includes a pair of angle and gain. The optimal estimator for the DOA and gain has been derived along

with a study of its statistical performance. Furthermore, for the moving targets, we have proposed two efficient DOA estimation and tracking methods based on linear regression models, one is called the block QRD-RLS technique and the other is the block regularized LS. Thanks to the temporal information in the incident signals, we can employ the linear regression models to deal with the spatial information of signals impinging on each sensor, and then employ the efficient QRD-RLS and regularized least squares methods to handle the linear regression models. Therefore, our proposed techniques can be utilized in general sparse arrays including those proposed in Chapters 2 and 3 with the help of some disambiguity methods.

5.2 Suggestions for Further Research

During my four years of study for DOA estimation and tracking based on sparse array systems, some original ideas have been proposed on designing sparse array systems and developing efficient algorithms for estimating and tracking DOAs of incident targets. There are still some issues that require further investigations.

- The proposed 2-D DOA estimation method using sparse L-shaped array and total least squares techniques can be generalized to the arbitrary planar or cubic sparse array to implement the accurate 2-D DOA estimation.

- Although the new total least squares-based 2-D DOA estimation proposed in Chapter 2 is computationally efficient with good performance, it is still a batch processing algorithm. In many applications, such as target tracking, the covariance matrix changes on a snapshot-by-snapshot or time interval-by-time interval

basis. Batch processing definitely becomes worse in such a scenario. Efficient and effective ways of updating the DOA will be more attractive for real-time applications. To the best of our knowledge, there is no such tracking technique that exploits the special array structure along with signal spatial signatures to track the signal's DOA. Also, its sensitivity analysis is worth-studying.

- The proposed total least squares-based 2-D DOA estimation algorithm is based on the assumption that the elevation angle is perfectly estimated. In practice, the measurements are noisy and the estimation of the elevation angle is not perfect. The use of (2.42) to estimate the source waveform may suffer from the bias due to the estimation error of the elevation angle. Therefore, it is necessary derive a more effective and robust total least squares-based method to minimize the error influence from the estimated elevation angle.

- The proposed DOA estimation method using Kalman filter and total least squares techniques can be extended to the situation where the noise is characterized by heavier tails and generating high-intensity distribution, named outliers. Robust statistical procedures would be desired to cope with the outlying data points and reduce the influence of the outliers [110–113].

- The least squares-based method for estimating and tracking the DOA of signals with known waveform could be extended to the case where the source waveform of interest is affected by random errors which may cause uncertainties about the source waveforms. As shown in Chapter 4, the least squares-based method can be implemented very well to estimate DOA due to the fact that the source

waveform of interest is completely known or perfectly estimated. Unfortunately, the source waveform of interest in many applications are contaminated by noise and thus must be estimated from noisy measurements. In such a case, the LS method suffers from bias and increased covariance due to the accumulation of noise errors in $\boldsymbol{R}_{pp}$ of (4.15). Therefore, a total least square-like method should be developed to alleviate the effect of the noise or the estimate error from the estimated waveforms.

Appendix A

Derivation of the Mean and Covariance of $b_{m_{\mathrm{MTLS}}}$ in Eq.(2.45)

Using the results in (2.43), we can write the matrix $\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}}$ as

$$\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}} = [\boldsymbol{S}+\boldsymbol{N}_s]^H[\boldsymbol{S}+\boldsymbol{N}_s] = \boldsymbol{S}^H\boldsymbol{S} + \boldsymbol{N}_s^H\boldsymbol{S} + \boldsymbol{S}^H\boldsymbol{N}_s + \boldsymbol{N}_s^H\boldsymbol{N}_s \tag{A.1}$$

where $\boldsymbol{N}_s = [\boldsymbol{n}_s(1), \boldsymbol{n}_s(2), \cdots, \boldsymbol{n}_s(N)]^T$. Following the same idea as in [82], we can express $\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}} - N\boldsymbol{C}_{n_s}$ as

$$\begin{aligned}\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}} - N\boldsymbol{C}_{n_s} &= \boldsymbol{S}^H\boldsymbol{S} + (\boldsymbol{N}_s^H\boldsymbol{S} + \boldsymbol{S}^H\boldsymbol{N}_s + \boldsymbol{N}_s^H\boldsymbol{N}_s - N\boldsymbol{C}_{n_s}) \\ &\triangleq \boldsymbol{R}_s + \triangle\boldsymbol{R}_s\end{aligned}. \tag{A.2}$$

It is easy to verify that

$$E[\triangle\boldsymbol{R}_s] = E[\boldsymbol{N}_s^H\boldsymbol{S} + \boldsymbol{S}^H\boldsymbol{N}_s + \boldsymbol{N}_s^H\boldsymbol{N}_s - N\boldsymbol{C}_{n_s}] = \boldsymbol{0}. \tag{A.3}$$

Assuming that the eigenvalues of $\boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s$ are less than unity [82], the first-order expansion of $(\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}} - N\boldsymbol{C}_{n_s})^{-1}$ in terms of $\triangle\boldsymbol{R}_s$ can be expressed as

$$(\hat{\boldsymbol{S}}^H\hat{\boldsymbol{S}} - N\boldsymbol{C}_{n_s})^{-1} = (\boldsymbol{R}_s + \triangle\boldsymbol{R}_s)^{-1} \simeq \boldsymbol{R}_s^{-1} - \boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s\boldsymbol{R}_s^{-1}. \tag{A.4}$$

Therefore, $\boldsymbol{b}_{m_{\mathrm{MTLS}}}$ can be rewritten as

$$\begin{aligned}
\boldsymbol{b}_{m_{\mathrm{MTLS}}} &\simeq (\boldsymbol{R}_s^{-1} - \boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s\boldsymbol{R}_s^{-1})(\boldsymbol{S} + \boldsymbol{N}_s)^H\boldsymbol{x}_m \\
&= (\boldsymbol{R}_s^{-1} - \boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s\boldsymbol{R}_s^{-1})(\boldsymbol{S} + \boldsymbol{N}_s)^H(\boldsymbol{S}\boldsymbol{b}_m + \boldsymbol{n}_{x,m}) \\
&= (\boldsymbol{R}_s^{-1} - \boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s\boldsymbol{R}_s^{-1})(\boldsymbol{R}_s\boldsymbol{b}_m + \boldsymbol{N}_s^H\boldsymbol{n}_{x,m} + \boldsymbol{N}_s^H\boldsymbol{S}\boldsymbol{b}_m + \boldsymbol{S}^H\boldsymbol{n}_{x,m}). \\
&= \boldsymbol{b}_m + (\boldsymbol{R}_s^{-1}\boldsymbol{N}_s^H\boldsymbol{S} - \boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s - \boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s\boldsymbol{R}_s^{-1}\boldsymbol{N}_s^H\boldsymbol{S})\boldsymbol{b}_m \\
&\quad + (\boldsymbol{R}_s^{-1} - \boldsymbol{R}_s^{-1}\triangle\boldsymbol{R}_s\boldsymbol{R}_s^{-1})(\boldsymbol{N}_s^H\boldsymbol{n}_{x,m} + \boldsymbol{S}^H\boldsymbol{n}_{x,m})
\end{aligned} \tag{A.5}$$

Obviously, if the number of snapshots is large enough, we can obtain $\lim_{N\to\infty}\frac{\triangle\boldsymbol{R}_s}{N} \backsimeq \boldsymbol{0}$, and then $\boldsymbol{b}_{m_{\mathrm{MTLS}}} \simeq \boldsymbol{b}_m + \boldsymbol{R}_s^{-1}\hat{\boldsymbol{S}}^H\boldsymbol{n}_{x,m}$. As such, it is easy to show that the mean of $\boldsymbol{b}_{m_{\mathrm{MTLS}}}$ equals $\boldsymbol{b}_m$ and its covariance can be obtained as

$$\begin{aligned}
\boldsymbol{C}_{b_{MTLS}} &= E[(\boldsymbol{b}_{m_{\mathrm{MTLS}}} - \boldsymbol{b}_m)(\boldsymbol{b}_{m_{\mathrm{MTLS}}} - \boldsymbol{b}_m)^H] \\
&= \boldsymbol{R}_s^{-1}E[\hat{\boldsymbol{S}}^H\boldsymbol{n}_{x,m}\boldsymbol{n}_{x,m}^H\hat{\boldsymbol{S}}]\boldsymbol{R}_s^{-H} \qquad . \\
&= \sigma^2\boldsymbol{R}_s^{-1}(\boldsymbol{R}_s + N\boldsymbol{C}_{n_s})\boldsymbol{R}_s^{-H}
\end{aligned} \tag{A.6}$$

