Department of Information and Communications Engineering

Advances and New Applications of Spectral Analysis

Majdoddin Esfandiari

Aalto University

DOCTORAL THESES

Aalto University publication series
DOCTORAL THESES 142/2023

Advances and New Applications of Spectral Analysis

Majdoddin Esfandiari

A doctoral thesis completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Electrical Engineering, at a public examination held at the lecture hall T1, Computer Science building of the school on 16 October 2023 at 13:00.

Aalto University
School of Electrical Engineering
Department of Information and Communications Engineering

Supervising professor
Prof. Sergiy A. Vorobyov, Aalto University, Finland

Preliminary examiners
Prof. A. Lee Swindlehurst, University of California, Irvine, USA
Prof. Andreas Jakobsson, Lund University, Sweden

Opponents
Prof. A. Lee Swindlehurst, University of California, Irvine, USA
Prof. Umberto Spagnolini, Politecnico di Milano, Italy

Aalto University publication series
DOCTORAL THESES 142/2023

ISBN 978-952-64-1420-1 (printed)
ISBN 978-952-64-1421-8 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-64-1421-8

Unigrafia Oy
Helsinki 2023

Finland

NORDIC SWAN ECOLABEL
Printed matter
4041-0619

Aalto University

Abstract

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Author
Majdoddin Esfandiari

Name of the doctoral thesis
Advances and New Applications of Spectral Analysis

Publisher School of Electrical Engineering

Unit Department of Information and Communications Engineering

Series Aalto University publication series DOCTORAL THESES 142/2023

Field of research Signal Processing Technology

Manuscript submitted 19 June 2023 **Date of the defence** 16 October 2023

Permission for public defence granted (date) 25 August 2023 **Language** English

☐ **Monograph** ☒ **Article thesis** ☐ **Essay thesis**

Abstract

Spectral analysis is a mathematical tool for modeling signals and extracting information from signals. Among the areas where it finds applications are radar, sonar, speech processing, and communication systems. One of applications of spectral analysis is to model random signals with the so-called rational models like autoregressive (AR) model. Spectral analysis is used for estimating direction-of-arrival (DOA) in array processing as well. In addition, spectral analysis is used for channel estimation in communication systems such as massive/mmWave MIMO systems. It is because channels in these systems can be modeled with multipaths' gains and angles. This thesis proposes methods for the problems of noisy AR parameter estimation, DOA estimation in unknown noise fields, and massive/mmWave MIMO channel estimation and data detection with one-bit ADCs.

The AR model offers a flexible yet simple tool for modeling complex signals. In practical scenarios, the observation noise may contaminate the AR signal. In this thesis, five methods for estimating noisy AR parameter estimation are proposed. In developing the methods, the concepts such as eigendecomposition (ED) and constrained minimization are used.

The most common assumption about the structure of the observation noise in DOA estimation problem is the uniform noise assumption. According to it, the noise covariance matrix is a scaled identity matrix. However, other noise covariance matrix structures such as nonuniform or block-diagonal are more accurate in some practical situations. A generalized least-squares (GLS)-based DOA estimator that takes into consideration the signal subspace perturbation and also enjoys a properly designed DOA selection strategy is proposed for the case of uniform sensor noise. For the case of nonuniform noise, a non-iterative subspace-based (NISB) method is developed which is computationally efficient compared to state-of-the-art competitors. Moreover, a unified approach to DOA estimation in uniform, nonuniform, and block-diagonal sensor noise is presented.

The use of one-bit ADCs instead of high-resolution ADCs is considered as an elegant solution for reducing power consumption of large-scale systems such as massive/mmWave MIMO and radar systems. In this thesis, we use the analogy between binary classification problem and one-bit parameter estimation to develop algorithms for massive MIMO and mmWave systems. In this regard, a method called SE-TMR is developed for one-bit mmWave UL channel estimation. Another method called L1-RLR-TMR is also offered for one-bit mmWave UL channel estimation. At last, the concept of AdaBoost is combined with Gaussian discriminant analysis (GDA) for developing computationally very efficient channel estimators and data detectors. It is shown that the proposed methods which use approximate versions of GDA as weak classifiers in iterations of the AdaBoost-based algorithms are exceptionally efficient, specifically in large-scale systems.

Keywords Autoregressive signals, noisy observation, DOA estimation, nonuniform and block-diagonal noise, one-bit ADC, channel estimation, data detection

ISBN (printed) 978-952-64-1420-1 **ISBN (pdf)** 978-952-64-1421-8

ISSN (printed) 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2023

Pages 158 **urn** http://urn.fi/URN:ISBN:978-952-64-1421-8

Preface

I wish to extend my deepest appreciation to my supervisor Prof. Sergiy A. Vorobyov, whose unwavering guidance and support have been indispensable throughout my doctoral journey. This work would not have been possible without patience, encouragement, and valuable support of Prof. Vorobyov. I consider myself fortunate to have him as my supervisor, and I remain deeply thankful for his substantial contributions to my intellectual development.

I would like to thank Prof. A. Lee Swindlehurst and Prof. Umberto Spagnolini for having taken the time to be the opponents for my defense. Additionally, I would like to express my gratitude to Prof. A. Lee Swindlehurst and Prof. Andreas Jakobsson for providing me valuable insights and feedback as the pre-examiners of my thesis.

I wish to thank my collaborators Prof. Robert W. Heath, Jr. and Prof. Mahmood Karimi, whose expertise and insights were pivotal to the development of my work. I am also thankful to the Nokia Foundation for granting me their incentive scholarship.

I would also like to extend my gratitude to fellow colleagues for their friendship, and support. I am grateful to them for making the workplace enjoyable.

Last but not least, I am grateful to my wife, my parents, and my siblings for their unconditional love and support in the ups and downs of this Journey.

Espoo, September 4, 2023,

Majdoddin Esfandiari

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I M. Esfandiari, S. A. Vorobyov, and M. Karimi. New estimation methods for autoregressive process in the presence of white observation noise. *Signal Processing*, vol. 171, pp. 1-11, Art. no. 107480, 2020.

II M. Esfandiari, S. A. Vorobyov, and M. Karimi. Non-iterative Subspace-based Method for Estimating AR Model Parameters in the Presence of White Noise with Unknown Variance. In *Proceedings of the 56th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, November 2019.

III M. Esfandiari, S. A. Vorobyov, S. Alibani, and M. Karimi. Non-Iterative Subspace-Based DOA Estimation in the Presence of Nonuniform Noise. *IEEE Signal Processing Letters*, vol. 26(6), pp. 848-852, June 2019.

IV M. Esfandiari and S. A. Vorobyov. Enhanced Standard ESPRIT for overcoming Imperfections in DOA Estimation. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, pp. 4375-4379, June 2021.

V M. Esfandiari and S. A. Vorobyov. A Novel Angular Estimation Method in the Presence of Nonuniform Noise. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, pp. 5023-5027, May 2022.

VI M. Esfandiari, S. A. Vorobyov, and R. W. Heath Jr.. Sparsity Enforcing with Toeplitz Matrix Reconstruction Method for mmWave UL Channel Estimation with One-bit ADCs. In *Proceedings of the IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)* , Trondheim, Norway, pp. 141-145,

June 2022.

VII M. Esfandiari, S. A. Vorobyov, and R. W. Heath Jr.. ADMM-Based Solution for mmWave UL Channel Estimation with One-Bit ADCs via Sparsity Enforcing and Toeplitz Matrix Reconstruction. In *Proc. 57th IEEE Int. Conf. Communications, IEEE ICC'23*, Rome, Italy, May 2023.

VIII M. Esfandiari and S. A. Vorobyov. A Unified Approach to DOA Estimation in Unknown Noise Fields Using ULA. , 18 pages, submitted to a journal, April 2023.

IX M. Esfandiari, S. A. Vorobyov, and R. W. Heath Jr.. AdaBoost-Based Efficient Channel Estimation and Data Detection in One-Bit Massive MIMO. , 10 pages, submitted to a journal, June 2023.

Author's Contribution

Publication I: "New estimation methods for autoregressive process in the presence of white observation noise"

The author proposed the idea, derived the main algorithms and results, implemented the numerical experiments and wrote the majority of the article, incorporating comments by the co-authors.

Publication II: "Non-iterative Subspace-based Method for Estimating AR Model Parameters in the Presence of White Noise with Unknown Variance"

The author proposed the idea, developed the algorithms and results for numerical experiments and wrote the majority of the article, incorporating comments by the co-authors.

Publication III: "Non-Iterative Subspace-Based DOA Estimation in the Presence of Nonuniform Noise"

The author derived the main algorithms and results, implemented the numerical experiments and wrote the majority of the article, incorporating comments by the co-authors.

Publication IV: "Enhanced Standard ESPRIT for overcoming Imperfections in DOA Estimation"

The author proposed the idea, developed the main algorithms and results, conducted the numerical experiments and wrote the majority of the article, taking into account the additional comments by the co-authors.

Publication V: "A Novel Angular Estimation Method in the Presence of Nonuniform Noise"

The main idea was proposed by the author, the algorithms and results were derived, the numerical experiments were implemented and the majority of the article was written by him, and the co-authors comments were taken into account.

Publication VI: "Sparsity Enforcing with Toeplitz Matrix Reconstruction Method for mmWave UL Channel Estimation with One-bit ADCs"

The author proposed the idea, developed the main algorithms and results, implemented the numerical experiments and wrote the majority of the article, taking into consideration comments by the co-authors.

Publication VII: "ADMM-Based Solution for mmWave UL Channel Estimation with One-Bit ADCs via Sparsity Enforcing and Toeplitz Matrix Reconstruction"

The main idea was proposed by the author, the algorithms and results were developed, the numerical experiments were conducted and the majority of the article was written by him, and the co-authors comments were taken into account.

Publication VIII: "A Unified Approach to DOA Estimation in Unknown Noise Fields Using ULA"

The main idea was proposed by the author, the algorithms and results were derived, the numerical experiments were conducted and the majority of the article was written by him, and the co-authors comments were taken into consideration.

Publication IX: "AdaBoost-Based Efficient Channel Estimation and Data Detection in One-Bit Massive MIMO"

The main idea was proposed by the author, the algorithms and results were developed, the numerical experiments were implemented and the majority of the article was written by him, and the co-authors comments were incorporated.

Abbreviations

AdaBoost Adaptive Boosting

ADC Analog-to-Digital Converter

ADMM Alternating Direction Method of Multipliers

AR Autoregressive

BCP Bias Compensation Principle

BS Base Station

CB Conventional Beamformer

CP Cyclic Prefix

DAC Digital-to-Analog Converter

DFT Discrete Fourier Transform

DML Deterministic Maximum Likelihood

DOA Direction-of-Arrival

ED Eigendecomposition

ES ESPRIT Enhanced Standard Estimation of Signal Parameters via Rotational Invariance Technique

EU ESPRIT Enhanced Unitary Estimation of Signal Parameters via Rotational Invariance Technique

FB Forward-Backward

GDA Gaussian Discriminant Analysis

GED Generalized Eigendecomposition

GLR Generalized Likelihood Ratio

GLS Generalized Least-Squares

HR Harmonic Retrieval

IMLSE Iterative Maximum Likelihood Subspace Estimation

i.i.d Independent and Identically Distributed

LS Least-Squares

L1-RLR-TMR ℓ_1 Regularized Logistic Regression with TOeplitz Matrix Reconstruction

MIMO Multiple-Input Multiple-Output

mmWave Millimeter Wave

MUSIC Multiple Signal Classification

NISB Non-Iterative Subspace-Based

OFDM Orthogonal Frequency Division Multiplexing

PSD Positive Semi-Definite

RCM Reduced Covariance Matrix

SCM Sample Covariance Matrix

SDR Semi-Definite Relaxation

SE-TMR Sparsity Enforcing with Toeplitz Matrix Reconstruction

SIE Shift Invariance Equation

SNR Signal-to-Noise Ratio

SVD Singular Value Decomposition

UL Uplink

ULA Uniform Linear Array

Symbols

$\mathbb{C}$ Complex numbers field

$\boldsymbol{d}_k$ k-th column of identity matrix

$\boldsymbol{I}_n$ $n \times n$ Identity matrix

$\mathbb{R}$ Real numbers field

$\boldsymbol{W}_D$ Normalized discrete Fourier transform matrix

θ DOA

λ Regularization parameter

$(\cdot)^*$ Element-wise complex-conjugate

$(\cdot)^T$ Transpose operator

$(\cdot)^H$ Hermitian operator

$(\cdot)^\dagger$ Moore-Penrose pseudo-inverse operator

$\mathbb{E}\{\cdot\}$ Expectation operator

$|\cdot|$ Cardinality of a set or absolute value of a scalar

$\|\cdot\|_p$ ℓ_p-norm of a vector

$\|\cdot\|_\mathrm{F}$ Frobenius norm of a matrix

$\odot$ Hadamard product

$\otimes$ Kronecker product

$\forall$ For all

bdiag{·} Block-diagonal matrix with the bracketed matrices being the main block-diagonal

$\mathscr{D}\{\cdot\}$ Diagonal matrix made by preserving the main diagonal of the bracketed square matrix

DFT{·} DFT of the bracketed argument

diag{·} Diagonal matrix with the bracketed vector being the main diagonal

$\Im\{\cdot\}$ Imaginary part of the bracketed argument

log(·) Natural logarithm operator

$\mathscr{O}(\cdot)$ Big O notation (algorithm complexity)

$\mathscr{Q}\{\cdot\}$ Element-wise one-bit quantizer

$\Re\{\cdot\}$ Real part of the bracketed argument

rem(a,b) Remainder in the division of a by b

$\mathscr{S}_\lambda\big(\cdot\big)$ Shrinkage operator

sign{·} Element-wise sign operator

$\mathscr{T}(\boldsymbol{\pi})$ Hermitian Toeplitz matrix with the vector $\boldsymbol{\pi}$ being its first column

trace{·} Trace of a square matrix

unvec{·} Unvectorization operator

vec{·} Vectorization operator

$\boldsymbol{\Pi} \succeq 0$ $\boldsymbol{\Pi}$ is a positive semi-definite matrix

$[\boldsymbol{\pi}]_i$ i-th entry of the vector $\boldsymbol{\pi}$

$\mathbf{1}\{\cdot\}$ Indicator function

1. Introduction

The spectral analysis of signals provides a rigorous tool for modeling signals and extracting information from them [1]. It finds applications in radar, sonar, speech processing, and communication to mention just a few. In this thesis, advances and new applications of spectral analysis in three specific areas are considered. These three areas are noisy autoregressive (AR) parameter estimation, direction-of-arrival (DOA) estimation, and one-bit massive/mmWave multiple-input multiple-output (MIMO) uplink (UL) channel estimation/data detection. We briefly review the challenges and open research directions corresponding to each of these three research areas in the sequel.

Employing AR model for characterizing the behavior of a random signal is a good fit in numerous signal processing applications such as speech processing, digital communication, spectral estimation, noise cancellation, biomedical signal processing, and image processing to name just a few [2]– [9].

The noisy p-th order real-valued AR model is given as

$$x(t) = a_1 x(t-1) + a_2 x(t-2) + \ldots + a_p x(t-p) + e(t) = \mathbf{a}^T \mathbf{x}_t + e(t) \tag{1.1}$$

$$y(t) = x(t) + w(t) \tag{1.2}$$

where $e(t)$ denotes the zero mean white driving noise with variance of σ_e^2, $\mathbf{a} = [a_1, a_2, \ldots, a_p]^T$ contains coefficients of the AR model, $\mathbf{x}_t = [x(t-1), x(t-2), \ldots, x(t-p)]^T$, and $w(t)$ is the white observation noise with zero mean and variance σ_w^2. Because of the presence of $w(t)$ in (1.2), the zero lag autocorrelation of the process $y(t)$ is biased. Therfore, the noiseless conventional least-squares (LS)-based solution leads to a biased estimation for the AR coefficients $\{a_i\}_{i=1}^p$ [19]. To remedy this issue, several methods have been proposed in the literature where the bias compensation principle (BCP) is used as the key idea [20]– [24].

DOA estimation problem is another area of spectral analysis covered in this thesis. The received signal by a uniform linear array (ULA) at the time instant t is expressed as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \tag{1.3}$$

where $\boldsymbol{\theta} \triangleq [\theta_1, \theta_2, \ldots, \theta_L]^T$ is the vector of the source DOAs, $\mathbf{A}(\boldsymbol{\theta}) \triangleq [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_L)]$ is the array manifold with $\mathbf{a}(\theta_l) =$

$[1,\ e^{-j2\pi\sin(\theta_l)d/\lambda},\ldots,e^{-j2\pi(M-1)\sin(\theta_l)d/\lambda}]^T \in \mathbb{C}^M$ being the steering vector corresponding to DOA θ_l for $l = 1,\ldots,L$, $\mathbf{s}(t) \triangleq [s_1(t),s_2(t),\ldots,s_L(t)]^T \in \mathbb{C}^L$ are the source signals, and $\mathbf{n}(t) \in \mathbb{C}^M$ denotes the complex Gaussian sensor noise vector. Several assumptions can be regarded concerning the structure of the second-order statistics of the observation noise in (1.3). Most common assumptions are uniform white and nonuniform white noise. A spatially block-correlated noise assumption may be more accurate in some applications as well [25]. Numerous DOA estimation methods have been proposed in the literature for the cases of uniform [26]– [53], nonuniform [54]– [66], and some for block-correlated [67]– [72] sensor noise.

Spectral analysis also finds novel applications in communications. Due to high propagation loss, mmWave channels can be considered to be sparse in angular domain. Hence, the channel between a base station (BS) with a ULA containing M antennas and the k-th user (equipped with single antenna) can be formulated as

$$\begin{aligned}\mathbf{h}_k &= \sum_{l=1}^{L_k}\sum_{m=1}^{M_{\text{path}}^{k,l}} \gamma_{k,l,m}\mathbf{a}(\theta_{k,l,m}) \\ &= [\mathbf{A}(\boldsymbol{\theta}_{k,1}),\mathbf{A}(\boldsymbol{\theta}_{k,2}),\ldots,\mathbf{A}(\boldsymbol{\theta}_{k,L_k})]\begin{bmatrix}\boldsymbol{\gamma}_{k,1}\\ \boldsymbol{\gamma}_{k,2}\\ \vdots\\ \boldsymbol{\gamma}_{k,L_k}\end{bmatrix} = \mathbf{A}(\boldsymbol{\theta}_k)\boldsymbol{\gamma}_k \end{aligned} \tag{1.4}$$

where L_k is the number of multipath clusters, $M_{\text{path}}^{k,l}$ denotes the number of paths existing in the l-th cluster scattered in an angular area [73], $\gamma_{k,l,m}$ and $\theta_{k,l,m}$ represent the gain and DOA the m-th path of the l-th cluster, respectively, $\mathbf{a}(\theta_{k,l,m}) \triangleq [1,e^{-j\pi\sin(\theta_{k,l,m})},\ldots,e^{-j(M-1)\pi\sin(\theta_{k,l,m})}]^T \in \mathbb{C}^{M\times 1}$, $\boldsymbol{\theta}_{k,l} \triangleq [\theta_{k,l,1},\theta_{k,l,2},\ldots,\theta_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{R}^{M_{\text{path}}^{k,l}\times 1}$ for $l = 1,2,\ldots,L_k$, $\mathbf{A}(\boldsymbol{\theta}_{k,l}) \triangleq [\mathbf{a}(\theta_{k,l,1}),\mathbf{a}(\theta_{k,l,2}),\ldots,\mathbf{a}(\theta_{k,l,M_{\text{path}}^{k,l}})] \in \mathbb{C}^{M\times M_{\text{path}}^{k,l}}$, $\boldsymbol{\gamma}_{k,l} \triangleq [\gamma_{k,l,1},\gamma_{k,l,2},\ldots,\gamma_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{C}^{M_{\text{path}}^{k,l}\times 1}$, $\boldsymbol{\theta}_k \triangleq [\boldsymbol{\theta}_{k,1}^T,\boldsymbol{\theta}_{k,2}^T,\ldots,\boldsymbol{\theta}_{k,L_k}^T]^T$, $\mathbf{A}(\boldsymbol{\theta}_k) \triangleq [\mathbf{A}(\boldsymbol{\theta}_{k,1}),\mathbf{A}(\boldsymbol{\theta}_{k,2}),\ldots,\mathbf{A}(\boldsymbol{\theta}_{k,L_k})]$, and $\boldsymbol{\gamma}_k \triangleq [\boldsymbol{\gamma}_{k,1}^T,\boldsymbol{\gamma}_{k,2}^T,\ldots,\boldsymbol{\gamma}_{k,L_k}^T]^T$. In (1.4), the dependency of channels on DOAs clearly appears. The use of one-bit analog-to-digital converters (ADCs) instead of high-resolution ADCs is considered as an elegant solution for reducing power consumption of large-scale systems like massive/mmWave MIMO systems. Due to preserving signs of the received signal only, the conventional algorithms developed for high-resolution ADCs may not be suitable for the one-bit ADCs configuration. Therefore, new methods should be devised for tasks like channel estimation and data detection when one-bit ADCs are used. Numerous one-bit channel estimators and data detectors have been proposed in the literature [74]– [105]. In this thesis, we use the angular sparsity of mmWave channels in (1.4), and also the analogy between binary classification problem and one-bit parameter estimation to develop one-bit channel estimators and data detectors for

massive MIMO and mmWave systems.

1.1 Objectives

The objective of this thesis is to develop accurate yet efficient algorithms for estimating noisy AR parameters, estimating DOAs in unknown noise fields, estimating channels and detecting transmitted data in large-scale MIMO systems when one-bit ADCs are deployed.

1.2 Contributions

- In Publication I, four methods are developed for noisy AR parameter estimation. These methods exploits both low-order and high-order Yule-Walker equations to find out the AR coefficients.

- In Publication II, a non-iterative subspace-based method for estimating noisy AR parameters is proposed. The essence of this method is to transform the problem into a generalized eigenvalue problem and then find the variance of the observation noise.

- In Publication III, a non-iterative subspace-based method called NISB is developed for the case of nonuniform sensor noise. NISB has two phases. In the first phase an initial estimate of the noise subspace is obtained with the help of eigendecomposition (ED) of a reduced covariance matrix (RCM), while a refined noise subspace estimate and a noise covariance matrix estimate are obtained in the second phase.

- In Publication IV, the enhanced standard ESPRIT (ES ESPRIT) and its unitary extension are presented as DOA estimators for uniform sensor noise case. These methods take into account the signal subspace perturbation and also use a DOA selection strategy designed for picking up the final DOAs from previously generated DOA candidates.

- In Publication V, an iterative DOA estimation method for the case of nonuniform sensor noise is developed. This method uses generalized eigendecomposition (GED) and LS to update the noise subspace estimate and the noise covariance estimate, respectively. The main advantage of this method is that only a few iterations is sufficient for achieving proper accuracy.

- In Publication VI, we use the concept of Toeplitz matrix reconstruction along with the sparsity of mmWave channels in the discrete Fourier transform (DFT) domain to propose a one-bit mmWave UL channel estimator.

- In Publication VII, an optimization problem for estimating one-bit mmWave UL channels is designed that combines ℓ_1 logistic regression with Toeplitz matrix reconstruction. Then, a computationally efficient alternating direction method of multipliers (ADMM)-based [120] solution is developed for that optimization problem.

- In Publication VIII, a unified approach for DOA estimation in the presence of unknown noise fields is presented. This approach has three connected steps. In the first step, the unknown noise covariance matrix is estimated for nonuniform and block-diagonal noise. Then, double number of DOA candidates is generated using a rooting-based method in the second step. In the third step, a DOA selection strategy is proposed to pick up the final DOA estimates.

- In Publication IX, adaptive boosting (AdaBoost)-based [118], [119] channel estimator and data detector are proposed for one-bit MIMO-OFDM system operating over frequency selective channels. The Gaussian discriminant analysis (GDA) classifier/approximate GDA classifiers [117] are used as weak classifiers in each iteration of the proposed AdaBoost-based algorithms. The main advantage of those AdaBoost-based methods that use approximate versions of the GDA classifier is that they are highly efficient.

1.3 Thesis Structure

The remainder of this thesis is organized as follows. Chapter 2 discusses the noisy AR parameter estimation problem and presents the methods proposed in Publications I and II. Chapter 3 presents the DOA estimation algorithms proposed in Publications III, IV, V, and VIII. Chapter 4 presents the one-bit channel estimation and data detection algorithms for large-scale MIMO systems proposed in Publications VI, VII, and IX.

2. Noisy Autoregressive (AR) Parameter Estimation

Employing AR model for characterizing the behavior of a random signal is a good fit in numerous signal processing applications such as speech processing, digital communication, spectral estimation, noise cancellation, biomedical signal processing, and image processing to name just a few [2]– [9]. Among modern data science applications, the use of AR modeling in, for example, annual population assessment [10], climate and river flow forecasting [11], [12], and financial time series analysis [13]– [16] is notable. In general, the AR parameter estimation problem can be further sub-categorized as one-dimensional AR estimation problem, multichannel AR estimation problem, and nonlinear AR estimation problem [17], [18].

The AR estimation problem is conventionally solved by applying the LS method to the low-order Yule-Walker equations. In practical scenarios, the existence of observation noise hinders the use of the LS solution of the aforementioned Yule-Walker equations [19]. The reason is rooted in a bias contaminating the zero lag autocorrelation of data caused by white observation noise. The objective of this chapter is to present five noisy AR parameter estimation algorithms. The first four algorithms are from Publication I, whereas the fifth one has been proposed in Publication II.

2.1 Signal Model

The noisy p-th order real-valued AR model is formulated as given in (1.1) and (1.2). Using (1.1) and (1.2), the autocorrelation functions of $y(t)$, $r_x(0)$, and $r_y(0)$ are respectively obtained as

$$r_y(k) = r_x(k) + \sigma_w^2 \delta(k) \tag{2.1}$$

$$r_x(0) = E\{x(t)^2\} = \mathbf{r}_x^T \mathbf{a} + \sigma_e^2 \tag{2.2}$$

$$r_y(0) = \mathbf{r}_x^T \mathbf{a} + \sigma_e^2 + \sigma_w^2 \tag{2.3}$$

where $\mathbf{r}_x = [r_x(1), r_x(2), \ldots, r_x(p)]^T$. For $k \geq 1$, the Yule-Walker equations are introduced as $r_x(k) = \sum_{i=1}^{p} a_i r_x(k-i)$ [2]. Then, the p low-order and q high-order

Yule-Walker equations can be respectively derived as $\mathbf{R}_x\mathbf{a} = \mathbf{r}_x$ and $\mathbf{R}'_x\mathbf{a} = \mathbf{r}'_x$ by considering $1 \le k \le p$ and $p+1 \le k \le p+q$, where

$$\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(-1) & \dots & r_x(1-p) \\ r_x(1) & r_x(0) & \dots & r_x(2-p) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(p-1) & r_x(p-2) & \dots & r_x(0) \end{bmatrix} \tag{2.4}$$

$$\mathbf{R}'_x = \begin{bmatrix} r_x(p) & r_x(p-1) & \dots & r_x(1) \\ r_x(p+1) & r_x(p) & \dots & r_x(2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(p+q-1) & r_x(p+q-2) & \dots & r_x(q) \end{bmatrix} \tag{2.5}$$

$$\mathbf{r}'_x = [r_x(p+1), r_x(p+2), \dots, r_x(p+q)]^T \tag{2.6}$$

Exploiting (2.1) and (2.4)-(2.6), we obtain $\mathbf{R}_y = \mathbf{R}_x + \sigma_w^2\mathbf{I}_p$, $\mathbf{R}'_y = \mathbf{R}'_x$, $\mathbf{r}_y = \mathbf{r}_x$, and $\mathbf{r}'_y = \mathbf{r}'_x$. As a result, the p low-order and q high-order Yule-Walker equations with respect to $y(t)$ can be written as

$$\mathbf{R}_y\mathbf{a} - \sigma_w^2\mathbf{a} = \mathbf{r}_y \tag{2.7}$$

$$\mathbf{R}'_y\mathbf{a} = \mathbf{r}'_y \tag{2.8}$$

We multiply (2.7) by $\mathbf{R}_y^{-1}$ and rearrange the result to get $\mathbf{a} = \mathbf{R}_y^{-1}\mathbf{r}_y + \sigma_w^2\mathbf{R}_y^{-1}\mathbf{a}$, in which the term $\mathbf{a}_{\mathrm{LS}} = \mathbf{R}_y^{-1}\mathbf{r}_y$ is the so-called conventional LS estimate of $\mathbf{a}$. This estimate is biased though. In order to compensate the bias term $\sigma_w^2\mathbf{R}_y^{-1}\mathbf{a}$, σ_w^2 needs to be estimated. Consequently, the objective of the noisy AR parameter estimation task is to estimate σ_e^2 and σ_w^2, and use the latter to correct the biased LS solution of the noisy AR problem.