Appendix B

Derivation of the Mean and Covariance of $\boldsymbol{A}_{TLS}^{m}$ in Eq.(3.41)

Let us define $\hat{\boldsymbol{s}}(i) \triangleq \boldsymbol{s}(i) + \tilde{\boldsymbol{s}}(i) = [\hat{p}_1(i|i), \hat{p}_2(i|i), \cdots, \hat{p}_K(i|i)]^T$, where $\tilde{\boldsymbol{s}}(i)$ is a K dimensional random Gaussian vector with zero mean and the covariance $\boldsymbol{C}_{\tilde{\boldsymbol{s}}(i)} = \boldsymbol{C}_{\hat{\boldsymbol{s}}(i)} = \boldsymbol{\Gamma} P(i|i) \boldsymbol{\Gamma}^T$. Thus, we can obtain

$$\hat{\boldsymbol{S}}^* \hat{\boldsymbol{S}}^T = [\boldsymbol{S} + \tilde{\boldsymbol{S}}]^* [\boldsymbol{S} + \tilde{\boldsymbol{S}}]^T = \boldsymbol{S}^* \boldsymbol{S}^T + \tilde{\boldsymbol{S}}^* \boldsymbol{S}^T + \boldsymbol{S}^* \tilde{\boldsymbol{S}}^T + \tilde{\boldsymbol{S}}^* \tilde{\boldsymbol{S}}^T \tag{B.1}$$

It can easily be verified that the covariance

$$E[\hat{\boldsymbol{S}}^* \hat{\boldsymbol{S}}^T] = E[\boldsymbol{S}^* \boldsymbol{S}^T] + \sum_{i=1}^{N} \boldsymbol{C}_{\hat{\boldsymbol{s}}(i)}. \tag{B.2}$$

In a manner similar to the derivation in Appendix A, we denote

$$(\hat{\boldsymbol{S}}^* \hat{\boldsymbol{S}}^T) - \sum_{i=1}^{N} \boldsymbol{C}_{\hat{\boldsymbol{s}}(i)} \triangleq \boldsymbol{R}_S(N) + \triangle \boldsymbol{R}_S(N), \tag{B.3}$$

where $\boldsymbol{R}_S(N) \triangleq \boldsymbol{S}^* \boldsymbol{S}^T$ and $\triangle \boldsymbol{R}_S(N) \triangleq \tilde{\boldsymbol{S}}^* \boldsymbol{S}^T + \boldsymbol{S}^* \tilde{\boldsymbol{S}}^T + \tilde{\boldsymbol{S}}^* \tilde{\boldsymbol{S}}^T - \sum_{i=1}^{N} \boldsymbol{C}_{\hat{\boldsymbol{s}}(i)}$. Recalling (B.2), wean obtain

$$E[\triangle \boldsymbol{R}_S(N)] = \boldsymbol{0}. \tag{B.4}$$

Similar to the result of (A.4), the first-order expansion of $\boldsymbol{R}_S(N)+\triangle\boldsymbol{R}_S(N)$ in terms of $\triangle\boldsymbol{R}_S(N)$ can be expressed as

$$(\boldsymbol{R}_S(N)+\triangle\boldsymbol{R}_S(N))^{-1} \simeq \boldsymbol{R}_S^{-1}(N)-\boldsymbol{R}_S^{-1}(N)\triangle\boldsymbol{R}_S(N)\boldsymbol{R}_S^{-1}(N). \tag{B.5}$$

Thus, the TLS estimate for $\boldsymbol{A}^{mT}$ can be given by

$$\begin{aligned}\boldsymbol{A}_{TLS}^{mT} &\simeq (\boldsymbol{R}_S^{-1}(N)-\boldsymbol{R}_S^{-1}(N)\triangle\boldsymbol{R}_S(N)\boldsymbol{R}_S^{-1}(N))\hat{\boldsymbol{S}}^*\boldsymbol{x}^m \\ &\simeq \left(\boldsymbol{I}_K+\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)\triangle\boldsymbol{R}_{\boldsymbol{S}}(N)\right)\boldsymbol{A}^{mT} \\ &\quad+\left(\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)-\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)\triangle\boldsymbol{R}_{\boldsymbol{S}}(N)\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)\right)\hat{\boldsymbol{S}}^*\boldsymbol{e}_m,\end{aligned} \tag{B.6}$$

and its mean is

$$\begin{aligned}E[\boldsymbol{A}_{TLS}^{mT}] &\simeq \left(\boldsymbol{I}_K+\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)E[\triangle\boldsymbol{R}_{\boldsymbol{S}}(N)]\right)\boldsymbol{A}^{mT} \\ &\quad+\left(\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)-\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)E[\triangle\boldsymbol{R}_{\boldsymbol{S}}(N)]\boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)\right)\boldsymbol{S}^*E[\boldsymbol{e}_m] \\ &= \boldsymbol{A}^{mT},\end{aligned} \tag{B.7}$$

and its auto-covariance is

$$\begin{aligned}\boldsymbol{C}_{\boldsymbol{A}_{TLS}^{mT}} &\triangleq E[(\boldsymbol{A}_{TLS}^m-\boldsymbol{A}^m)^T(\boldsymbol{A}_{TLS}^m-\boldsymbol{A}^m)^*] \\ &= \boldsymbol{R}_S^{-1}(N)E\left[\hat{\boldsymbol{S}}^*\boldsymbol{e}_m\boldsymbol{e}_m^H\hat{\boldsymbol{S}}\right]\boldsymbol{R}_S^{-H}(N) \\ &= \sigma_m^2\boldsymbol{R}_S^{-1}(N)\left(\boldsymbol{R}_S(N)+\sum_{i=1}^{N}\boldsymbol{C}_{\hat{\boldsymbol{s}}(i)}\right)\boldsymbol{R}_S^{-H}(N),\end{aligned} \tag{B.8}$$

where we have made use of the fact that $\lim\limits_{N\to\infty}\triangle\boldsymbol{R}_S(N)\simeq\boldsymbol{0}$. Similarly, the cross-covariance is given by

$$\begin{aligned}\boldsymbol{C}_{\boldsymbol{A}_{TLS}^{pT},\boldsymbol{A}_{TLS}^{qT}} &\triangleq E[(\boldsymbol{A}_{TLS}^p-\boldsymbol{A}^p)^T(\boldsymbol{A}_{TLS}^q-\boldsymbol{A}^q)^*] \\ &= \boldsymbol{R}_S^{-1}(N)E\left[\hat{\boldsymbol{S}}^*\boldsymbol{e}_p\boldsymbol{e}_q^H\hat{\boldsymbol{S}}\right]\boldsymbol{R}_S^{-H}(N) \\ &= \boldsymbol{R}_{\boldsymbol{e}}(p,q)\boldsymbol{R}_S^{-1}(N)\left(\boldsymbol{R}_S(N)+\sum_{i=1}^{N}\boldsymbol{C}_{\hat{\boldsymbol{s}}(i)}\right)\boldsymbol{R}_S^{-H}(N),\quad p\neq q.\end{aligned} \tag{B.9}$$

Appendix C

Derivation of the Mean and Covariance of $\triangle \boldsymbol{d}_k$ in Eq.(3.43)

Using the result in Appendix B, we can get the asymptotic estimate of $\boldsymbol{A}^{mT}$

$$\hat{\boldsymbol{A}}^{mT} \triangleq \lim_{N\to\infty} \boldsymbol{A}_{TLS}^{mT} \simeq \boldsymbol{A}^{mT} + \boldsymbol{R}_{\boldsymbol{S}}^{-1}(N)\hat{\boldsymbol{S}}^{*}\boldsymbol{e}_m \triangleq \boldsymbol{A}^{mT} + \triangle\boldsymbol{A}^{mT}, \tag{C.1}$$

where $\triangle\boldsymbol{A}^{mT} \triangleq [\triangle a^m(\theta_1), \triangle a^m(\theta_2), \cdots, \triangle a^m(\theta_K)]$ denotes the error vector of the estimated $\hat{\boldsymbol{A}}^{mT}$. Using (B.8) and (B.9) and omitting the second noise term, it is easy to show that $\triangle a^m(\theta_k)$ is a zero mean Gaussian random variable with variance $\sigma^2_{\triangle a^m(\theta_k)} = [\boldsymbol{C}_{\boldsymbol{A}^m_{TLS}}]_{k,k}$. As we assume that the array noise is uncorrelated or i.i.d Gaussian noise, the covariance of $\triangle a^m(\theta_k)$ equals zero, i.e., $\sigma_{\triangle a^p(\theta_k),\triangle a^q(\theta_k)} \triangleq E[\triangle a^p(\theta_k)\triangle a^{q*}(\theta_k)] = 0, \quad p \neq q$. Let us define

$$\hat{d}_k^m \triangleq \hat{a}^{m-1}(\theta_k)\hat{a}^{m*}(\theta_k) \simeq e^{jd_m\psi_k} + \triangle a^{m-1}(\theta_k)e^{j\varphi_k^m} + \triangle a^{m*}(\theta_k)e^{-j\varphi_k^{m-1}}, \tag{C.2}$$

where $\triangle d_k^m \triangleq \triangle a^{m-1}(\theta_k)e^{j\varphi_k^m} + \triangle a^{m*}(\theta_k)e^{-j\varphi_k^{m-1}}$ denotes the estimation error of $d_k^m \triangleq e^{jd_m\psi_k}$. From (C.1) and (C.2), we can easily verify that $\hat{d}_k^m$ is a random Gaussian variable. Therefore, its mean equals $E[\hat{d}_k^m] = e^{jd_m\psi_k}$ and its variance is given by

$$\sigma^2_{\triangle d_k^m} = E[\triangle d_k^m \triangle d_k^{m*}] = \sigma^2_{\triangle a^m(\theta_k)} + \sigma^2_{\triangle a^{m-1}(\theta_k)}. \tag{C.3}$$

Similarly, the covariance of $\hat{d}_k^m$ is

$$\begin{aligned}
\sigma_{\triangle d_k^p, \triangle d_k^q} &= E[\triangle d_k^p \triangle d_k^{q*}] \\
&= E[\triangle a^{p-1}(\theta_k)\triangle a^{(q-1)*}(\theta_k)]e^{j\varphi_k^p}e^{-j\varphi_k^q} \\
&\quad + E[\triangle a^{p*}(\theta_k)\triangle a^q(\theta_k)]e^{-j\varphi_k^{p-1}}e^{j\varphi_k^{q-1}} \\
&= 0, \quad p \neq q.
\end{aligned} \tag{C.4}$$

Appendix D

Statistics of $\mathbf{Re}(\triangle g_k^{(m)})$ and $\mathbf{Im}(\triangle g_k^{(m)})$ in Eq.(4.26)

From (4.18) and (4.19), we have the estimation error vector $\triangle \boldsymbol{g}_m = [\triangle g_1^{(m)}, \triangle g_2^{(m)}, \cdots, \triangle g_K^{(m)}]^T = \boldsymbol{\Pi} \boldsymbol{n}_m$, where $\boldsymbol{\Pi} \triangleq (\boldsymbol{P}^* \boldsymbol{P}^T)^{-1} \boldsymbol{P}^*$ is a deterministic complex matrix and $\boldsymbol{g}_m$ is a zero-mean complex Gaussian random vector with covariance $\frac{\sigma_m^2 [\boldsymbol{R}_{pp}^{(-1)}]_{k,k}}{N}$. We can rewrite the estimation error $\triangle g_k^{(m)}$ as

$$\begin{aligned}
\triangle g_k^{(m)} &= \mathrm{Re}(\triangle g_k^{(m)}) + j\mathrm{Im}(\triangle g_k^{(m)}) \\
&= [\mathrm{Re}(\boldsymbol{\Pi}(k,:)) + j\mathrm{Im}(\boldsymbol{\Pi}(k,:))][\mathrm{Re}(\boldsymbol{n}_m) + j\mathrm{Im}(\boldsymbol{n}_m)] \\
&= \begin{bmatrix} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{bmatrix}^T \begin{bmatrix} \mathrm{Re}\left(\boldsymbol{n}_m\right) \\ -\mathrm{Im}\left(\boldsymbol{n}_m\right) \end{bmatrix} + j \begin{bmatrix} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{bmatrix}^T \begin{bmatrix} \mathrm{Im}\left(\boldsymbol{n}_m\right) \\ \mathrm{Re}\left(\boldsymbol{n}_m\right) \end{bmatrix}.
\end{aligned} \tag{D.1}$$

Since $\boldsymbol{n}_m$ is i.i.d. complex Gaussian noise, it is easy to verify that $\triangle g_k^{(m)}$, as a linear combination of $\boldsymbol{n}_m$, is a complex Gaussian random variable. Hence, the real part $\mathrm{Re}\left(\triangle g_k^{(m)}\right) = \left(\begin{bmatrix} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{bmatrix}^T \begin{bmatrix} \mathrm{Re}\left(\boldsymbol{n}_m\right) \\ -\mathrm{Im}\left(\boldsymbol{n}_m\right) \end{bmatrix} \right)$ is also Gaussian and its

expectation can be obtained as

$$E\left[\mathrm{Re}\left(\triangle g_k^{(m)}\right)\right] = \left[\mathrm{Re}\left(\boldsymbol{\Pi}(k,:)\right) \mathrm{Im}\left(\boldsymbol{\Pi}(k,:)\right)\right] E\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{n}_m\right) \\ -\mathrm{Im}\left(\boldsymbol{n}_m\right) \end{array}\right] = 0 \tag{D.2}$$

where we have used the assumptions $E\left[\mathrm{Re}\left(\boldsymbol{n}_m\right)\right] = \mathbf{0}$ and $E\left[\mathrm{Im}\left(\boldsymbol{n}_m\right)\right] = \mathbf{0}$. The variance of $\mathrm{Re}\left(\triangle b_k^{(m)}\right)$ can then be computed as

$$\begin{aligned}
&\mathrm{Var}\left(\mathrm{Re}\left(\triangle g_k^{(m)}\right)\right) \triangleq E\left[\left(\mathrm{Re}\left(\triangle g_k^{(m)}\right)\right)^2\right] \\
&= \left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right]^T E\left[\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{n}_m\right) \\ -\mathrm{Im}\left(\boldsymbol{n}_m\right) \end{array}\right]\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{n}_m\right) \\ -\mathrm{Im}\left(\boldsymbol{n}_m\right) \end{array}\right]^T\right]\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right] \\
&= \frac{\sigma_m^2}{2}\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right]^T \left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right] \\
&= \frac{\sigma_m^2}{2}\left(\boldsymbol{\Pi}(k,:)\boldsymbol{\Pi}^H(k,:)\right) = \frac{\sigma_m^2}{2N}\left[\boldsymbol{R}_{pp}^{-1}\right]_{k,k}.
\end{aligned} \tag{D.3}$$

Similarly, we can obtain $E\left[\mathrm{Im}\left(\triangle g_k^{(m)}\right)\right] = 0$ and $\mathrm{Var}\left(\mathrm{Im}\left(\triangle g_k^{(m)}\right)\right) = \frac{\sigma_m^2}{2N}\left[\boldsymbol{R}_{pp}^{-1}\right]_{k,k}$. Therefore, with the estimated noise variance $\hat{\sigma}_m^2$ in (4.25), the statistics of the real and imaginary parts of $\triangle g_k^{(m)}$ have been obtained. Finally, we show that $\mathrm{Re}(\triangle g_k^{(m)})$ and $\mathrm{Im}(\triangle g_k^{(m)})$ are uncorrelated, namely,

$$\begin{aligned}
&\mathrm{Cov}\left(\mathrm{Re}\left(\triangle g_k^{(m)}\right), \mathrm{Im}\left(\triangle g_k^{(m)}\right)\right) \triangleq E\left[\mathrm{Re}\left(\triangle g_k^{(m)}\right) \mathrm{Im}\left(\triangle g_k^{(m)}\right)\right] \\
&= \left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right]^T E\left[\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{n}_m\right) \\ -\mathrm{Im}\left(\boldsymbol{n}_m\right) \end{array}\right]\left[\begin{array}{c} \mathrm{Im}\left(\boldsymbol{n}_m\right) \\ \mathrm{Re}\left(\boldsymbol{n}_m\right) \end{array}\right]^T\right]\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right] \\
&= \frac{\sigma_m^2}{2}\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right]^T \left[\begin{array}{cc} \mathbf{0} & \boldsymbol{I}_N \\ -\boldsymbol{I}_N & \mathbf{0} \end{array}\right]\left[\begin{array}{c} \mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right) \end{array}\right] = \frac{\sigma_m^2}{2} 0 = 0.
\end{aligned} \tag{D.4}$$

Appendix E

Statistics of $\triangle a_k^{(m)}$ in Eq.(4.28)