2.2 Proposed Methods

2.2.1 The first proposed method

In the first proposed method of Publication I, we presented an iterative method with the aim of reducing the detrimental impact of the term $\sigma_w^2\mathbf{a}$ in (2.7) in each iteration. Let $\mathbf{c} = \mathbf{a} - \mathbf{\Delta}$ denote the estimate of $\mathbf{a}$ obtained in the previous iteration with $\|\mathbf{\Delta}\|_2 \ll \|\mathbf{a}\|_2$. Then, there are $p-1$ pair-wise orthonormal vectors $\mathbf{b}_i$ ($i = 1, 2, \dots, p-1$) that span the null space of $\mathbf{c}$, that is,

$$\mathbf{b}_i^T\mathbf{c} = 0\ ,\ \|\mathbf{b}_i\|_2^2 = 1\ ,\ \mathbf{b}_i^T\mathbf{b}_j = 0\ ,\ \ i,j = 1,\dots,p-1\ ,\ \ i \ne j\ . \tag{2.9}$$

Therefore, by multiplying (2.7) by $\mathbf{b}_i^T$'s we obtain

$$\mathbf{b}_i^\intercal\mathbf{R}_y\mathbf{a} = \mathbf{b}_i^\intercal\mathbf{r}_y + \sigma_w^2\mathbf{b}_i^\intercal\mathbf{a}\ , \qquad i = 1,\dots,p-1\ . \tag{2.10}$$

placing the definition of $\mathbf{c}$ into (2.10) and using (2.9), we get

$$\sigma_w^2 \mathbf{b}_i^T \mathbf{a} = \sigma_w^2 \mathbf{b}_i^T (\mathbf{c} + \Delta) = \sigma_w^2 \mathbf{b}_i^T \Delta \approx 0 \ , \qquad i = 1, \ldots, p-1 \ . \tag{2.11}$$

where the term $\sigma_w^2 \mathbf{b}_i^T \Delta$ can be interpreted as a negligible error. Note that as $\mathbf{c}$ approaches the actual $\mathbf{a}$, the approximation of (2.11) becomes more precise. Adding q (arbitrary integer larger than one) high-order Yule-Walker equations of (2.8) to $p-1$ equations obtain by combining (2.10) and (2.11), a linear system of equations can be formed as $\mathbf{Ha} = \mathbf{h}$ with the following definitions:

$$\mathbf{H} = \begin{bmatrix} \mathbf{B}(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \\ \mathbf{R}'_y \end{bmatrix} \ , \ \mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{p-1}]^T \ , \ \mathbf{h} = \begin{bmatrix} \mathbf{B}\mathbf{r}_y \\ \mathbf{r}'_y \end{bmatrix} \ . \tag{2.12}$$

Thus, a new update of $\mathbf{a}$ in the current iteration can be derived as $\mathbf{a} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{h}$.

Instead of initializing the first proposed method with $\mathbf{a}_{\mathrm{LS}}$, which is a popular initial vector, we developed a method in Publication I to estimate a proper initial value for σ_w^2, denoted by $\hat{\sigma}_w^{2(0)}$ (see Publication I for details). Then, $\hat{\mathbf{a}}(0)$ can be constructed using (2.7) as $\hat{\mathbf{a}}^{(0)} = (\hat{\mathbf{R}}_y - \hat{\sigma}_w^{2(0)}\mathbf{I}_p)^{-1}\hat{\mathbf{r}}_y$. In the l^{th} iteration, the updates of $\mathbf{a}$ and σ_w^2 can be expressed as $\hat{\mathbf{a}}^{(l)} = (\hat{\mathbf{H}}^{(l)T}\hat{\mathbf{H}}^{(l)})^{-1}\hat{\mathbf{H}}^{(l)T}\hat{\mathbf{h}}^{(l)}$ and $\hat{\sigma}_w^{2(l)} = \frac{\hat{\mathbf{a}}^{(l)T}(\hat{\mathbf{R}}_y\hat{\mathbf{a}}^{(l)} - \hat{\mathbf{r}}_y)}{\|\hat{\mathbf{a}}^{(l)}\|^2}$, respectively. After terminating the iterations, we calculate $\hat{\sigma}_e^2 = \hat{r}_y[0] - \hat{\mathbf{r}}_y^T\hat{\mathbf{a}} - \hat{\sigma}_w^2$ where $\hat{\mathbf{a}}$ and $\hat{\sigma}$ are the output of the aforementioned iterations. Algorithms 1 and 2 in Publication I outlines the steps of the first proposed method.

2.2.2 The second proposed method

The essence of the second proposed method is to design a constrained optimization problem, in which the LS cost function of the low-order Yule-Walker equations is regarded as the objective function of the optimization problem, while the first high-order Yule-Walker equation is imposed as an equality constraint. Using (2.7) and (2.8), the aforementioned optimization problem can be written as

$$\begin{aligned} &\underset{\mathbf{a}\ ,\ \sigma_w^2}{\text{minimize}} && \left(\left(\mathbf{R}_y - \sigma_w^2\mathbf{I}_p\right)\mathbf{a} - \mathbf{r}_y\right)^{\mathsf{T}}\left(\left(\mathbf{R}_y - \sigma_w^2\mathbf{I}_p\right)\mathbf{a} - \mathbf{r}_y\right) \\ &\text{subject to} && \bar{\mathbf{r}}^{\mathsf{T}}\mathbf{a} = r_y(p+1) \end{aligned} \tag{2.13}$$

where $\bar{\mathbf{r}}$ represents the first row of $\mathbf{R}'_y$. We adopt the Lagrangian multiplier method here to engage the equality constraint in the updated optimization objective function $L(\mathbf{a}, \sigma_w^2)$ as

$$\begin{aligned} L(\mathbf{a}, \sigma_w^2) &= \left(\left(\mathbf{R}_y - \sigma_w^2\mathbf{I}_p\right)\mathbf{a} - \mathbf{r}_y\right)^{\mathsf{T}}\left(\left(\mathbf{R}_y - \sigma_w^2\mathbf{I}_p\right)\mathbf{a} - \mathbf{r}_y\right) + \lambda\left(\mathbf{a}^{\mathsf{T}}\bar{\mathbf{r}} - r_y(p+1)\right) \\ &= \mathbf{a}^{\mathsf{T}}\left(\mathbf{R}_y - \sigma_w^2\mathbf{I}_p\right)^2\mathbf{a} - 2\mathbf{a}^{\mathsf{T}}\left(\mathbf{R}_y - \sigma_w^2\mathbf{I}_p\right)\mathbf{r}_y + \|\mathbf{r}_y\|^2 + \lambda\left(\mathbf{a}^{\mathsf{T}}\bar{\mathbf{r}} - r_y(p+1)\right) \end{aligned} \tag{2.14}$$

where λ is the Lagrangian multiplier. Taking partial derivative of (2.14) with respect to $\mathbf{a}$ and σ_w^2 first, and then equating the resultants lead us to the following relations

$$\mathbf{a} = \left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^{-1} \mathbf{r}_y - \frac{\lambda}{2} \left[\left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^2\right]^{-1} \bar{\mathbf{r}} \,. \tag{2.15}$$

$$\sigma_w^2 = \frac{\mathbf{a}^\mathsf{T} \mathbf{R}_y \mathbf{a} - \mathbf{a}^\mathsf{T} \mathbf{r}_y}{\|\mathbf{a}\|_2^2} \tag{2.16}$$

where (2.16) is equivalent to the result reached in the first proposed method. To determine λ, the result of (2.15) should satisfy the constraint (2.13). Consequently, plugging (2.15) into (2.13) results in

$$\lambda = 2 \frac{\bar{\mathbf{r}}^\mathsf{T} \left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^{-1} \mathbf{r}_y - r_y(p+1)}{\bar{\mathbf{r}}^\mathsf{T} \left[\left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^2\right]^{-1} \bar{\mathbf{r}}} \,. \tag{2.17}$$

At last, combining (2.15) and (2.17) together leads us to the second proposed estimator of $\mathbf{a}$, that is,

$$\mathbf{a} = \left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^{-1} \mathbf{r}_y - \left(\frac{\bar{\mathbf{r}}^\mathsf{T} \left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^{-1} \mathbf{r}_y - r_y(p+1)}{\bar{\mathbf{r}}^\mathsf{T} \left[\left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^2\right]^{-1} \bar{\mathbf{r}}} \right) \left[\left(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p\right)^2\right]^{-1} \bar{\mathbf{r}} \,. \tag{2.18}$$

Analogous to the first proposed method, we employ an iterative method to calculate $\mathbf{a}$ and σ_w^2 using (2.18) and (2.16), respectively. Note that the initialization here is the same as what presented for the first proposed method (See Publication I for details).

2.2.3 The third proposed method

The aim of the third proposed method is to reduce the dimension of the original noisy AR parameters estimation problem from p to only two. We exploit this observation here that σ_w^2 is usually much smaller than the $p-2$ largest eigenvalues of $\mathbf{R}_y$. First, we write $\mathbf{a}$ as a linear combination of the eigenvectors of $\mathbf{R}_y$, and then show that different values of σ_w^2 have negligible impact on $p-2$ of unknown parameters. Using the eigenvectors of $\mathbf{R}_y$ denoted by $\mathbf{v}_m$ for $m = 1, \ldots, p$, we write

$$\mathbf{a} = \sum_{m=1}^{p} \alpha_m \mathbf{v}_m \,. \tag{2.19}$$

Note that we represent the eigenvalues of $\mathbf{R}_y$ by λ_m for $m = 1, \ldots, p$ where $\lambda_1 < \lambda_2 < \ldots < \lambda_p$. Plugging (2.19) into (2.7) and also taking into account the characteristic equations $\mathbf{R}_y \mathbf{v}_m = \lambda_m \mathbf{v}_m$ (for $m = 1, \ldots, p$), we have

$$\mathbf{R}_y \sum_{m=1}^{p} \alpha_m \mathbf{v}_m = \mathbf{r}_y + \sigma_w^2 \sum_{m=1}^{p} \alpha_m \mathbf{v}_m \implies \sum_{m=1}^{p} \alpha_m \lambda_m \mathbf{v}_m = \mathbf{r}_y + \sigma_w^2 \sum_{m=1}^{p} \alpha_m \mathbf{v}_m \,. \tag{2.20}$$

Consequently, by multiplying (2.20) by $\mathbf{v}_m^T$ for $m = 1, \ldots, p$, we obtain

$$\alpha_m \lambda_m = \mathbf{v}_m^{\mathsf{T}} \mathbf{r}_y + \sigma_w^2 \alpha_m \implies \alpha_m = \frac{\mathbf{v}_m^{\mathsf{T}} \mathbf{r}_y}{\lambda_m - \sigma_w^2} , \qquad m = 1, \ldots, p \ . \tag{2.21}$$

Given the fact that $0 < \sigma_w^2 < \lambda_1$, (2.21) implies that varying σ_w^2 does not substantially change the values of α_m for $m \geq 3$ as λ_m is usually much larger than σ_w^2 for $m \geq 3$. In other words, updating α_m for $m \geq 3$ is not necessary in each iteration since changes are negligible. Therefore, only initializing α_m ($m \geq 3$) via selecting a proper σ_w^2 is sufficient which results in having to update only α_1, α_2, and σ_w^2 in each iteration. Using this approximation, (2.19) can be reformulated as

$$\mathbf{a} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \bar{\mathbf{x}}, \ \ \bar{\mathbf{x}} = \sum_{m=3}^{p} \alpha_m \mathbf{v}_m \ . \tag{2.22}$$

Lastly, by exploiting (2.8), and (2.21)-(2.22), the following system of linear equations can be formed:

$$\mathbf{H}_2 \boldsymbol{\alpha} = \mathbf{h}_2 \tag{2.23}$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \ , \ \mathbf{H}_2 = \begin{bmatrix} \bar{\boldsymbol{\Lambda}} \\ \mathbf{R}_y' \mathbf{V} \end{bmatrix} \ , \ \mathbf{h}_2 = \begin{bmatrix} \mathbf{V}^{\mathsf{T}} \mathbf{r}_y \\ \mathbf{r}_y' - \mathbf{R}_y' \bar{\mathbf{x}} \end{bmatrix}$$

$$\bar{\boldsymbol{\Lambda}} = \begin{bmatrix} \lambda_1 - \sigma_w^2 & 0 \\ 0 & \lambda_2 - \sigma_w^2 \end{bmatrix} \ , \ \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \ . \tag{2.24}$$

Here, the LS solution of (2.23) is $\boldsymbol{\alpha} = (\mathbf{H}_2^T \mathbf{H}_2)^{-1} \mathbf{H}_2^T \mathbf{h}_2$. Analogous to previously proposed methods, an iterative procedure can be used to update $\boldsymbol{\alpha}$ and σ_w^2 via (2.23) and (2.16), respectively. Note that initialization is carried out similarly as in the previously proposed methods.