From (4.26)-(4.28), we get

$$\begin{aligned}
\hat{a}_k^{(m)} &\triangleq \frac{\hat{g}_k^{(m)}\hat{g}_k^{(0)*}}{|\varepsilon_k|^2} \\
&= \frac{(\varepsilon_k e^{-j\varphi_k^{(m)}} + \triangle g_k^{(m)})(\varepsilon_k + \triangle g_k^{(0)*})}{|\varepsilon_k|^2} \\
&= \frac{|\varepsilon_k|^2 e^{-j\varphi_k^{(m)}} + \triangle g_k^{(m)}\varepsilon_k^* + \varepsilon_k e^{-j\varphi_k^{(m)}}\triangle g_k^{(0)*} + \triangle g_k^{(m)}\triangle g_k^{(0)*}}{|\varepsilon_k|^2} \qquad \text{(E.1)} \\
&\approx e^{-j\varphi_k^{(m)}} + \frac{\triangle g_k^{(m)}\varepsilon_k^* + \varepsilon_k e^{-j\varphi_k^{(m)}}\triangle g_k^{(0)*}}{|\varepsilon_k|^2} \\
&\triangleq e^{-j\varphi_k^{(m)}} + \triangle a_k^{(m0)} + \triangle a_k^{(0m)},
\end{aligned}$$

where $\triangle a_k^{(0m)} \triangleq \frac{\triangle g_k^{(m)}\varepsilon_k^*}{|\varepsilon_k|^2}$ and $\triangle a_k^{(m0)} \triangleq \frac{\varepsilon_k e^{-j\varphi_k^{(m)}}\triangle g_k^{(0)*}}{|\varepsilon_k|^2}$ are the first-order estimation errors. In obtaining (E.1), the second-order estimation error term has been ignored due to the fact that $\frac{\triangle g_k^{(m)}\triangle g_k^{(0)*}}{|\varepsilon_k|^2}$ is too small relative to $\hat{a}_k^{(m)}$ to affect the estimation result. By using the results of $\triangle g_k^{(m)}$ in Appendix D, it can easily be verified that both $\triangle a_k^{(0m)}$ and $\triangle a_k^{(m0)}$ are complex Gaussian random variables. Their expectations are, respectively, given by

$$E\left[\triangle a_k^{(0m)}\right] = \frac{E\left[\triangle g_k^{(m)}\right]}{\varepsilon_k} = 0$$

and

$$E\left\{\triangle a_k^{(m0)}\right\} = \frac{E\left[\triangle g_k^{(0)*}\right] e^{-} j\varphi_k^{(m)}}{\varepsilon_k^*} = 0$$

and thus, the expected value of $\triangle a_m^{(k)} = \triangle a_k^{(0m)} + \triangle a_k^{(m0)}$ is also zero. Moreover, the variances of $\triangle a_k^{(0m)}$ and $\triangle a_k^{(m0)}$ are given by

$$\mathrm{Var}(\triangle a_k^{(0m)}) = E\left[\triangle a_k^{(0m)} \triangle a_k^{(0m)*}\right] = \frac{E\left[\triangle g_k^{(m)} \triangle g_k^{(m)*}\right]}{|\varepsilon_k|^2}$$

$$\mathrm{Var}(\triangle a_k^{(m0)}) = E\left[\triangle a_k^{(m0)} \triangle a_k^{(m0)*}\right] = \frac{E\left[\triangle g_k^{(0)} \triangle g_k^{(0)*}\right]}{|\varepsilon_k|^2}$$

Therefore, the variance and covariance of $\triangle a_k^{(m)}$ can be obtained as follows.

$$\begin{aligned}
\mathrm{Var}(\triangle a_k^{(m)}) &= E\left[\left(\triangle a_k^{(0m)} + \triangle a_k^{(m0)}\right)\left(\triangle a_k^{(0m)} + \triangle a_k^{(m0)}\right)^*\right] \\
&= \mathrm{Var}(\triangle a_k^{(0m)}) + \mathrm{Var}(\triangle a_k^{(m0)}) \\
&= \frac{\left(\mathrm{Var}\left(\triangle g_k^{(m)}\right) + \mathrm{Var}\left(\triangle g_k^{(0)}\right)\right)}{|\varepsilon_k|^2} \\
&= \frac{(\sigma_m^2 + \sigma_0^2)\left[\boldsymbol{R}_{pp}^{-1}\right]_{k,k}}{N|\varepsilon_k|^2},
\end{aligned} \tag{E.2}$$

and

$$\begin{aligned}
\mathrm{Cov}(\triangle a_k^{(p)}, \triangle a_k^{(q)}) &\triangleq E[(\triangle a_k^{(0p)} + \triangle a_k^{(p0)})(\triangle a_k^{(0q)} + \triangle a_k^{(q0)})^*] \\
&= E\left[\Delta a_k^{(0p)} \Delta a_k^{(0q)*}\right] + E\left[\Delta a_k^{(p0)} \Delta a_k^{(0q)*}\right] \\
&\quad + E\left[\Delta a_k^{(p0)} \Delta a_k^{(q0)*}\right] + E\left[\Delta a_k^{(0p)} \Delta a_k^{(q0)*}\right] \\
&= \frac{E\left[\Delta b_k^{(q)} \Delta b_k^{(p)*}\right] + e^{-j\varphi_k^{(m)}} E\left[\Delta b_k^{(q)} \Delta b_k^{(0)*}\right]}{|\varepsilon_k|^2} \\
&\quad + \frac{e^{j\varphi_k^{(m)}} E\left[\Delta b_k^{(0)} \Delta b_k^{(p)*}\right] + E\left[\Delta b_k^{(0)} \Delta b_k^{(0)*}\right]}{|\varepsilon_k|^2} \\
&= \frac{\mathrm{Var}\left(\Delta g_k^{(0)}\right)}{|\varepsilon_k|^2} = \frac{\sigma_0^2\left[\boldsymbol{R}_{pp}^{-1}\right]_{k,k}}{N|\varepsilon_k|^2}, p \neq q \in [1, 2, \cdots, M].
\end{aligned} \tag{E.3}$$

Appendix F

Derivation of Eq.(4.38)

Since the ambiguities are solved by (4.32)-(4.37), the estimated value of $\mathbf{\Phi}_k$ can be obtained as

$$\hat{\mathbf{\Phi}}_k = \left[\hat{\varphi}_k^{(1)}, \hat{\varphi}_k^{(2)}, \cdots, \hat{\varphi}_k^{(M)}\right]^T \tag{F.1}$$

where $\hat{\varphi}_k^{(m)} = -\angle\left(\hat{a}_k^{(m)}\right) + 2\pi l_k^{(m)} = -\angle\left(a_k^{(m)}\right) + 2\pi l_k^{(m)} + \triangle\varphi_k^{(m)}$ and $\triangle\varphi_k^{(m)} \triangleq \hat{\varphi}_k^{(m)} - \varphi_k^{(m)}$. Using the first-order Taylor series expansion, we have

$$e^{-j\hat{\varphi}_k^{(m)}} \approx e^{-j\varphi_k^{(m)}} - j\triangle\varphi_k^{(m)} e^{-j\varphi_k^{(m)}} = e^{-j\varphi_k^{(m)}} + \triangle a_k^{(m)}, \tag{F.2}$$

where $\triangle\varphi_k^{(m)} = j\triangle a_k^{(m)} e^{j\varphi_k^{(m)}}$. Considering that the estimation error $\triangle\varphi_k^{(m)}$ is a real quantity, we let

$$\triangle\varphi_k^{(m)} = \mathrm{Re}\left[j\triangle a_k^{(m)} e^{j\varphi_k^{(m)}}\right] = -\mathrm{Re}\left[\triangle a_k^{(m)}\right]\sin(\varphi_k^{(m)}) - \mathrm{Im}\left[\triangle a_k^{(m)}\right]\cos(\varphi_k^{(m)}). \tag{F.3}$$

From Appendices D and E, it is easy to prove that both $\mathrm{Re}\left[\triangle a_k^{(m)}\right]$ and $\mathrm{Im}\left[\triangle a_k^{(m)}\right]$ have an identical distribution $\mathbb{N}(\mathbf{0}, \frac{1}{2}\mathbf{\Omega}_k)$. Therefore, $\triangle\varphi_k^{(m)}$ as a linear combination of $\mathrm{Re}\left[\triangle a_k^{(m)}\right]$ and $\mathrm{Im}\left[\triangle a_k^{(m)}\right]$ is a real-valued Gaussian random variable, and its expectation is

$$\begin{aligned} E\left[\triangle\varphi_k^{(m)}\right] &= E\left[-\mathrm{Re}\left(\triangle a_k^{(m)}\right)\sin(\varphi_k^{(m)}) - \mathrm{Im}\left(\triangle a_k^{(m)}\right)\cos(\varphi_k^{(m)})\right] \\ &= -E\left[\mathrm{Re}\left(\triangle a_k^{(m)}\right)\right]\sin(\varphi_k^{(m)}) - E\left[\mathrm{Im}\left(\triangle a_k^{(m)}\right)\right]\cos(\phi_k^{(m)}) \\ &= 0. \end{aligned} \tag{F.4}$$

The covariance of $\triangle\varphi_k^{(m)}$ can be calculated as

$$\begin{aligned} \mathrm{Cov}\left(\triangle\varphi_k^{(p)}, \triangle\varphi_k^{(q)}\right) &\triangleq E\left[\triangle\varphi_k^{(p)}\triangle\varphi_k^{(q)}\right] \\ &= E\left[\mathrm{Re}\left(\triangle a_k^{(p)}\right)\mathrm{Re}\left(\triangle a_k^{(q)}\right)\right]\sin(\varphi_k^{(p)})\sin(\varphi_k^{(q)}) \\ &\quad + E\left[\mathrm{Im}\left(\triangle a_k^{(p)}\right)\mathrm{Im}\left(\triangle a_k^{(q)}\right)\right]\cos(\varphi_k^{(p)})\cos(\varphi_k^{(q)}) \\ &= \frac{1}{2}\mathrm{Cov}\left(\triangle a_k^{(p)}, \triangle a_k^{(q)}\right). \end{aligned} \tag{F.5}$$

where we have used the fact that $\sin(\varphi_k^{(p)}) = \sin(\varphi_k^{(q)})$ and $\cos(\varphi_k^{(p)}) = \cos(\varphi_k^{(q)})$. Clearly, the variance of $\triangle\varphi_k^{(m)}$ is given by $\mathrm{Var}\left(\triangle\varphi_k^{(m)}\right) \triangleq E\left[\left(\triangle\varphi_k^{(m)}\right)^2\right] = \frac{1}{2}\mathrm{Var}\left(\triangle a_k^{(m)}\right)$.

Appendix G

Statistics of $\hat{\varepsilon}_k^{(m)}$ in Eq.(4.43)

We rewrite (4.40) as

$$\hat{\psi}_k \triangleq \psi_k + \triangle\psi_k = \frac{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\hat{\boldsymbol{\Phi}}_k}{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{d}} \triangleq \frac{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\Phi}_k}{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{d}} + \frac{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\triangle\boldsymbol{\Phi}_k}{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{d}}. \tag{G.1}$$

where $\Delta\boldsymbol{\Phi}_k \triangleq \left[\Delta\varphi_k^{(1)}, \Delta\varphi_k^{(2)}, \ldots, \Delta\varphi_k^{(M)}\right]^T$, and $\Delta\psi_k = \sum_{m=1}^{M} \omega_m \Delta\varphi_k^{(m)}$ with $\omega_m = \left(\frac{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}}{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{d}}\right)_{1,m}$ being the m^{th} element of $\frac{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}}{\boldsymbol{d}^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{d}}$. From Appendix E and (F.5), we can obtain

$$\begin{aligned} e^{j\sum_{i=1}^{m} d_i(\psi_k+\triangle\psi_k)} &\approx e^{j\sum_{i=1}^{m} d_i\psi_k} + j\sum_{i=1}^{m} d_i\triangle\psi_k e^{j\sum_{i=1}^{m} d_i\psi_k} \\ &= e^{j\sum_{i=1}^{m} d_i\psi_k} \\ &\quad + j\sum_{i=1}^{m} d_i e^{j\sum_{i=1}^{m} d_i\psi_k \sum_{m=1}^{M}\omega_m j\Delta a_k^{(m)}} e^{j\sum_{i=1}^{m} d_i\psi_k} \\ &= e^{j\sum_{i=1}^{m} d_i\psi_k} - \sum_{i=1}^{m} d_i \sum_{m=1}^{M} \omega_m \triangle a_k^{(m)} e^{j\sum_{i=1}^{m}(d_i+d_m)\psi_k} \end{aligned} \tag{G.2}$$

such that

$$
\begin{aligned}
\hat{\varepsilon}_k^{(m)} &= \hat{g}_k^{(m)} e^{j\hat{\psi}_k \sum_{i=1}^{m} d_i} \\
&= (\varepsilon_k e^{-j\psi_k \sum_{i=1}^{m} d_i} + \triangle g_k^{(m)})(e^{j \sum_{i=1}^{m} d_i \psi_k} - \sum_{i=1}^{m} d_i \sum_{m=1}^{M} \omega_m \triangle a_k^{(m)} e^{j \sum_{i=1}^{m} (d_i + d_m)\psi_k}) \\
&= \varepsilon_k + \varepsilon_k \sum_{i=1}^{m} d_i \sum_{m=1}^{M} \omega_m \triangle a_k^{(m)} e^{j\psi_k \sum_{i=1}^{m} d_i} + \triangle g_k^{(m)} e^{j\psi_k \sum_{i=1}^{m} d_i} \\
&= \varepsilon_k + \frac{\varepsilon_k \sum_{i=1}^{m} d_i}{|\varepsilon_k|^2} \sum_{m=1}^{M} \omega_m (\Delta g_k^{(m)} \varepsilon_k^* + \varepsilon_k \Delta g_k^{(0)*} e^{-j\psi_k \sum_{i=1}^{m} d_i}) e^{j\psi_k \sum_{i=1}^{m} d_i} \\
&\quad + \Delta g_k^{(m)} e^{j\psi_k \sum_{i=1}^{m} d_i} \\
&= \varepsilon_k + (\sum_{i=1}^{m} d_i \omega_m + 1) \Delta g_k^{(m)} e^{j\psi_k \sum_{i=1}^{m} d_i} + \sum_{i=1}^{m} d_i \sum_{i \neq m}^{M} \omega_i \triangle g_k^{(i)} e^{j\psi_k \sum_{i=1}^{m} d_i} \\
&\quad + \frac{\varepsilon_k \sum_{i=1}^{m} d_i}{\varepsilon_k^*} \triangle g_k^{(0)*} \sum_{i=1}^{M} \omega_i \\
&\triangleq \varepsilon_k + \triangle \varepsilon_k^{(m)}
\end{aligned}
\tag{G.3}
$$

where the second or higher order estimation errors are omitted. (G.3) shows that $\triangle \varepsilon_k^{(m)}$ is a zero-mean Gaussian random variable, whereby one can obtain its covariance

as follows.

$$
\begin{aligned}
[\mathbf{H}_k]_{p,q} &\triangleq E\left\{\Delta\varepsilon_k^{(p)}\Delta\varepsilon_k^{(q)}\right\} \quad p \neq q \\
&= \sum_{i=1}^{p} d_i \sum_{i=1}^{q} d_i E\left\{\Delta g_k^{(0)}\Delta g_k^{(0)*}\right\}\left(\sum_{i=1}^{M}\omega_i\right)^2 + (\sum_{i=1}^{p} d_i\omega_p + 1)\sum_{i=1}^{q} d_i E\left\{\Delta b_k^{(p)}\Delta b_k^{(p)*}\right\} \\
&\quad + (d_q\omega_q + 1)d_p E\left\{\Delta g_k^{(q)}\Delta g_k^{(q)*}\right\} + \sum_{i=1}^{p} d_i \sum_{i=1}^{q} d_i \sum_{i\neq p,q}^{M} \omega_i^2 E\left\{\Delta g_k^{(i)}\Delta g_k^{(i)*}\right\} \\
&= \mathrm{Var}\left(\Delta g_k^{(0)}\right)\left(\sum_{i=1}^{M}\omega_i\right)^2 \sum_{i=1}^{p} d_i \sum_{i=1}^{q} d_i + \mathrm{Var}\left(\Delta g_k^{(p)}\right)(\sum_{i=1}^{p} d_i\omega_p + 1)\sum_{i=1}^{p} d_i \\
&\quad + \mathrm{Var}\left(\Delta g_k^{(q)}\right)(\sum_{i=1}^{q} d_i\omega_q + 1)\sum_{i=1}^{q} d_i + \sum_{i=1}^{p} d_i \sum_{i=1}^{q} d_i \sum_{i\neq p,q}^{M} \omega_i^2 \mathrm{Var}\left(\Delta g_k^{(i)}\right)
\end{aligned}
\tag{G.4}
$$