2.2.4 The fourth proposed method

Contrary to three previously proposed methods, the fourth one is a non-iterative method. The objective of this method is to estimate σ_w^2 as the minimum eigenvalue of a properly enlarged autocorrelation matrix. In doing so, we write

$$\mathbf{R}_y(m) = \mathbf{R}_x(m) + \sigma_w^2 \mathbf{I}_m \ , \qquad m \geq 1 \tag{2.25}$$

where

$$\mathbf{R}_y(m) \triangleq \begin{bmatrix} r_y[0] & r_y[-1] & \ldots & r_y[1-m] \\ r_y[1] & r_y[0] & \ldots & r_y[2-m] \\ \vdots & \vdots & \ddots & \vdots \\ r_y[m-1] & r_y[m-2] & \ldots & r_y[0] \end{bmatrix} . \tag{2.26}$$

We show in Publication I that the minimum eigenvalues of $\mathbf{R}_y(m+1)$ and $\mathbf{R}_x(m+1)$ are smaller than the minimum eigenvalues of $\mathbf{R}_y(m)$ and $\mathbf{R}_x(m)$, respectively, for $m \geq p$. Since the minimum eigenvalue of $\mathbf{R}_y(m)$ is equal to the minimum eigenvalue of $\mathbf{R}_x(m)$ plus σ_w^2 according to (2.25), it can be concluded that as m increases, the minimum eigenvalue of $\mathbf{R}_y$ gets closer to σ_w^2. Thus, the minimum eigenvalue of $\mathbf{R}_y(2p)$, for example, is a better estimate for σ_w^2 than the minimum eigenvalue of $\mathbf{R}_y(p)$. Increasing m beyond a certain number is not always beneficial though, as it makes the computational complexity of calculating the minimum eigenvalue of $\mathbf{R}_y(2p)$ prohibitive. We used $m = 2p$ in generating results in Publication I.

2.2.5 The fifth proposed method

In Publication II, another non-iterative method for estimating noisy AR parameters is developed. First, we combine (2.7) and (2.8) to write

$$\mathbf{A}\mathbf{a} = \mathbf{b} + \sigma_w^2 \mathbf{c} \tag{2.27}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{R}_y \\ \mathbf{R}_y' \end{bmatrix} \quad , \quad \mathbf{b} = \begin{bmatrix} \mathbf{r}_y \\ \mathbf{r}_y' \end{bmatrix} \quad , \quad \mathbf{c} = \begin{bmatrix} \mathbf{a} \\ \mathbf{0}_q \end{bmatrix} . \tag{2.28}$$

As $\mathbf{b} \in \mathbb{R}^{(p+q)\times 1}$, a matrix $\mathbf{D}$ with $p+q-1$ orthonormal rows can be found that satisfies $\mathbf{D}\mathbf{b} = \mathbf{0}_{p+q-1}$. As a result, by multiplying (2.27) by $\mathbf{D}$ and reorganizing terms, we obtain

$$(\mathbf{D}\mathbf{A} - \sigma_w^2 \mathbf{E})\mathbf{a} = \mathbf{0}_{p+q-1} \tag{2.29}$$

where $\mathbf{E}$ is composed of the first p columns of $\mathbf{D}$. Note that (2.30) is in the form of a generalized eigenvalue problem with $\mathbf{a}$ and σ_w^2 being a generalized eigenvector and the corresponding generalized eigenvalue, respectively. One can multiply (2.29) by $(\mathbf{D}\mathbf{A} - \sigma_w^2 \mathbf{E})^T$ to build the following quadratic eigenvalue problem:

$$(\mathbf{G}_0 + \sigma_w^2 \mathbf{G}_1 + (\sigma_w^2)^2 \mathbf{G}_2)\mathbf{a} = \mathbf{0}_p \tag{2.30}$$

where

$$\mathbf{G}_0 = \mathbf{A}^T\mathbf{D}^T\mathbf{D}\mathbf{A},\ \mathbf{G}_1 = -(\mathbf{A}^T\mathbf{D}^T\mathbf{E} + \mathbf{E}^T\mathbf{D}\mathbf{A}),$$
$$\mathbf{G}_2 = \mathbf{E}^T\mathbf{E} . \tag{2.31}$$

Although multiple methods can be found in the literature for solving (2.31) , we selected the method that transforms (2.31) into a generalized eigenvalue problem (now with squares matrices) as

$$(\mathbf{P} - \sigma_w^2 \mathbf{Q})\bar{\mathbf{a}} = \mathbf{0}_{2p} \tag{2.32}$$

where

$$\mathbf{P} = \begin{bmatrix} \mathbf{G}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_p \end{bmatrix}, \ \mathbf{Q} = \begin{bmatrix} -\mathbf{G}_1 & -\mathbf{G}_2 \\ \mathbf{I}_p & \mathbf{0} \end{bmatrix}, \ \bar{\mathbf{a}} = \begin{bmatrix} \mathbf{a} \\ \sigma_w^2 \mathbf{a} \end{bmatrix} . \tag{2.33}$$

With respect to the observation noise, we propose to recognize the absolute value of the eigenvalue that has the minimum imaginary component, as the estimated σ_w^2. Finally, $\mathbf{a}$ can be estimated as the LS solution of (2.27).

2.3 Experimental Results

In this section, in addition to simulation examples in Publication I and Publication II, two numerical examples are considered for evaluating the performance of the proposed noisy AR estimators. The normalized root mean squared error (NRMSE) is used here for comparing the accuracy of the methods tested, which is defined as

$$\text{NRMSE} = \frac{\sqrt{((\sum_{m=1}^{M} \|\hat{\mathbf{a}}_m - \mathbf{a}\|^2)/M)}}{\|\mathbf{a}\|}$$

where $\hat{\mathbf{a}}_m$ is the estimate of $\mathbf{a}$ in the m-th trial. The hyperparameters used in this section are the same as in Publication I and Publication II. In the first example, a fourth-order noisy AR process with $\mathbf{a} = [0.55, 0.1550, -0.5495, 0.6241]^T$ and $\sigma_e^2 = 1$ is considered. Fig. 2.1 compares the performance of the proposed methods when the number of data points varies from 200 to 2000 for SNR = 20 dB. It can be observed that the first and third proposed methods outperform other methods. Moreover, the use of the constraint in the optimization problem of (2.13) is crucial for improving the performance of the second proposed method.

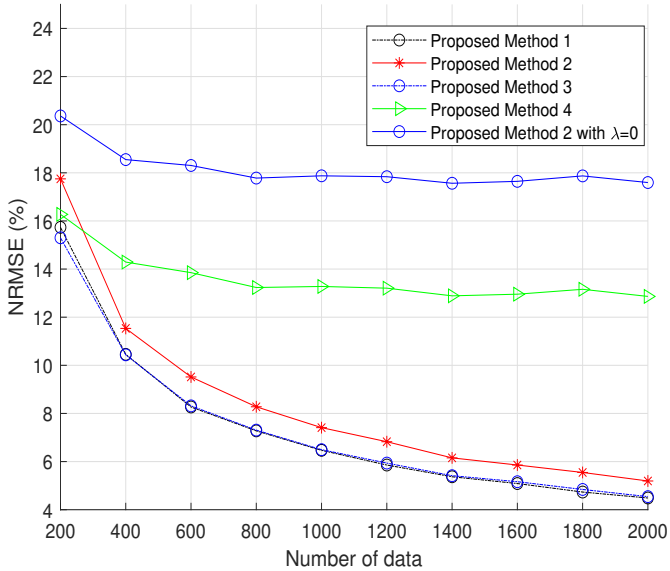

Figure 2.1. NRMSE vs. the number of data points for the first example.

In the second example, a fourth-order noisy AR process with $\mathbf{a} = [1.6771, -1.6875, 0.9433, -0.3164]^T$ and $\sigma_e^2 = 1$ is considered. The number of trials and the number of data points are set to $M = 4000$ and $N = 1000$, respectively. Table 2.1 shows the means and standard deviations obtained from implementing the proposed noisy AR estimators. It can be seen that all proposed methods provide good results in this scenario.

Table 2.1. Computed results of estimated parameters for SNR = 1 dB for the second example.

True value	Xia-Zheng method	Proposed method I	Proposed method II	Proposed method III	Proposed method IV	Proposed method V
$a_1 = 1.6771$	1.5039 ± 0.1547	1.5879 ± 0.1589	1.6298 ± 0.1604	1.6442 ± 0.1446	1.5224 ± 0.1777	1.6088 ± 0.1445
$a_2 = -1.6875$	-1.3542 ± 0.3379	-1.5241 ± 0.2721	-1.5967 ± 0.2682	-1.6138 ± 0.2548	-1.5094 ± 0.2967	-1.5638 ± 0.2456
$a_3 = 0.9433$	0.6325 ± 0.3721	0.8030 ± 0.2258	0.8621 ± 0.2180	0.8698 ± 0.2175	0.7967 ± 0.2412	0.8369 ± 0.2035
$a_4 = -0.3164$	-0.1840 ± 0.2103	-0.2654 ± 0.0816	-0.2850 ± 0.0770	-0.2842 ± 0.0795	-0.2654 ± 0.0848	-0.2772 ± 0.0745
$\sigma_w^2 = 4.6$	4.5279 ± 0.2342	4.5666 ± 0.1992	4.6360 ± 0.2017	4.5986 ± 0.1818	4.5652 ± 0.1921	4.5783 ± 0.1892
$\sigma_e^2 = 1$	1.3014 ± 0.3862	1.1722 ± 0.2934	1.0502 ± 0.2740	1.0487 ± 0.2498	1.2207 ± 0.3478	1.1369 ± 0.2546
$NRMSE(\%)$	29.4308	17.9332	15.9501	15.1377	19.4546	15.5602

3. DOA Estimation in the Presence of Uniform, Nonuniform, and Block-diagonal Sensor Noise

3.1 Signal Model

Considering a ULA composed of M sensors and L narrowband signals emitted by L sources located in the far-field, the signal received by the ULA at the time instant t is given by (1.3). For notation convenience, we use $\mathbf{A}$ instead of $\mathbf{A}(\boldsymbol{\theta})$ throughout this chapter.[1] Exploiting (1.3), the array covariance matrix can be formed as

$$\mathbf{R} \triangleq \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q} \tag{3.1}$$

where $\mathbf{P} \in \mathbb{C}^{L\times L}$ is the signal covariance matrix, and $\mathbf{Q} \in \mathbb{C}^{M\times M}$ is the noise covariance matrix. These matrices are defined as

$$\mathbf{P} \triangleq \mathbb{E}\{\mathbf{s}(t)\mathbf{s}^H(t)\}, \quad \mathbf{Q} \triangleq \mathbb{E}\{\mathbf{n}(t)\mathbf{n}^H(t)\}. \tag{3.2}$$

In this chapter, the problem of DOA estimation for the cases of uniform, nonuniform, and block-diagonal noise covariance matrices is solved. The uniform, nonuniform, and block-diagonal noise covariance matrices are respectively represented as $\mathbf{Q}_{\text{uni}} = \sigma^2\mathbf{I}_M$, $\mathbf{Q}_{\text{nonuni}} = \text{diag}\left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2\right\}$, and $\mathbf{Q}_{\text{bdiag}} = \text{bdiag}\left\{\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_q\right\}$. In the latter, it is worth noting that $\mathbf{Q}_j \in \mathbb{C}^{n_j\times n_j}$ for $j = 1, \ldots, q$.

The sample covariance matrix (SCM) calculated as $\hat{\mathbf{R}} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)\mathbf{x}^H(t) = \frac{1}{N}\mathbf{X}\mathbf{X}^H$ is typically used instead of $\mathbf{R}$, as the latter is unknown in practical scenarios. Note that $\mathbf{X}$ is expressed as

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \tag{3.3}$$

where $\mathbf{X} \triangleq [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(N)]$, $\mathbf{S} \triangleq [\mathbf{s}(1), \mathbf{s}(2), \ldots, \mathbf{s}(N)]$, $\mathbf{N} \triangleq [\mathbf{n}(1), \mathbf{n}(2), \ldots, \mathbf{n}(N)]$, and N is the number of snapshots.

[1] Note that we use the complete notation wherever the parameter of $\mathbf{A}$ is not $\boldsymbol{\theta}$.

3.2 Proposed Methods

3.2.1 ES ESPRIT and EU ESPRIT

We have developed ES ESPRIT and EU ESPRIT DOA estimators in Publication IV for the uniform sensor noise case. These methods first transform the shift invariance equation (SIE) equations into the DFT domain, and then generate $2L$ DOA candidates by solving two different systems of linear equations. These two systems of linear equations are solved using a generalized least-squares (GLS)-based method which takes into account the second-order statistic of the signal subspace perturbation. Afterwards, a properly designed DOA selection strategy is introduced for selecting the final L DOA estimates from $2L$ DOA candidates produced previously.