Obviously, the variance of $\Delta\varepsilon_k^{(m)}$ is given by

$$
\begin{aligned}
[\mathbf{H}_k]_{m,m} &\triangleq \mathrm{Var}\left(\triangle\varepsilon_k^{(m)}\right) = E\left\{\Delta\varepsilon_k^{(m)}\triangle\varepsilon_k^{(m)*}\right\} \\
&= (\sum_{i=1}^{m} d_i\omega_m + 1)^2 E\left\{\triangle g_k^{(m)}\triangle g_k^{(m)*}\right\} + \left(\sum_{i=1}^{m} d_i\right)^2 \sum_{i\neq m}^{M} \omega_i^2 E\left\{\triangle g_k^{(i)}\triangle g_k^{(i)*}\right\} \\
&\quad + E\left\{\triangle g_k^{(0)}\triangle g_k^{(0)*}\right\}\left(\sum_{i=1}^{m} d_i \sum_{i=1}^{M}\omega_i\right)^2 \\
&= \left(\sum_{i=1}^{m} d_i\right)^2 \sum_{i=1}^{M} \omega_i^2 \mathrm{Var}\left(\triangle g_k^{(i)}\right) + \left(\sum_{i=1}^{m} d_i \sum_{i=1}^{M}\omega_i\right)^2 \mathrm{Var}\left(\triangle g_k^{(0)}\right) \\
&\quad + (2\sum_{i=1}^{m} d_i\omega_m + 1)\mathrm{Var}\left(\triangle g_k^{(m)}\right).
\end{aligned}
\tag{G.5}
$$

Appendix H

Proof of Lemma 1 in Section 4.3

Let us consider the following partitioned $(m+n) \times K$ matrix,

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{X}^T & \boldsymbol{Y}^T \end{bmatrix}^T, \tag{H.1}$$

both $\boldsymbol{X}$ and $\boldsymbol{Y}$ being full column rank matrices with the dimensions of $m \times K$ and $m \times K$, respectively, such that we can implement the QRD for $\boldsymbol{X}$ and $\boldsymbol{Y}$ as

$$\boldsymbol{X} = \boldsymbol{Q}_X \begin{bmatrix} \boldsymbol{A}^T & \mathbf{0}_{(m-K)\times K}^T \end{bmatrix}^T \quad \text{and} \quad \mathbf{Y} = \mathbf{Q}_Y \begin{bmatrix} \mathbf{B}^T & \mathbf{0}_{(n-K)\times K}^T \end{bmatrix}^T, \tag{H.2}$$

where $\boldsymbol{Q}_X$ and $\boldsymbol{Q}_Y$ are unitary matrices, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are the corresponding unique upper triangular matrices of full rank. Next, constructing an unitary matrix as

$$\boldsymbol{Q}_Z \triangleq \begin{bmatrix} \boldsymbol{Q}_X & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \boldsymbol{Q}_Y \end{bmatrix} \tag{H.3}$$

and using (H.2), we have

$$\boldsymbol{Q}_z^H \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{A}^T & \mathbf{0}_{(m-K)\times K}^T & \boldsymbol{B}^T & \mathbf{0}_{(n-K)\times K}^T \end{bmatrix}^T. \tag{H.4}$$

Since the null submatrices in (H.4) have no effect on the construction of matrix $\boldsymbol{Q}_Z$ based on the QRD technique, the construction of $\boldsymbol{Q}_Z$ is equivalent to finding a unitary matrix for evaluating the QRD of $\begin{bmatrix} \boldsymbol{A}^T & \boldsymbol{B}^T \end{bmatrix}^T$. End of the proof.

Appendix I

Proof of Lemma 2 in Section 4.3

Let us consider the following two block matrices

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I}_K & \boldsymbol{S} \\ -\boldsymbol{S}^H & \boldsymbol{I}_m \end{bmatrix} \quad \text{and} \quad \boldsymbol{T} = \begin{bmatrix} \boldsymbol{A}^T & \boldsymbol{B}^T \end{bmatrix}^T, \tag{I.1}$$

where $\boldsymbol{A}$ is an invertible matrix of order K, $\boldsymbol{B}$ is an $m \times K$ matrix and $\boldsymbol{S}$ is a $K \times m$ matrix. From (I.1), we can get

$$\boldsymbol{G}\mathbf{T} = \begin{bmatrix} \boldsymbol{I}_K & \boldsymbol{S} \\ -\boldsymbol{S}^H & \boldsymbol{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A}^T & \boldsymbol{B}^T \end{bmatrix}^T = \begin{bmatrix} \boldsymbol{A} + \boldsymbol{S}\boldsymbol{B} \\ -\boldsymbol{S}^H\mathbf{A} + \boldsymbol{B} \end{bmatrix}. \tag{I.2}$$

Obviously, by setting $-\mathbf{S}^H\mathbf{A} + \mathbf{B} = \mathbf{0}$, we can obtain a solution for $\boldsymbol{S}$, namely, $\boldsymbol{S} = \boldsymbol{A}^{-H}\boldsymbol{B}^H$, which is a $(m+K) \times (m+K)$ nonsingular lower triangular matrix.

Let us consider another $(m+K)\times(m+K)$ block diagonal matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{C}_1 & \mathbf{0}_{K\times m} \\ \mathbf{0}_{m\times K} & \boldsymbol{C}_2 \end{bmatrix}, \tag{I.3}$$

where C_1 is a $K\times K$ invertible matrix and C_2 is a $m\times m$ invertible matrix. By using (I.1) and (I.3), we can obtain

$$\begin{aligned} \boldsymbol{\Sigma G}(\boldsymbol{\Sigma G})^H &= \begin{bmatrix} \boldsymbol{C}_1 & \mathbf{0}_{K\times m} \\ \mathbf{0}_{m\times K} & \boldsymbol{C}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_K & \boldsymbol{S} \\ -\boldsymbol{S}^H & \boldsymbol{I}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_K & -\boldsymbol{S} \\ \boldsymbol{S}^H & \boldsymbol{I}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_1^H & \mathbf{0}_{K\times m} \\ \mathbf{0}_{m\times K} & \boldsymbol{C}_2^H \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{C}_1 & \mathbf{0}_{K\times m} \\ \mathbf{0}_{m\times K} & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_K{+}\mathbf{S}\boldsymbol{S}^H & \mathbf{0}_{K\times m} \\ \mathbf{0}_{m\times K} & \boldsymbol{I}_m{+}\boldsymbol{S}^H\boldsymbol{S} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_1^H & \mathbf{0}_{K\times m} \\ \mathbf{0}_{m\times K} & \boldsymbol{C}_2^H \end{bmatrix}. \end{aligned} \tag{I.4}$$

It is easy to verify that the RHS of (I.4) is an identity matrix if we set $\boldsymbol{C}_1\left(\boldsymbol{I}_K+\boldsymbol{S}\boldsymbol{S}^H\right)\boldsymbol{C}_1^H=\boldsymbol{I}_K$ and $\boldsymbol{C}_2\left(\boldsymbol{I}_m+\boldsymbol{S}\boldsymbol{S}^H\right)\boldsymbol{C}_2^H=\boldsymbol{I}_m$. Then, the matrix $\boldsymbol{\Sigma G}$ is unitary. Therefore, we get

$$\boldsymbol{C}_1^{-1}\boldsymbol{C}_1^{-H}=\left(\boldsymbol{I}_K+\boldsymbol{S}\boldsymbol{S}^H\right) \text{ and } \quad \boldsymbol{C}_2^{-1}\boldsymbol{C}_2^{-H}=\left(\boldsymbol{I}_m+\boldsymbol{S}^H\boldsymbol{S}\right). \tag{I.5}$$

Since the two symmetric nonnegative definite matrices in (I.5) can be rewritten as $\left(\boldsymbol{I}_K+\boldsymbol{S}\boldsymbol{S}^H\right)\triangleq\boldsymbol{D}_1^H\boldsymbol{D}_1$ and $\left(\boldsymbol{I}_P+\boldsymbol{S}^H\boldsymbol{S}\right)\triangleq\boldsymbol{D}_2^H\boldsymbol{D}_2$, where $\boldsymbol{D}_1\triangleq\boldsymbol{C}_1^{-H}$ and $\boldsymbol{D}_2\triangleq\boldsymbol{C}_2^{-H}$ are the corresponding Cholesky factors, by substituting $\boldsymbol{S}=\boldsymbol{A}^{-H}\boldsymbol{B}^H$ in (I.4) we have that $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are the unique upper triangular matrices of $\boldsymbol{A}^{-H}\left(\boldsymbol{A}^H\boldsymbol{A}+\boldsymbol{B}^H\boldsymbol{B}\right)\boldsymbol{A}^{-1}$ and $\left(\boldsymbol{B}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{B}^H+\boldsymbol{I}_m\right)$. Thus, there exists a unitary matrix such that

$\boldsymbol{Q}^H = \begin{bmatrix} \boldsymbol{C}_1 & \boldsymbol{C}_1\boldsymbol{A}^{-H}\boldsymbol{B}^H \\ -\boldsymbol{C}_2\boldsymbol{B}\boldsymbol{A}^{-1} & \boldsymbol{C}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_1^{-H} & \boldsymbol{D}_1^{-H}\boldsymbol{A}^{-H}\boldsymbol{B}^H \\ -\boldsymbol{D}_2^{-H}\mathbf{B}\boldsymbol{A}^{-1} & \boldsymbol{D}_2^{-H} \end{bmatrix}$ to annihilate the matrix $\boldsymbol{B}$ and produce a new matrix, that is,

$$\begin{aligned} \boldsymbol{Q}^H\mathbf{T} &= \begin{bmatrix} \boldsymbol{D}_1^{-H} & \boldsymbol{D}_1^{-H}\boldsymbol{A}^{-H}\boldsymbol{B}^H \\ -\boldsymbol{D}_2^{-H}\boldsymbol{B}\boldsymbol{A}^{-1} & \boldsymbol{D}_2^{-H} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}^T & \boldsymbol{B}^T \end{bmatrix}^T \\ &= \begin{bmatrix} \boldsymbol{D}_1^{-H}(\boldsymbol{A} + \boldsymbol{A}^{-H}\boldsymbol{B}^H\boldsymbol{B}) \\ \mathbf{0}_{m\times K} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_1\boldsymbol{A} \\ \mathbf{0}_{m\times K} \end{bmatrix}. \end{aligned} \tag{I.6}$$

This ends the proof.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. Wiley-Interscience, 2004.

[2] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Mag.*, vol. 13, no. 4, pp. 67–94, 1996.

[3] R. Zekavat and R. M. Buehrer, *Handbook of Position Location: Theory, Practice and Advances*. John Wiley and Sons, 2011.

[4] P. S. Naidu, *Sensor Array Signal Processing*. CRC Press, 2010.

[5] G. W. Stimson, *Introduction to Airborne Radar*. Scitech Pub., 1998.

[6] S. Stergiopoulos, *Advanced Signal Processing Handbook*. CRC Press, 2010.

[7] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[8] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Springer, 2008.

[9] D. H. Rogstad, A. Mileant, and T. T. Pham, *Antenna Arraying Techniques in the Deep Space Network*. John Wiley and Sons, 2005.

[10] L. C. Godara, "Application of antenna arrays to mobile communications: I-I. beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[11] ——, *Smart Antennas*. CRC Press, 2004.

[12] D. H. Johnson and D. E. Dudgeon, *Array signal processing: concepts and techniques*. Simon and Schuster, 1992.

[13] M. C. Dogan and J. M. Mendel, "Applications of cumulants to array processing. I. aperture extension and array calibration," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1200–1216, 1995.

[14] J.-F. Gu, W.-P. Zhu, and M. N. S. Swamy, "A spatio-temporal stacking approach for estimating two-dimensional direction-of-arrival," in *Proc. of IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2010, pp. 1–4.

[15] L. Gan, J.-F. Gu, and P. Wei, "Estimation of 2-D DOA for noncircular sources using simultaneous SVD technique," *IEEE Antennas Wireless Propag. Lett.*, vol. 7, pp. 385–388, 2008.

[16] W. A. Gardner, "Simplification of MUSIC and ESPRIT by exploitation of cyclostationarity," *Proc. IEEE*, vol. 76, no. 7, pp. 845–847, 1988.

[17] I. Ziskind and Y. Bar-Ness, "Localization by narrow-band autoregressive sources by passive sensor arrays," *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 484–487, 1992.

[18] I. Ziskind and D. Hertz, "Maximum-likelihood localization of narrow-band autoregressive sources via the EM algorithm," *IEEE Trans. Signal Process.*, vol. 41, no. 8, pp. 2719–2724, 1993.

[19] W.-K. Ma, T.-H. Hsieh, and C.-Y. Chi, "DOA estimation of quasi-stationary signals with less sensors than sources and unknown spatial noise covariance: a Khatri–Rao subspace approach," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2168–2180, 2010.

[20] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propag.*, vol. 16, no. 2, pp. 172–175, 1968.

[21] M. Rubsamen and A. B. Gershman, "Sparse array design for azimuthal direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5957–5969, 2011.

[22] P. Pal and P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, 2010.

[23] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, 2011.

[24] V. Vasylyshyn, "Direction of arrival estimation using ESPRIT with sparse arrays," in *European Radar Conference*. IEEE, 2009, pp. 246–249.

[25] M. D. Zoltowski and K. T. Wong, "Closed-form eigenstructure-based direction finding using arbitrary but identical subarrays on a sparse uniform Cartesian array grid," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2205–2210, 2000.

[26] J.-F. Gu, W.-P. Zhu, and M. N. S. Swamy, "Minimum redundancy linear sparse subarrays for direction of arrival estimation without ambiguity," in *Proc. of International Symposium on Circuits and Systems (ISCAS)*, 2011, pp. 390–393.

[27] ——, "Accurate DOA estimation via sparse sensor array," in *Proc. of International Symposium on Circuits and Systems (ISCAS)*, 2012, pp. 2585–2588.

[28] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[29] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[30] Y. Hua and T. K. Sarkar, "On SVD for estimating generalized eigenvalues of singular matrix pencil in noise," *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 892–900, 1991.

[31] M. D. Zoltowski and D. Stavrinides, "Sensor array signal processing via a procrustes rotations based eigenanalysis of the ESPRIT data pencil," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 6, pp. 832–861, 1989.

[32] A.-J. van der Veen, P. B. Ober, and E. F. Deprettere, "Azimuth and elevation computation in high resolution DOA estimation," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1828–1832, 1992.

[33] V. S. Kedia and B. Chandna, "A new algorithm for 2-D DOA estimation," *Signal Processing*, vol. 60, no. 3, pp. 325–332, 1997.

[34] A. Swindlehurst and T. Kailath, "Azimuth/elevation direction finding using regular array geometries," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 145–156, 1993.

[35] P. Heidenreich, A. M. Zoubir, and M. Rubsamen, "Joint 2-D DOA estimation and phase calibration for uniform rectangular arrays," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4683–4693, 2012.

[36] Y. Hua, T. K. Sarkar, and D. Weiner, "L-shaped array for estimating 2-D directions of wave arrival," in *Proc. of the 32nd Midwest Symposium on Circuits and Systems*, 1989, pp. 390–393.

[37] N. Tayem and H. M. Kwon, "L-shape 2-dimensional arrival angle estimation with propagator method," *IEEE Trans. Antennas and Propag.*, vol. 53, no. 5, pp. 1622–1630, 2005.

[38] ——, "Reply to comments on L-shape two-dimensional arrival angle estimation with propagator method'," *IEEE Trans. Antennas and Propag.*, vol. 56, no. 5, pp. 1503–1506, 2008.

[39] S. Kikuchi, H. Tsuji, and A. Sano, "Pair-matching method for estimating 2-D angle of arrival with a cross-correlation matrix," *IEEE Antennas and Wireless Propag. Lett.*, vol. 5, no. 1, pp. 35–40, 2006.

[40] J.-F. Gu and P. Wei, "Joint SVD of two cross-correlation matrices to achieve automatic pairing in 2-D angle estimation problems," *IEEE Antennas Wireless Propag. Lett.*, vol. 6, pp. 553–556, 2007.

[41] L. Bai, C.-Y. Peng, and S. Biswas, "Association of DOA estimation from two ULAs," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 6, pp. 1094–1101, 2008.

[42] J.-F. Gu, P. Wei, and H.-M. Tai, "Two-dimensional DOA estimation by cross-correlation matrix stacking," *Circuits, Systems, and Signal Processing*, vol. 30, no. 2, pp. 339–353, 2011.

[43] J. Liang and D. Liu, "Joint elevation and azimuth direction finding using L-shaped array," *IEEE Trans. Antennas and Propag.*, vol. 58, no. 6, pp. 2136–2141, 2010.

[44] G. Wang, J. Xin, N. Zheng, and A. Sano, "Computationally efficient subspace-based method for two-dimensional direction estimation with L-shaped array," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3197–3212, 2011.

[45] J. Gu, P. Wei, and H.-M. Tai, "DOA estimation using cross-correlation matrix," in *Proc. of IEEE International Symposium on Phased Array Systems and Technology (ARRAY)*, 2010, pp. 593–598.

[46] J.-F. Gu, W.-P. Zhu, and M. N. S. Swamy, "Performance analysis of 2-D DOA estimation via L-shaped array," in *Proc. of Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2012, pp. 1–4.

[47] D. H. Johnson, "The application of spectral estimation methods to bearing estimation problems," *Proc. IEEE*, vol. 70, no. 9, pp. 1018–1028, 1982.

[48] A.-J. Van Der Veen, E. F. Deprettere, and A. L. Swindlehurst, "Subspace-based signal analysis using singular value decomposition," *Proc. IEEE*, vol. 81, no. 9, pp. 1277–1308, 1993.