The noiseless SIE is given as

$$\mathbf{J}_1\mathbf{U}_s\mathbf{\Psi} = \mathbf{J}_2\mathbf{U}_s \ . \tag{3.4}$$

where $\mathbf{J}_1 = [\mathbf{I}_{M-1},\ \mathbf{0}_M]$, $\mathbf{J}_2 = [\mathbf{0}_M,\ \mathbf{I}_{M-1}]$, $\mathbf{U}_s \in \mathbb{C}^{M\times L}$ is the actual signal subspace obtained by applying the truncated singular value decomposition (SVD) on $\mathbf{X}$, and $\mathbf{\Psi} \in \mathbb{C}^{L\times L}$ is a matrix whose eigenvalues λ_l's are related to θ_l's through $\lambda_l = e^{-j\pi\sin(\theta_l)}$ for $l = 1,\ldots,L$. As a result, the aim is to first estimate $\mathbf{\Psi}$, and then obtain θ_l's from its eigenvalues. Multiplying (3.4) by first the DFT matrix $\mathbf{W}_D$ and then the selecting matrix $\mathbf{Z}_{\mathscr{I}} \in \mathbb{R}^{|\mathscr{I}|\times M-1}$, we have

$$\mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_1\mathbf{U}_s\mathbf{\Psi} = \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_2\mathbf{U}_s \tag{3.5}$$

where members of the set $\mathscr{I}$ are the indices of the selected equations. Note that all entries of the i-th row of $\mathbf{Z}_{\mathscr{I}}$ are zero except one entry which is set to 1. The index of this nonzero entry is specified by the i-th member of $\mathscr{I}$. The main reason of multiplying (3.4) by the DFT matrix $\mathbf{W}_D$ is rooted in the relation $\mathbf{U}_s = \mathbf{A}\mathbf{T}^{-1}$, which indicates that the columns of $\mathbf{U}_s$ can be formed by linear combinations of the columns of $\mathbf{A}$. Therefore, multiplying (3.4) by $\mathbf{W}_D$ is a proper choice as the columns of $\mathbf{W}_D$ are structurally matched to the columns of $\mathbf{J}_1\mathbf{U}_s$. Inserting $\mathbf{U}_s = \hat{\mathbf{U}}_s + \mathbf{\Delta U}_s$ into (3.5) and reorganizing the terms, we obtain

$$\mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_1\hat{\mathbf{U}}_s\mathbf{\Psi} + \mathbf{E} = \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_2\hat{\mathbf{U}}_s \tag{3.6}$$

where $\hat{\mathbf{U}}_s$ denotes the estimated signal subspace by applying truncated SVD on the received data, $\mathbf{\Delta U}_s$ is the signal subspace estimation error caused by the observation noise, and $\mathbf{E} \triangleq \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_1\mathbf{\Delta U}_s\mathbf{\Psi} - \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_2\mathbf{\Delta U}_s$. Vectorizing (3.6), we have

$$\hat{\mathbf{f}}_{\mathrm{S}} \triangleq \mathrm{vec}\left\{\mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_2\hat{\mathbf{U}}_s\right\} = \hat{\mathbf{F}}_{\mathrm{S}}\boldsymbol{\psi} + \mathbf{e} = \hat{\mathbf{F}}_{\mathrm{S}}\boldsymbol{\psi} + \hat{\mathbf{G}}_{\mathrm{S}}\mathbf{\Delta u}_s \tag{3.7}$$

with $\hat{\mathbf{F}}_{\mathrm{S}} \triangleq \mathbf{I}_L \otimes \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_1\hat{\mathbf{U}}_s \in \mathbb{C}^{|\mathscr{I}|L\times L^2}$, $\hat{\mathbf{G}}_{\mathrm{S}} \triangleq (\mathbf{\Psi}^T \otimes \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_1) - (\mathbf{I}_L \otimes \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{J}_2) \in \mathbb{C}^{|\mathscr{I}|L\times ML}$, $\boldsymbol{\psi} \triangleq \mathrm{vec}\{\mathbf{\Psi}\} \in \mathbb{C}^{L^2\times 1}$, and $\mathbf{\Delta u}_s \triangleq \mathrm{vec}\{\mathbf{\Delta U}_s\} \in \mathbb{C}^{ML\times 1}$. Analogous to [41],

the covariance matrix of $\mathbf{\Delta u}_s$ can be employed in a GLS-based approach [41], [109], [110] to find the optimal solution of (3.7) as

$$\hat{\boldsymbol{\psi}}_{\mathrm{GLS}} = \left(\hat{\mathbf{F}}_{\mathrm{S}}^{H}\hat{\mathbf{W}}_{\mathrm{S}}\hat{\mathbf{F}}_{\mathrm{S}}\right)^{-1}\hat{\mathbf{F}}_{\mathrm{S}}^{H}\hat{\mathbf{W}}_{\mathrm{S}}\hat{\mathbf{f}}_{\mathrm{S}} \tag{3.8}$$

where as shown in [41] $\hat{\mathbf{W}}_{\mathrm{S}} = \left[\hat{\mathbf{G}}_{\mathrm{S}}\left(\hat{\mathbf{\Sigma}}_s^{-2}\otimes\mathbf{I}_M\right)\hat{\mathbf{G}}_{\mathrm{S}}^{H}\right]^{-1}$ with $\hat{\mathbf{\Sigma}}_s \in \mathbb{R}^{L\times L}$ being a diagonal matrix that contains the L principal singular values of $\mathbf{X}$ (for more detail, see [41] and Publication IV). The DOA's can be extracted from the arguments of the L eigenvalues of $\hat{\mathbf{\Psi}}_{GLS} = \mathrm{unvec}\left\{\hat{\boldsymbol{\psi}}_{GLS}\right\}$. Due to the dependency of $\hat{\mathbf{G}}_{\mathrm{S}}$ on $\mathbf{\Psi}$, it is natural to use an iterative method to estimate $\hat{\mathbf{W}}_{\mathrm{S}}$ and $\hat{\mathbf{\Psi}}_{GLS}$. Based on our observations, few iterations are sufficient to reach accurate results.

The unitary extension of (3.5) is given as

$$\mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{K}_1\mathbf{E}_s\mathbf{\Upsilon} = \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{K}_2\mathbf{E}_s \tag{3.9}$$

where $\mathbf{K}_1 \triangleq 2\Re\{\mathbf{Q}_{M-1}^{H}\mathbf{J}_2\mathbf{Q}_M\} \in \mathbb{R}^{(M-1)\times M}$, $\mathbf{K}_2 \triangleq 2\Im\{\mathbf{Q}_{M-1}^{H}\mathbf{J}_2\mathbf{Q}_M\} \in \mathbb{R}^{(M-1)\times M}$, and the columns of $\mathbf{E}_s$ are the L principal left singular vectors of $\varphi(\overline{\mathbf{X}}) = \mathbf{Q}_M^{H}\overline{\mathbf{X}}\mathbf{Q}_{2N} \in \mathbb{R}^{M\times 2N}$ with $\mathbf{Q}_M$ and $\mathbf{Q}_{2N}$ being left $\mathbf{\Pi}$-real matrices [35]. Similar to the steps presented for solving (3.5), a GLS-based solution of (3.9) is obtained as

$$\hat{\boldsymbol{\nu}}_{GLS} = \left(\hat{\mathbf{F}}_{\mathrm{U}}^{H}\hat{\mathbf{W}}_{\mathrm{U}}\hat{\mathbf{F}}_{\mathrm{U}}\right)^{-1}\hat{\mathbf{F}}_{\mathrm{U}}^{H}\hat{\mathbf{W}}_{\mathrm{U}}\hat{\mathbf{f}}_{\mathrm{U}} \tag{3.10}$$

where $\hat{\boldsymbol{\nu}}_{GLS} \triangleq \mathrm{vec}\{\hat{\mathbf{\Upsilon}}_{GLS}\}$, $\hat{\mathbf{F}}_{\mathrm{U}} \triangleq (\mathbf{I}_L \otimes \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{K}_1\hat{\mathbf{E}}_s)$, $\hat{\mathbf{f}}_{\mathrm{U}} \triangleq \mathrm{vec}\{\mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{K}_2\hat{\mathbf{E}}_s\}$. Moreover, it is showed in [41] that $\hat{\mathbf{W}}_{\mathrm{U}} \triangleq \left[\hat{\mathbf{G}}_{\mathrm{U}}(\hat{\overline{\mathbf{\Sigma}}}_s^{-2}\otimes\mathbf{I}_M)\hat{\mathbf{G}}_{\mathrm{U}}^{H}\right]^{-1}$ with $\hat{\mathbf{G}}_{\mathrm{U}} \triangleq (\hat{\mathbf{\Upsilon}}^{T}\otimes \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{K}_1) - (\mathbf{I}_L \otimes \mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\mathbf{K}_2)$, and $\hat{\overline{\mathbf{\Sigma}}}_s \in \mathbb{R}^{L\times L}$ denoting a diagonal matrix which contains the principal singular values of $\varphi(\overline{\mathbf{X}})$.

In Publication IV, we propose to select members of $\mathscr{I}$ in (3.5) as those indices associated with $|\mathscr{I}|$ largest absolute values of $\mathbf{W}_D\mathbf{J}_1\mathbf{u}_1$. Here, $\mathbf{u}_1$ denotes the left singular vector of $\mathbf{X}$ which corresponds to the largest singular value.

DOA Selection Strategy

For improving DOA estimation accuracy, we first generate $2L$ DOA candidates by implementing the proposed ESPRIT-based methods twice with $|\mathscr{I}| = M-1$ and $|\mathscr{I}| = M-2$, and then select the best L DOAs from $2L$ DOA candidates.

The first DOA selection strategy is to employ the deterministic ML (DML) cost function analogous to papers such as [39], [40], [114]. In this method, the final L DOA estimates are a subset of $2L$ DOA candidates that minimizes the following DML cost function:

$$\hat{\mathbf{\Theta}}_{DML} = \arg\min_{\mathbf{\Theta}_i} \mathrm{trace}((\mathbf{I}_M - \mathbf{A}(\mathbf{\Theta}_i)(\mathbf{A}(\mathbf{\Theta}_i)^{H}\mathbf{A}(\mathbf{\Theta}_i))^{-1}\mathbf{A}(\mathbf{\Theta}_i)^{H})\hat{\mathbf{R}})$$

$$\forall \quad i = 1,\ldots,P_{ESE(EUE)} \tag{3.11}$$

where $\mathbf{\Theta}_i$ represents the i-th DOA subset, and $P_{ESE(EUE)} \triangleq \frac{2L!}{L!L!}$ is the total number of different subsets.

The second DOA selection strategy relies on a generalized likelihood ratio (GLR) method presented in [115]. We propose in Publication IV that the final L DOA estimates are recognized one by one in a sequential manner. Towards this end, we select the l-th (for $l = 1,\ldots,L$) DOA estimate as the member of the set of $2L-(l-1)$ remaining DOAs that maximizes the GLR cost function [115], that is,

$$\hat{\theta}_l = \arg\max_{\theta_i} \frac{\mathbf{a}(\theta_i)^H \mathbf{P}_{l-1}^{\perp} \hat{\mathbf{R}} \mathbf{P}_{l-1}^{\perp} \mathbf{a}(\theta_i)}{\mathbf{a}(\theta_i)^H \mathbf{P}_{l-1}^{\perp} \mathbf{a}(\theta_i)}$$
$$\text{for} \quad i = 1,\ldots,2L-(l-1) \ , \ l = 1,\ldots,L \tag{3.12}$$

where

$$\mathbf{P}_{l-1}^{\perp} \triangleq \begin{cases} \mathbf{I}_M - \mathbf{A}_{l-1}(\mathbf{A}_{l-1}^H \mathbf{A}_{l-1})^{-1} \mathbf{A}_{l-1}^H & l > 1 \\ \mathbf{I}_M & l = 1 \end{cases} \tag{3.13}$$

with $\mathbf{A}_{l-1} \triangleq [\mathbf{a}(\hat{\theta}_1), \mathbf{a}(\hat{\theta}_2), \ldots, \mathbf{a}(\hat{\theta}_{l-1})] \in \mathbb{C}^{M\times(l-1)}$. The advantage of the GLR-based DOA selection strategy over the DML-based one is that the former requires considerably less computations for selecting the final DOA estimates.

3.2.2 NISB

In Publication III, we develop the NISB method for DOA estimation in the presence of nonuniform sensor noise. The NISB is a non-iterative method comprised of two consecutive phases. In the first phase, an initial estimate of the noise subspace is identified by applying ED of a RCM [58]. In the second phase, this initial noise subspace estimate is used for estimating the noise covariance matrix, and then a refined estimate of the noise subspace is found by applying the generalized ED to the pair of SCM and estimated noise covariance matrix. Well-known subspace-based methods such as multiple signal classification (MUSIC), and root-MUSIC can exploit the noise subspace estimate for identifying the unknown DOAs. The NISB method requires substantially lower computational complexity to implement as compared to the iterative methods like IMLSE. It should be pointed out that the performance of the NISB method degrades in the presence of correlated sources.

Recall that the noise subspace matrix $\mathbf{U} \in \mathbb{C}^{M\times(M-L)}$ satisfies the following condition:

$$\mathbf{A}^H \mathbf{U} = \mathbf{0}_{L\times(M-L)} \ . \tag{3.14}$$

Consequently, multiplying (3.1) by a noise subspace estimate denoted by $\hat{\mathbf{U}}$ leads to

$$\hat{\mathbf{R}}\hat{\mathbf{U}} \approx \hat{\mathbf{Q}}_{\text{nonuni}} \hat{\mathbf{U}} \ . \tag{3.15}$$

where $\mathbf{R}$ is replaced by $\hat{\mathbf{R}}$, and $\hat{\mathbf{Q}}_{\text{nonuni}}$ represents an estimate of the nonuniform noise covariance matrix. It is proved in Publication III that $\hat{\mathbf{U}}$ can be obtained

as the $M-L$ eigenvectors of the generalized ED of the pair $\hat{\mathbf{R}}$ and $\hat{\mathbf{Q}}_{\text{nonuni}}$ which correspond to the $M-L$ smallest eigenvalues. The NISB method uses this result to find the refined noise subspace. A proper $\hat{\mathbf{Q}}_{\text{nonuni}}$, however, needs to be found first. Towards this end, the initial noise subspace estimate, denoted by $\hat{\mathbf{U}}_{\text{ini}}$, can be found as the $M-L$ eigenvectors of the RCM corresponding to the $M-L$ smallest eigenvalues [58]. Note that the RCM is formed as $\hat{\mathbf{R}}_{\text{RCM}} = \hat{\mathbf{R}} - \mathscr{D}\{\hat{\mathbf{R}}\}$. We can write $\hat{\mathbf{Q}}_{\text{nonuni}}$ as

$$\hat{\mathbf{Q}}_{\text{nonuni}} = \sigma^2 \mathbf{I}_M + \mathbf{Q}_{\text{nun}} \tag{3.16}$$

where σ^2 denotes the uniform part of sensor noise variances, and $\mathbf{Q}_{\text{nun}}$ is a diagonal matrix with one of its diagonal entries being zero. We consider the position of this zero entry as the position of the smallest diagonal entry of $\hat{\mathbf{R}}$. With k being the index of the aforementioned zero diagonal entry, we can first insert (3.16) into (3.15) and then multiply the resultant by the unit vector $\mathbf{d}_k^T$ which yields

$$\mathbf{d}_k^T \hat{\mathbf{R}} \hat{\mathbf{U}}_{\text{ini}} \approx \mathbf{d}_k^T (\sigma^2 \mathbf{I}_M + \mathbf{Q}_{\text{nun}}) \hat{\mathbf{U}}_{\text{ini}} = \sigma^2 \mathbf{d}_k^T \hat{\mathbf{U}}_{\text{ini}} \tag{3.17}$$

where $\hat{\mathbf{U}}$ is replaced by $\hat{\mathbf{U}}_{\text{ini}}$, and the relation $\mathbf{d}_k^T \mathbf{Q}_{\text{nun}} = \mathbf{0}_{1\times M}$ is used. Therefore, an estimate of σ^2 is obtained as

$$\hat{\sigma}^2 = \left| \frac{\mathbf{d}_k^T \hat{\mathbf{R}} \hat{\mathbf{U}}_{\text{ini}} \hat{\mathbf{U}}_{\text{ini}}^H \mathbf{d}_k}{\mathbf{d}_k^T \hat{\mathbf{U}}_{\text{ini}} \hat{\mathbf{U}}_{\text{ini}}^H \mathbf{d}_k} \right|. \tag{3.18}$$

As the last step of forming $\hat{\mathbf{Q}}_{\text{nonuni}}$, an estimate of $\mathbf{Q}_{\text{nun}}$ is considered as

$$\hat{\mathbf{Q}}_{\text{nun}} = \text{diag}\left\{ [\hat{\mathbf{R}}]_{1,1} - c, \ldots, [\hat{\mathbf{R}}]_{M,M} - c \right\} \tag{3.19}$$

where c is the smallest diagonal entry of $\hat{\mathbf{R}}$. The matrix $\hat{\mathbf{Q}}_{\text{nonuni}}$ can be constructed using (3.16), (3.17) and (3.19). At last, the refined noise subspace estimate is computed as the $M-L$ eigenvectors corresponding to the $M-L$ smallest eigenvalues of the generalized ED of the pair matrices $\hat{\mathbf{R}}$ and $\hat{\mathbf{Q}}_{\text{nonuni}}$.

3.2.3 Unified Approach to DOA Estimation in Unknown Noise Fields

In Publication VIII, a unified approach for DOA estimation problem in unknown noise fields is proposed. This approach comprises of three phases designed carefully to handle challenging scenarios with small sample size and/or closely located sources and/or relatively low signal-to-noise ratios (SNRs).

In the first phase, a general approach for nonuniform and also block-diagonal noise covariance estimation is developed, which is applicable to arbitrary array configurations. In the second phase, a GLS-based forward-only DOA estimation method is devised that uses the output of the first phase. In addition, a forward-backward (FB) extension of the aforementioned DOA estimator is developed. Using the GLS-based estimators (forward-only or FB versions) twice leads to the output of the second phase being a set which contains $2L$ DOA candidates. In the third phase, a DOA selection strategy is designed to select the final DOA estimates from the set of DOA candidates.

Nonuniform Noise Covariance Matrix Estimation

The noise covariance matrix estimator is an iterative estimator that utilizes (3.15) to update $\hat{\mathbf{Q}}_{\text{nonuni}}$ and $\hat{\mathbf{U}}$. The LS minimization criteria is used to obtain the update rule for $\hat{\mathbf{Q}}_{\text{nonuni}}$, whereas the GED concept is adopted here for updating $\hat{\mathbf{U}}$.

In the i-th iteration, the columns of $\hat{\mathbf{U}}^{(i)}$ can be estimated as the $M-L$ eigenvectors corresponding to the $M-L$ smallest eigenvalues obtained by performing the GED of the matrices $\left\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}_{\text{nonuni}}^{(i)}\right\}$. This choice is similar to the last step of the NISB method, where it is used in a non-iterative manner though. Noteworthy to mention that $\hat{\mathbf{U}}^{(i)}$ and $\hat{\mathbf{Q}}^{(i)}$ represent the estimates in the i-th iteration.

To update $\hat{\mathbf{Q}}_{\text{nonuni}}$ in the $(i+1)$-st iteration, the following LS minimization problem can be considered using (3.15):

$$\hat{\mathbf{Q}}_{\text{nonuni}}^{(i+1)} = \arg\min_{\mathbf{Q}} f(\mathbf{Q}) = \|(\hat{\mathbf{R}} - \mathbf{Q})\hat{\mathbf{U}}^{(i)}\|_{\text{F}}^2 \,. \tag{3.20}$$

Note that (3.20) should be solved given the constraint that $\mathbf{Q}$ is a diagonal matrix. We show in Publication VIII that the partial derivative of $f(\mathbf{Q})$ with respect to $\mathbf{Q}$ after excluding the constant term can be written as [116]

$$\frac{\partial f(\mathbf{Q})}{\partial \mathbf{Q}} = 2\mathscr{D}\left\{\hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H\right\}\mathbf{Q} - \mathscr{D}\left\{\hat{\mathbf{R}}\hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H\hat{\mathbf{R}}\right\}. \tag{3.21}$$

As a result, the m-th diagonal entry of $\hat{\mathbf{Q}}_{\text{nonuni}}^{(i+1)}$ can be found by equating (3.21) to zero as

$$\hat{\sigma}_m^{2\,(i+1)} = \mathbf{d}_m^T\left(\frac{1}{2}\mathscr{D}\left\{\hat{\mathbf{R}}\hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H\hat{\mathbf{R}}\right\}\mathscr{D}\left\{\hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H\right\}^{-1}\right)\mathbf{d}_m,$$
$$m = 1, 2, \ldots, M \,. \tag{3.22}$$

In the i-th iteration of the proposed nonuniform noise covariance matrix estimator, only the m-th diagonal element of $\hat{\mathbf{Q}}_{\text{nonuni}}$ is updated via (3.22) where $m \triangleq \text{rem}(i, M) + 1$ with $\text{rem}(a, b)$ denoting the remainder in the division of a by b. All columns of $\hat{\mathbf{U}}$ are updated in each iteration. Employing the element-wise update rule for estimating $\hat{\mathbf{Q}}_{\text{nonuni}}$ results in boosting the convergence of the proposed iterative algorithm, in which we terminate the algorithm when the condition $|f^{(i+1)} - f^{(i)}| < \epsilon$ is met. Note that we set $\epsilon = 10^{-4}$ and $\hat{\mathbf{Q}}_{\text{nonuni}}^{(0)} = \mathscr{D}\{\hat{\mathbf{R}}\}$ in Publication VIII.

Noteworthy to mention that we also propose another nonuniform noise covariance matrix estimator in Publication V that use the GED-based approach to update $\hat{\mathbf{U}}$, and use (3.22) to update $\hat{\mathbf{Q}}_{\text{nonuni}}$. The difference here is that all diagonal elements of $\hat{\mathbf{Q}}_{\text{nonuni}}$ are updated together using (3.22), and also the proposed iterative algorithm is terminated early after a few iterations. The reason for the latter is to make the proposed algorithm in Publication V more computationally efficient.

Block-diagonal Noise Covariance Matrix Estimation

The objective here is to develop an iterative block-diagonal noise covariance estimator. First, we rewrite (3.15) for the case of block-diagonal noise as

$$\hat{\mathbf{R}}\hat{\mathbf{U}} \approx \hat{\mathbf{Q}}_{\text{bdiag}}\hat{\mathbf{U}} \,. \tag{3.23}$$

Analogous to the nonuniform noise case, it is observed from (3.23) that $\hat{\mathbf{U}}^{(i)}$ can be considered as the $M-L$ eigenvectors associated with the $M-L$ smallest eigenvalues computed by applying the GED on the pair of matrices $\left\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}_{\text{bdiag}}^{(i)}\right\}$.

For estimating $\hat{\mathbf{Q}}_{\text{bdiag}}$ in the $(i+1)$-st iteration, the following LS minimization problem can be written using (3.23):

$$\hat{\mathbf{Q}}_{\text{bdiag}}^{(i+1)} = \arg\min_{\mathbf{Q}_{\text{bdiag}}} f_{\text{bdiag}}(\mathbf{Q}_{\text{bdiag}}) = \|(\hat{\mathbf{R}} - \mathbf{Q}_{\text{bdiag}})\hat{\mathbf{U}}^{(i)}\|_{\text{F}}^2 \,. \tag{3.24}$$

We show in Publication VIII that the partial derivative of $f_{\text{bdiag}}(\mathbf{Q}_{\text{bdiag}})$ with respect to the Hermitian matrix $\mathbf{Q}_j$ after eliminating the constant term can be expressed as [106]

$$\frac{\partial f_{\text{bdiag}}(\mathbf{Q}_{\text{bdiag}})}{\partial \mathbf{Q}_j} = \mathbf{Q}_j^*(\overline{\mathbf{V}}_{jj}^{(i)})^T + (\overline{\mathbf{V}}_{jj}^{(i)})^T\mathbf{Q}_j^T - (\overline{\mathbf{R}}_{jj}^{(i)})^T \,,$$

$$j = 1, \ldots, q \tag{3.25}$$

where $\overline{\mathbf{V}}_{jj}^{(i)} \in \mathbb{C}^{n_j \times n_j}$ and $\overline{\mathbf{R}}_{jj}^{(i)} \in \mathbb{C}^{n_j \times n_j}$ are respectively defined as the jth block on the main diagonal of $\overline{\mathbf{R}}^{(i)} \triangleq \hat{\mathbf{R}}\hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H\hat{\mathbf{R}}$ and $\overline{\mathbf{V}}^{(i)} \triangleq \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H$. Applying first the transposition operator to (3.25), followed by vectorizing and equating the result to zero, we have

$$\mathbf{V}_j^{(i)}\mathbf{q}_j = \overline{\mathbf{r}}_j^{(i)}, \; j = 1, \ldots, q \tag{3.26}$$

where $\mathbf{V}_j^{(i)} \triangleq \left[\left((\overline{\mathbf{V}}_{jj}^{(i)})^T \otimes \mathbf{I}_{n_j}\right) + \left(\mathbf{I}_{n_j} \otimes \overline{\mathbf{V}}_{jj}^{(i)}\right)\right]$ is a square matrix, $\mathbf{q}_j \triangleq \text{vec}\{\mathbf{Q}_j\}$, and $\overline{\mathbf{r}}_j^{(i)} \triangleq \text{vec}\{\overline{\mathbf{R}}_{jj}^{(i)}\}$. Consequently, by solving the systems of linear equations in (3.26), and using the unvectorization operator, we obtain

$$\hat{\mathbf{Q}}_j^{(i+1)} = \text{unvec}\left\{(\mathbf{V}_j^{(i)})^{-1}\overline{\mathbf{r}}_j^{(i)}\right\}, \quad j = 1, \ldots, q. \tag{3.27}$$

In the i-th iteration, we compute $j = \text{rem}(i, q) + 1$ and update only the j-th block of $\hat{\mathbf{Q}}_{\text{bdiag}}$.

Subspace-Based DOA Estimation via GLS

In the second phase, we first pre-whiten the received signal using the noise covariance matrix estimated in the first phase, and then develop a GLS-based DOA estimation method by taking into account the signal subspace error. For notation simplicity, we use $\mathbf{Q}$ to represent the noise covariance matrix regardless of its structure.

Multiplying (3.3) by $\mathbf{Q}^{-\frac{1}{2}}$, we have

$$\overline{\mathbf{X}} \triangleq \mathbf{Q}^{-\frac{1}{2}}\mathbf{X} = \mathbf{Q}^{-\frac{1}{2}}\mathbf{A}\mathbf{S} + \mathbf{Q}^{-\frac{1}{2}}\mathbf{N} = \mathbf{Q}^{-\frac{1}{2}}\mathbf{A}\mathbf{S} + \overline{\mathbf{N}} \tag{3.28}$$

where the covariance matrix of the columns of $\overline{\mathbf{N}}$ is $\mathbf{I}_M$. Using the truncated SVD, $\overline{\mathbf{X}}$ can be decomposed as

$$\overline{\mathbf{X}} = \mathbf{U}_\mathrm{s}\mathbf{\Sigma}_\mathrm{s}\mathbf{V}_\mathrm{s}^H \tag{3.29}$$

where $\mathbf{U}_\mathrm{s} \in \mathbb{C}^{M\times L}$ and $\mathbf{V}_\mathrm{s} \in \mathbb{C}^{N\times L}$ are respectively the left and right singular vectors corresponding to the L largest singular values on the diagonal of $\mathbf{\Sigma}_\mathrm{s} \in \mathbb{R}^{L\times L}$. Taking (3.28) and (3.29) into account, it can be concluded that the columns of $\mathbf{U}_\mathrm{s}$ and $\mathbf{Q}^{-\frac{1}{2}}\mathbf{A}$ reside in the same vector space. In other words, $\mathbf{U}_\mathrm{s}$ can be written as $\mathbf{U}_\mathrm{s} = \mathbf{Q}^{-\frac{1}{2}}\mathbf{A}\mathbf{G}$ with $\mathbf{G} \in \mathbb{C}^{L\times L}$ being a non-singular matrix. Consequently, we have $\widetilde{\mathbf{U}}_\mathrm{s} \triangleq \mathbf{Q}^{\frac{1}{2}}\mathbf{U}_\mathrm{s} = \mathbf{A}\mathbf{G}$. It is shown in Publication VIII that the DFT of the p-th column of $\widetilde{\mathbf{U}}_\mathrm{s}$ can be expressed as

$$\bar{\mathbf{u}}_p + \mathrm{diag}(\bar{\mathbf{u}}_p)\mathbf{W}_a\mathbf{a} = \overline{\mathbf{W}}\mathbf{b}_p, \quad p = 1,\ldots,L \tag{3.30}$$

where $\bar{\mathbf{u}}_p \triangleq \mathrm{DFT}\{\widetilde{\mathbf{u}}_p\} = \mathbf{W}_D\widetilde{\mathbf{u}}_p$, $\widetilde{\mathbf{u}}_p$ is the p-th column of $\widetilde{\mathbf{U}}_\mathrm{s}$, $\mathbf{W}_a \triangleq [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_M]^T \in \mathbb{C}^{M\times L}$, $\overline{\mathbf{W}} \triangleq [\overline{\mathbf{w}}_1, \overline{\mathbf{w}}_2, \ldots, \overline{\mathbf{w}}_M]^T \in \mathbb{C}^{M\times L}$, $\mathbf{w}_k \triangleq [W_M^k,\ (W_M^k)^2,\ (W_M^k)^3,\ \ldots,\ (W_M^k)^L]^T$, $\overline{\mathbf{w}}_k \triangleq [1,\ W_M^k,\ (W_M^k)^2,\ \ldots,\ (W_M^k)^{L-1}]^T$, and $W_M^k \triangleq e^{-j\frac{2\pi k}{M}}$. In addition, it is shown in Publication VIII that the roots of the polynomial $\gamma^L + \sum_{l=1}^{L}[\mathbf{a}]_l\ \gamma^{L-l} = 0$ are related to the unknown DOAs θ_l via $\gamma_l = e^{-j2\pi d\sin(\theta_l)/\lambda}$ for $l = 1,\ldots,L$ [107, 108]. Therefore, the DOA estimation problem boils down to estimating $\mathbf{a}$.