[49] D. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, 2005.

[50] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, 2011.

[51] Y. Jiang, P. Stoica, and J. Li, "Array signal processing in the known waveform and steering vector case," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 23–35, 2004.

[52] L. Xu, P. Stoica, and J. Li, "Complex amplitude estimation in the known steering matrix and generalized waveform case," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1716–1726, 2006.

[53] J. Li and R. Compton Jr, "Maximum likelihood angle estimation for signals with known waveforms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2850–2862, 1993.

[54] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2154–2163, 1995.

[55] M. Cedervall and R. L. Moses, "Efficient maximum likelihood DOA estimation for signals with known waveforms in the presence of multipath," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 808–811, 1997.

[56] H. Li, G. Liu, and J. Li, "Angle estimator for signals with known waveforms," *Electron. Lett.*, vol. 35, no. 23, pp. 1992–1994, 1999.

[57] L. l. Najjar Atallah and S. Marcos, "DOA estimation and association of coherent multipaths by using reference signals," *Signal Processing*, vol. 84, no. 6, pp. 981–996, 2004.

[58] N. Wang, P. Agathoklis, and A. Antoniou, "A new DOA estimation technique based on subarray beamforming," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3279–3290, 2006.

[59] J.-F. Gu, P. Wei, and H.-M. Tai, "Fast direction-of-arrival estimation with known waveforms and linear operators," *IET Signal Process.*, vol. 2, no. 1, pp. 27–36, 2008.

[60] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.,*, vol. 43, no. 1, pp. 95–107, 1995.

[61] P. Comon and G. H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, no. 8, pp. 1327–1343, 1990.

[62] Y. Bar-Shalom, *Tracking and Ddata Association.* Academic Press Professional, Inc., 1987.

[63] R. E. Zarnich, K. L. Bell, and H. L. Van Trees, "A unified method for measurement and tracking of contacts from an array of sensors," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 2950–2961, 2001.

[64] C. Sword, M. Simaan, and E. Kamen, "Multiple target angle tracking using sensor array outputs," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, no. 2, pp. 367–373, 1990.

[65] K. Lo and C. K. Li, "An improved multiple target angle tracking algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 797–805, 1992.

[66] S. Park, C. Ryu, and K. Lee, "Multiple target angle tracking algorithm using predicted angles," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 2, pp. 643–648, 1994.

[67] C. Sastry, E. Kamen, and M. Simaan, "An efficient algorithm for tracking the angles of arrival of moving targets," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 242–246, 1991.

[68] C. Rao, L. Zhang, and L. Zhao, "Multiple target angle tracking using sensor array outputs," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 268–271, 1993.

[69] A. Satish and R. L. Kashyap, "Multiple target tracking using maximum likelihood principle," *IEEE Trans. Signal Process.*, vol. 43, no. 7, pp. 1677–1695, 1995.

[70] P.-J. Chung, J. F. BÖ, A. O. Hero, and Others, "Tracking of multiple moving sources using recursive EM algorithm," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 1, pp. 50–60, 1900.

[71] C. Rao, C. Sastry, and B. Zhou, "Tracking the direction of arrival of multiple moving targets," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1133–1144, 1994.

[72] Y. Zhou, P. C. Yip, and H. Leung, "Tracking the direction-of-arrival of multiple moving targets by passive arrays: Algorithm," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2655–2666, 1999.

[73] H. Yan and H. H. Fan, "Signal-selective DOA tracking for wideband cyclostationary sources," *IEEE Trans. Signal Process.*, vol. 55, no. 5, 2007.

[74] J.-F. Gu, S. C. Chan, W.-P. Zhu, and M. N. S. Swamy, "Joint DOA estimation and source signal tracking with Kalman filtering and regularized QRD RLS algorithm," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 60, no. 1, pp. 46–50, 2013.

[75] J.-F. Gu, W.-P. Zhu, and M. N. S. Swamy, "Sparse linear arrays for estimating and tracking DOAs of signals with known waveforms," in *Proc. of International Symposium on Circuits and Systems (ISCAS)*, 2013.

[76] J.-E. Nilsson and H. Warston, "Radar with separated subarray antennas," in *Proc. of International Radar Conference*. IEEE, 2003, pp. 194–199.

[77] N.-J. Li, J.-F. Gu, and P. Wei, "2-D DOA estimation via matrix partition and stacking technique," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, 2009.

[78] Y. Hua and K. Abed-Meraim, "Techniques of eigenvalues estimation and association," *Digital signal processing*, vol. 7, no. 4, pp. 253–259, 1997.

[79] F. Gao and A. B. Gershman, "A generalized ESPRIT approach to direction-of-arrival estimation," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 254–257, 2005.

[80] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 10, pp. 1553–1560, 1988.

[81] Å. Björck, *Numerical Methods for Least Squares Problems*. SIAM, 1996.

[82] S. Hodges and P. Moore, "Data uncertainties and least squares regression," *Applied Statistics*, pp. 185–195, 1972.

[83] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM Journal on Numerical Analysis*, vol. 17, no. 6, pp. 883–893, 1980.

[84] I. Markovsky and S. Van Huffel, "Overview of total least-squares methods," *Signal processing*, vol. 87, no. 10, pp. 2283–2302, 2007.

[85] T. F. Chan, "An improved algorithm for computing the singular value decomposition," *ACM Trans. Mathematical Software (TOMS)*, vol. 8, no. 1, pp. 72–83, 1982.

[86] P. Stoica and N. Arye, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.

[87] W. X. Zheng, "Fast identification of autoregressive signals from noisy observations," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 52, no. 1, pp. 43–48, 2005.

[88] R. E. Kalman *et al.*, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[89] S. A. BROWN, "Introduction to random signals and applied Kalman filtering with matlab exercises and solutions 3E," 1997.

[90] R. L. Eubank, *A Kalman Filter Primer*. Chapman And Hall/CRC, 2010.

[91] J. A. Apolinário, *QRD-RLS Adaptive Filtering*. Springer, 1980.

[92] W. M. Gentleman, "Least squares computations by Givens transformations without square roots," *IMA Journal of Applied Mathematics*, vol. 12, no. 3, pp. 329–336, 1973.

[93] M. Haardt and J. A. Nossek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1232–1242, 1995.

[94] T. Kariya and H. Kurata, *Generalized Least Squares*. John Wiley and Sons, 2004.

[95] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Pearson/Prentice Hall Upper Saddle River, NJ, 2005.

[96] A. Charnes, E. Frome, and P.-L. Yu, "The equivalence of generalized least squares and maximum likelihood estimates in the exponential family," *Journal of the American Statistical Association*, vol. 71, no. 353, pp. 169–171, 1976.

[97] S. Chatterjee and A. S. Hadi, *Sensitivity Analysis in Linear Regression*. John Wiley and Sons, 2009.

[98] ——, *Regression Analysis by Example*. John Wiley and Sons, 2013.

[99] M. Pesavento and A. B. Gershman, "Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1310–1324, 2001.

[100] M. H. Gruber, *Regression estimators: A comparative study.* JHU Press, 2010.

[101] D. A. Harville, *Matrix Algebra from a Statistician's Perspective.* Springer, 2008.

[102] W. S. McCormick, J. B. Tsui, and V. Bakkie, "A noise insensitive solution to an ambiguity problem in spectral estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 5, pp. 729–732, 1989.

[103] S. M. Kay, *Fundamentals of statistical signal processing, Volume I: Estimation theory.* Prentice Hall, 1993.

[104] G. A. Seber and A. J. Lee, *Linear Regression Analysis.* John Wiley and Sons, 2012.

[105] W. W. Hager, "Updating the inverse of a matrix," *SIAM Review*, vol. 31, no. 2, pp. 221–239, 1989.

[106] S. S. Haykin, *Adaptive Filter Theory, 4E.* Pearson Education India, 2005.

[107] A. H. Sayed, *Adaptive Filters.* John Wiley and Sons, 2008.

[108] K. Liu, S.-F. Hsieh, and K. Yao, "Systolic block Householder transformation for RLS algorithm with two-level pipelined implementation," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 946–958, 1992.

[109] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[110] P. J. Huber, *Robust statistics.* Springer, 2011.

[111] C. Masreliez and R. Martin, "Robust Bayesian estimation for the linear model and robustifying the Kalman filter," *IEEE Trans. Automat. Contr.*, vol. 22, no. 3, pp. 361–371, 1977.

[112] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, 1985.

[113] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," *IEEE Signal Processing Mag.*, vol. 29, no. 4, pp. 61–80, 2012.