As the aim is to form a set of DOA candidates with $2L$ members, the selection matrix $\mathbf{Z}_\mathscr{I} \in \mathbb{R}^{|\mathscr{I}|\times M}$, introduced earlier in this chapter, is used to revise (3.30) as

$$\mathbf{Z}_\mathscr{I}\bar{\mathbf{u}}_p + \mathrm{diag}(\mathbf{Z}_\mathscr{I}\bar{\mathbf{u}}_p)\mathbf{Z}_\mathscr{I}\mathbf{W}_a\mathbf{a} = \mathbf{Z}_\mathscr{I}\overline{\mathbf{W}}\mathbf{b}_p, \quad p = 1,\ldots,L. \tag{3.31}$$

Note that different selecting matrices choose different sets of equations in (3.31). It is shown in Publication VIII that choosing $|\mathscr{I}| = M$ and $|\mathscr{I}| = M-1$ ends up in generating the most accurate $2L$ candidates.

To estimate $\mathbf{a}$ via (3.31), we need to first remove the impact of the unknown vectors $\mathbf{b}_p$'s. To do so, $\mathbf{B} \in \mathbb{C}^{|\mathscr{I}|\times(|\mathscr{I}|-L)}$ is obtained such that $\mathbf{B}^H\widetilde{\mathbf{Z}}_\mathscr{I} = \mathbf{0}_{(|\mathscr{I}|-L)\times L}$ with $\widetilde{\mathbf{Z}}_\mathscr{I} \triangleq \mathbf{Z}_\mathscr{I}\overline{\mathbf{W}} \in \mathbb{C}^{|\mathscr{I}|\times L}$. Then, multiplying (3.31) by $\mathbf{B}^H$ results in

$$\mathbf{B}^H(\mathbf{Z}_\mathscr{I}\bar{\mathbf{u}}_p + \mathrm{diag}(\mathbf{Z}_\mathscr{I}\bar{\mathbf{u}}_p)\mathbf{Z}_\mathscr{I}\mathbf{W}_a\mathbf{a}) = \mathbf{0}_{(|\mathscr{I}|-L)},\ p = 1,\ldots,L. \tag{3.32}$$

Reorganizing (3.32), we form $\mathbf{H}_p\mathbf{a} = \mathbf{h}_p$ for $p = 1,\ldots,L$, where $\mathbf{H}_p \triangleq \mathbf{B}^H\mathrm{diag}(\mathbf{Z}_\mathscr{I}\bar{\mathbf{u}}_p)\mathbf{Z}_\mathscr{I}\mathbf{W}_a \in \mathbb{C}^{(|\mathscr{I}|-L)\times L}$ and $\mathbf{h}_p \triangleq -\mathbf{B}^H\mathbf{Z}_\mathscr{I}\bar{\mathbf{u}}_p \in \mathbb{C}^{(|\mathscr{I}|-L)}$. Piling up the L matrices $\mathbf{H}_p$ and the L vectors $\mathbf{h}_p$ into a matrix $\mathbf{H}$ and a vector $\mathbf{h}$, respectively, (3.32) can be recast as

$$\mathbf{H}\mathbf{a} = \mathbf{h} \tag{3.33}$$

where $\mathbf{H} \triangleq [\mathbf{H}_1^T \ldots \mathbf{H}_L^T]^T \in \mathbb{C}^{L(|\mathscr{I}|-L)\times L}$ and $\mathbf{h} \triangleq [\mathbf{h}_1^T \ldots \mathbf{h}_L^T]^T \in \mathbb{C}^{L(|\mathscr{I}|-L)}$.

As only an estimate of $\mathbf{Q}$ can be calculated, (3.28)-(3.33) should be rewritten with this consideration (see Publication VIII). By doing so, (3.33) is rewritten as $\hat{\mathbf{H}}\mathbf{a} \approx \hat{\mathbf{h}}$. Then, a GLS-based estimator of $\mathbf{a}$ is given as

$$\hat{\mathbf{a}} = (\hat{\mathbf{H}}^H\mathbf{W}\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^H\mathbf{W}\hat{\mathbf{h}}. \tag{3.34}$$

where $\mathbf{W} \triangleq \left(\mathbb{E}\{\hat{\mathbf{e}}\hat{\mathbf{e}}^H\}\right)^{-1} \in \mathbb{C}^{L(|\mathscr{I}|-L)\times L(|\mathscr{I}|-L)}$ and $\hat{\mathbf{e}} \triangleq \hat{\mathbf{H}}\mathbf{a} - \hat{\mathbf{h}} \in \mathbb{C}^{L(|\mathscr{I}|-L)}$. We show in Publication VIII that a proper estimate of $\mathbf{W}$ can be calculated as

$$\hat{\mathbf{W}} = (\hat{\mathbf{\Sigma}}_{\mathrm{s}}^2 \otimes (\mathbf{C}(\mathbf{a})\mathbf{C}^H(\mathbf{a}))^{-1}). \tag{3.35}$$

where $\hat{\mathbf{\Sigma}}_{\mathrm{s}}$ is a diagonal matrix which contains the L largest singular values of the matrix $\hat{\bar{\mathbf{X}}} \triangleq \hat{\mathbf{Q}}^{-\frac{1}{2}}\mathbf{X}$, and $\mathbf{C}(\mathbf{a}) \triangleq \mathbf{B}^H\left(\mathbf{I}_{|\mathscr{I}|} + \mathrm{diag}\{\mathbf{Z}_{\mathscr{I}}\mathbf{W}_a\mathbf{a}\}\right)\mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\hat{\mathbf{Q}}^{\frac{1}{2}} \in \mathbb{C}^{(|\mathscr{I}|-L)\times M}$. It is clear from (3.34) and (3.35) that an iterative algorithm should be used to update $\hat{\mathbf{a}}$ and $\hat{\mathbf{W}}$. The algorithm is initialized by $\hat{\mathbf{a}}^{(0)} = \hat{\mathbf{a}}_{LS} = \hat{\mathbf{H}}^{\dagger}\hat{\mathbf{h}}$. Noteworthy to mention that only a few iterations are sufficient to reach a precise result.

At last, we show in Publication VIII that the FB extension of (3.34) can be obtained as

$$\hat{\mathbf{a}} = \left(\hat{\bar{\mathbf{H}}}^H\hat{\mathbf{W}}_{\mathrm{FB}}\hat{\bar{\mathbf{H}}}\right)^{-1}\hat{\bar{\mathbf{H}}}^H\hat{\mathbf{W}}_{\mathrm{FB}}\hat{\bar{\mathbf{h}}} \tag{3.36}$$

where $\hat{\bar{\mathbf{H}}} = \left[\hat{\bar{\mathbf{H}}}_1^T,\ldots,\hat{\bar{\mathbf{H}}}_L^T\right]^T \in \mathbb{C}^{L(|\mathscr{I}|-L)\times L}$, $\hat{\bar{\mathbf{h}}} = \left[\hat{\bar{\mathbf{h}}}_1^T,\ldots,\hat{\bar{\mathbf{h}}}_L^T\right]^T \in \mathbb{C}^{L(|\mathscr{I}|-L)}$, $\hat{\bar{\mathbf{H}}}_p = \mathbf{B}^H\mathrm{diag}(\mathbf{Z}_{\mathscr{I}}\hat{\bar{\mathbf{e}}}_p)\mathbf{Z}_{\mathscr{I}}\mathbf{W}_a \in \mathbb{C}^{(|\mathscr{I}|-L)\times L}$, $\hat{\bar{\mathbf{h}}}_p = -\mathbf{B}^H\mathbf{Z}_{\mathscr{I}}\hat{\bar{\mathbf{e}}}_p \in \mathbb{C}^{(|\mathscr{I}|-L)}$, $\hat{\bar{\mathbf{e}}}_p$ is the p-th column of $\hat{\bar{\mathbf{E}}}_{\mathrm{s}} \triangleq \mathrm{DFT}\{\hat{\bar{\mathbf{Q}}}^{\frac{1}{2}}\hat{\mathbf{E}}_{\mathrm{s}}\} \in \mathbb{C}^{M\times L}$, $\hat{\bar{\mathbf{Q}}} \triangleq \hat{\mathbf{Q}} + \mathbf{J}_M\hat{\mathbf{Q}}^*\mathbf{J}_M$, $\hat{\mathbf{E}}_{\mathrm{s}}$ and $\hat{\mathbf{\Pi}}_{\mathrm{s}}$ are respectively the matrix of the left singular vectors and the diagonal matrix of the L principal singular values of the matrix $\hat{\bar{\mathbf{X}}}_{\mathrm{FB}} = \left[\hat{\bar{\mathbf{Q}}}^{-\frac{1}{2}}\mathbf{X} \quad \hat{\bar{\mathbf{Q}}}^{-\frac{1}{2}}\mathbf{J}_M\mathbf{X}^*\mathbf{J}_N\right]$. Moreover, we have $\hat{\mathbf{W}}_{\mathrm{FB}} = \hat{\mathbf{\Pi}}_{\mathrm{s}}^2 \otimes \left(\mathbf{C}_{\mathrm{FB}}(\mathbf{a})\mathbf{C}_{\mathrm{FB}}(\mathbf{a})^H\right)^{-1}$, $\mathbf{C}_{\mathrm{FB}}(\mathbf{a}) \triangleq \mathbf{B}^H\left(\mathbf{I}_{|\mathscr{I}|} + \mathrm{diag}\{\mathbf{Z}_{\mathscr{I}}\mathbf{W}_a\mathbf{a}\}\right)\mathbf{Z}_{\mathscr{I}}\mathbf{W}_D\hat{\bar{\mathbf{Q}}}^{\frac{1}{2}} \in \mathbb{C}^{(|\mathscr{I}|-L)\times M}$. Similar to the forward-only case, an iterative algorithm should be used to update $\hat{\mathbf{a}}$ and $\hat{\mathbf{W}}_{\mathrm{FB}}$.

After calculating $\hat{\mathbf{a}}$, $\hat{\gamma}_l$'s are obtained as the roots of the polynomial $\gamma^L + \sum_{l=1}^{L}[\hat{\mathbf{a}}]_l\ \gamma^{L-l} = 0$. Then, $\hat{\theta}_l$ for $l = 1,\ldots,L$ are calculated as $\hat{\theta}_l = \arcsin\left(-\frac{\beta_l\lambda}{2\pi d}\right)$ where β_l is the phase argument of $\hat{\gamma}_l$.

DOA Selection Strategy

Given $2L$ DOA candidates, a properly designed DOA selection strategy is required to select the final L DOA estimates. We propose a three-step DOA selection strategy in Publication VIII which exploits the conventional beamformer (CB) [33], [111–113], DML cost function [39], [40], [114], and GLR technique [46], [115].

Step 1: Let $\boldsymbol{\theta}_{2\mathrm{L}}$ represent the vector which contains $2L$ DOA candidates. Then, we calculate the threshold η as the value of the $(L+1)$-st peak of the CB output $L_{\mathrm{CB}}(\theta) = \mathbf{a}(\theta)^H\hat{\mathbf{R}}\mathbf{a}(\theta)$ computed for a reasonable number of equidistant points (for example, 314 points) to cover the interval $[-\frac{\pi}{2},\frac{\pi}{2}]$. We compute the CB output for the elements of $\boldsymbol{\theta}_{2\mathrm{L}}$ and put those elements that have output larger than η into a new vector $\widetilde{\boldsymbol{\theta}}$. Note that if certain scenarios occur, we consider the L DOAs generated by $|\mathscr{I}| = M$ as the final DOAs and terminate the DOA selection strategy steps. These scenarios are the CB output has less number of peaks than $(L+1)$, and the number of elements of $\widetilde{\boldsymbol{\theta}}$ becomes smaller than L or equal to $2L$.

Step 2: Select the first DOA as that element of $\widetilde{\boldsymbol{\theta}}$ which maximizes the GLR, that is,

$$\hat{\theta}_1 = \arg\max_{\theta} \frac{\mathbf{a}^H(\theta)\hat{\mathbf{Q}}^{-1}\hat{\mathbf{R}}\hat{\mathbf{Q}}^{-1}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\hat{\mathbf{Q}}^{-1}\mathbf{a}(\theta)} \quad , \theta \in \widetilde{\boldsymbol{\theta}}. \tag{3.37}$$

Note that the GLR presented in [115] is extended here to the general noise case where $\hat{\mathbf{Q}}$ is an estimate of the noise covariance matrix.

Step 3: The remaining elements of $\widetilde{\boldsymbol{\theta}}$ are stored in $\bar{\boldsymbol{\theta}}$. Let $\overline{L}$ denote the length of $\bar{\boldsymbol{\theta}}$. Using the elements of $\bar{\boldsymbol{\theta}}$, we construct $\bar{G} = \frac{\bar{L}}{(L-1)!(\bar{L}-L+1)!}$ DOA subsets such that each subset has $(L-1)$ DOAs. Let $\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2, \ldots, \boldsymbol{\Theta}_{\bar{G}}$ and $\mathbf{A}(\boldsymbol{\Theta}_1), \mathbf{A}(\boldsymbol{\Theta}_2), \ldots, \mathbf{A}(\boldsymbol{\Theta}_{\bar{G}})$ be these DOA subsets and their corresponding array manifolds, respectively. Therefore, we identify the $(L-1)$ remaining DOAs as the subset that minimizes the following DML cost function

$$\hat{\boldsymbol{\Theta}}_R = \arg\min_{\boldsymbol{\Theta}_i} \text{trace}\left[\left(\mathbf{P}^{\perp}_{\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)} - \boldsymbol{\nu}_1\boldsymbol{\nu}_1^H\right)\hat{\mathbf{Q}}^{-\frac{1}{2}}\hat{\mathbf{R}}\hat{\mathbf{Q}}^{-\frac{1}{2}}\right], \quad i = 1, 2, \ldots, \bar{G} \tag{3.38}$$

where $\mathbf{P}^{\perp}_{\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)} \triangleq \mathbf{I}_M - \widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)\left(\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)^H\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)\right)^{-1}\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)^H$, $\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i) \triangleq \hat{\mathbf{Q}}^{-\frac{1}{2}}\mathbf{A}(\boldsymbol{\Theta}_i)$, and $\boldsymbol{\nu}_1 \triangleq \frac{\mathbf{P}^{\perp}_{\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)}\hat{\mathbf{Q}}^{-\frac{1}{2}}\mathbf{a}(\hat{\theta}_1)}{\|\mathbf{P}^{\perp}_{\widetilde{\mathbf{A}}(\boldsymbol{\Theta}_i)}\hat{\mathbf{Q}}^{-\frac{1}{2}}\mathbf{a}(\hat{\theta}_1)\|_2}$. Noteworthy to mention that $\hat{\mathbf{Q}}^{-\frac{1}{2}}\hat{\mathbf{R}}\hat{\mathbf{Q}}^{-\frac{1}{2}}$ is employed here to cover the general noise case.

4. MIMO Channel Estimation and Data Detection with One-Bit ADCs

4.1 Signal Model

4.1.1 MmWave UL Channel

Let a BS of a multi-user mmWave MIMO system be composed of a ULA with M antennas that deploy one-bit ADCs. Consider also K single antenna users equipped with high-resolution digital-to-analog converters (DACs). The UL channel between user k and the BS is mathematically expressed as in (1.4). As a result, by placing $\mathbf{h}_k$ for $k = 1,2,\ldots,K$ in columns of the matrix $\mathbf{H}$, we have

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K] = [\mathbf{A}(\boldsymbol{\theta}_1)\boldsymbol{\gamma}_1, \mathbf{A}(\boldsymbol{\theta}_2)\boldsymbol{\gamma}_2, \ldots, \mathbf{A}(\boldsymbol{\theta}_K)\boldsymbol{\gamma}_K] \,. \tag{4.1}$$

Consequently, the received signal at the BS in the training stage is formulated as

$$\mathbf{Y} = \mathcal{Q}(\mathbf{HS} + \mathbf{N}) \tag{4.2}$$

where $\mathcal{Q}(\cdot) \triangleq \text{sign}(\Re\{\cdot\}) + j\,\text{sign}(\Im\{\cdot\})$ represents the one-bit quantizer, $\mathbf{S} \in \mathbb{C}^{K \times N_s}$ is the pilot sequence transmitted by users, and $\mathbf{N} \in \mathbb{C}^{M \times N_s}$ is a matrix of complex-valued Gaussian noise with zero mean and variance σ^2.

4.1.2 OFDM Systems With Frequency Selective Channels

Consider a MIMO-OFDM system operating over a frequency selective channel with known number of channel taps, denoted by L_{tap}. The BS deploys M antennas equipped by one-bit ADCs. This MIMO-OFDM system serves K single-antenna users with high-resolution DACs. Moreover, N_{c} is the number of sub-carriers employed by the MIMO-OFDM system. The frequency domain symbol of the k-th user is $\mathbf{x}_k^{\text{FD}} \in \mathbb{C}^{N_{\text{c}} \times 1}$. We add a cyclic prefix (CP) of length N_{cp} with N_{cp} satisfying the relation $L_{\text{tap}} - 1 \leq N_{\text{cp}} \leq N_{\text{c}}$. Note that the superscripts "TD" and "FD" are used to specify Time Domain and Frequency Domain variables,

respectively. After excluding the CP, the one-bit quantized observed signal by the i-th antenna of the BS can be modeled as

$$\mathbf{y}_i^{\mathrm{TD}} = \mathcal{Q}\left(\sum_{k=1}^{K} \mathbf{G}_{i,k}^{\mathrm{TD}} \boldsymbol{W}_D^H \mathbf{x}_k^{\mathrm{FD}} + \mathbf{n}_i^{\mathrm{TD}}\right),\ i = 1,\ldots,M \tag{4.3}$$

where $\boldsymbol{W}_D \in \mathbb{C}^{N_c \times N_c}$ represents the normalized DFT matrix, and $\mathbf{G}_{i,k}^{\mathrm{TD}}$ is a circulant matrix, specified by its first column $\mathbf{g}_{i,k}^{\mathrm{TD}} = [(\mathbf{h}_{i,k}^{\mathrm{TD}})^T, 0, \ldots, 0]^T$ with $\mathbf{h}_{i,k}^{\mathrm{TD}} \in \mathbb{C}^{L_{\mathrm{tap}} \times 1}$ being the L_{tap} channel vector between the i-th antenna of the BS and the k-th user. It is assumed that the elements of $\mathbf{h}_{i,k}^{\mathrm{TD}}$ are independent and identically distributed (i.i.d.) as $\mathcal{CN}\left(0, \frac{1}{L_{\mathrm{tap}}}\right)$.

4.2 Proposed Methods

4.2.1 SE-TMR

The SE-TMR method is proposed in Publication VI for one-bit mmWave UL channel estimation. It is developed via leveraging the angular domain sparsity of mmWave channel and Toeplitz matrix reconstruction concept. It is observed from (1.4) that $\mathbf{h}_k$ can be still considered sparse in the angular domain, in spite of being made of many paths. As a result, we approximate the k-th column of $\mathbf{H}$ in (4.1) with only L_k path gains and DOAs. Therefore, (4.1) can be approximated as

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K] = [\mathbf{A}(\bar{\boldsymbol{\theta}}_1)\bar{\boldsymbol{\gamma}}_1, \mathbf{A}(\bar{\boldsymbol{\theta}}_2)\bar{\boldsymbol{\gamma}}_2, \ldots, \mathbf{A}(\bar{\boldsymbol{\theta}}_K)\bar{\boldsymbol{\gamma}}_K] \tag{4.4}$$

where $\bar{\boldsymbol{\theta}}_k \triangleq [\bar{\theta}_{k,1}, \bar{\theta}_{k,2}, \ldots, \bar{\theta}_{k,L_k}]^T \in \mathbb{R}^{L_k \times 1}$ and $\bar{\boldsymbol{\gamma}}_k \triangleq [\bar{\gamma}_{k,1}, \bar{\gamma}_{k,2}, \ldots, \bar{\gamma}_{k,L_k}]^T \in \mathbb{C}^{L_k \times 1}$ are respectively the DOAs and path gains of L_k paths, considered for reconstructing $\mathbf{h}_k$. Then, we recast (4.4) as

$$\mathbf{H} = \mathbf{A}\boldsymbol{\Gamma}\bar{\mathbf{G}} = \bar{\mathbf{H}}\bar{\mathbf{G}} \tag{4.5}$$

where

$$\mathbf{A} \triangleq [\mathbf{A}(\bar{\boldsymbol{\theta}}_1), \mathbf{A}(\bar{\boldsymbol{\theta}}_2), \ldots, \mathbf{A}(\bar{\boldsymbol{\theta}}_K)] \in \mathbb{C}^{M \times L} \tag{4.6}$$

$$\boldsymbol{\Gamma} \triangleq \begin{bmatrix} \mathrm{diag}(\bar{\boldsymbol{\gamma}}_1) & & \\ & \ddots & \\ & & \mathrm{diag}(\bar{\boldsymbol{\gamma}}_K) \end{bmatrix} \in \mathbb{C}^{L \times L} \tag{4.7}$$

$$\bar{\mathbf{G}} \triangleq \begin{bmatrix} \mathbf{1}_{L_1} & & & \\ & \mathbf{1}_{L_2} & & \\ & & \ddots & \\ & & & \mathbf{1}_{L_K} \end{bmatrix} \in \mathbb{R}^{L \times K} \tag{4.8}$$

$$\bar{\mathbf{H}} \triangleq \mathbf{A}\boldsymbol{\Gamma} \in \mathbb{C}^{M\times L} \tag{4.9}$$

and $L \triangleq \sum_{k=1}^{K} L_k$. Note that by estimating $\bar{\mathbf{H}}$, $\mathbf{H}$ can be recovered since $\bar{\mathbf{G}}$ is known. As $\mathbf{A}$ and $\boldsymbol{\Gamma}$ are Vandermonde and diagonal matrices, respectively, it can be shown that

$$\bar{\mathbf{H}}\bar{\mathbf{H}}^H = \mathscr{T}(\mathbf{u}) \tag{4.10}$$

with $\mathbf{u} \in \mathbb{C}^{M\times 1}$ and $[\mathbf{u}]_1$ being a real number. Combining (4.10) with the sparsity property of the columns of $\mathbf{H}$ in the angular domain, the following optimization problem can be formulated:

$$\min_{\bar{\mathbf{H}},\mathbf{u},\mathbf{E}^R,\mathbf{E}^I} \|\mathrm{vec}\{\boldsymbol{W}_D\bar{\mathbf{H}}\bar{\mathbf{G}}\}\|_1 + \lambda\left(\sum_{i=1}^{M}\sum_{j=1}^{N_s}([\mathbf{E}^R]_{i,j} + [\mathbf{E}^I]_{i,j})\right) \tag{4.11}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{I}_L & \bar{\mathbf{H}}^H \\ \bar{\mathbf{H}} & \mathscr{T}(\mathbf{u}) \end{bmatrix} \succeq 0$$

$$[\mathbf{u}]_1 = \frac{C}{M}$$

$$\Re\{[\bar{\mathbf{H}}\bar{\mathbf{G}}\mathbf{S}]_{i,j}\}\Re\{[\mathbf{Y}]_{i,j}\} \geqslant -[\mathbf{E}^R]_{i,j},$$

$$i = 1,\ldots,M,\ \ j = 1,\ldots,N_s$$

$$\Im\{[\bar{\mathbf{H}}\bar{\mathbf{G}}\mathbf{S}]_{i,j}\}\Im\{[\mathbf{Y}]_{i,j}\} \geqslant -[\mathbf{E}^I]_{i,j},$$

$$i = 1,\ldots,M,\ \ j = 1,\ldots,N_s$$

$$[\mathbf{E}^R]_{i,j} \geqslant 0,\ i = 1,\ldots,M,\ j = 1,\ldots,N_s$$

$$[\mathbf{E}^I]_{i,j} \geqslant 0,\ i = 1,\ldots,M,\ \ j = 1,\ldots,N_s$$

where $\boldsymbol{W}_D \in \mathbb{C}^{M\times M}$ is the normalized DFT matrix, $\lambda > 0$ is a regularization parameter, the entries of $\mathbf{E}^R \in \mathbb{R}^{M\times N_s}$ and $\mathbf{E}^I \in \mathbb{R}^{M\times N_s}$ are slack variables (see Publication VI for details). Note that the first constraint in (4.11) is imposed to enforce the Toeplitz property presented in (4.10). The optimization problem (4.11) is convex, and it is solved by CVX [121] in Publication VI. After recovering $\mathbf{H}$ using (4.5), the RELAX [122] (which is an one-dimensional harmonic retrieval (HR) method) is used in Publication VI to further refine the estimate of $\mathbf{H}$.

4.2.2 L1-RLR-TMR

The L1-RLR-TMR method is proposed in Publication VII for estimating mmWave UL channels with one-bit ADCs. This method leverages the combination of ℓ_1 regularized logistic regression and Toeplitz matrix reconstruction notions for designing a proper minimization problem. An ADMM-based approach is developed in Publication VII for handling the aforementioned minimization problem.

Plugging (4.5) into (4.2), and then applying the vectorization operator to the resultant, we obtain

$$\mathbf{y} \triangleq \text{vec}\{\mathbf{Y}\} = \mathcal{Q}\Big(\big((\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\big)\bar{\mathbf{h}} + \mathbf{n}\Big) \tag{4.12}$$

where $\bar{\mathbf{h}} \triangleq \text{vec}\{\bar{\mathbf{H}}\}$ and $\mathbf{n} \triangleq \text{vec}\{\mathbf{N}\}$. The real domain representation of (4.12) is given as

$$\mathbf{y}_R \triangleq [\Re\{\mathbf{y}\}^T, \Im\{\mathbf{y}\}^T]^T = \bar{\mathbf{S}}\bar{\mathbf{h}}_R \tag{4.13}$$

where $\bar{\mathbf{H}} \triangleq \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I = [\bar{\mathbf{h}}_1, \bar{\mathbf{h}}_2, \ldots, \bar{\mathbf{h}}_M]^T$, $\bar{\mathbf{h}}_R \triangleq [\text{vec}\{\bar{\mathbf{H}}_R\}^T, \text{vec}\{\bar{\mathbf{H}}_I\}^T]^T$, and

$$\begin{aligned}\bar{\mathbf{S}} &\triangleq \begin{bmatrix} \Re\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} & -\Im\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} \\ \Im\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} & \Re\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} \end{bmatrix} \\ &= [\bar{\mathbf{s}}_1, \bar{\mathbf{s}}_2, \ldots, \bar{\mathbf{s}}_{2MN_s}]^T. \end{aligned} \tag{4.14}$$

The following minimization problem can be formulated for finding $\bar{\mathbf{h}}_R$:

$$\begin{aligned} \min_{\bar{\mathbf{h}}_R, \mathbf{u}} \; & \|\bar{\mathbf{F}}\bar{\mathbf{h}}_R\|_1 + \lambda \sum_{t=1}^{2MN_s} \log\Big(1 + e^{-\kappa[\mathbf{y}_R]_t(\bar{\mathbf{s}}_t^T\bar{\mathbf{h}}_R)}\Big) \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I)^H \\ \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I & \mathcal{T}(\mathbf{u}) \end{bmatrix} \succeq 0 \\ & \|\bar{\mathbf{h}}_m\|_2^2 = c \;, \quad m = 1, \ldots, M \end{aligned} \tag{4.15}$$

where $\bar{\mathbf{F}} \triangleq \begin{bmatrix} \Re\{\bar{\mathbf{G}}^T \otimes \boldsymbol{W}_D\} & -\Im\{\bar{\mathbf{G}}^T \otimes \boldsymbol{W}_D\} \\ \Im\{\bar{\mathbf{G}}^T \otimes \boldsymbol{W}_D\} & \Re\{\bar{\mathbf{G}}^T \otimes \boldsymbol{W}_D\} \end{bmatrix}$, $\lambda > 0$ is a regularization parameter, and $\bar{\mathbf{h}}_m^T$ is the m-th row of $\bar{\mathbf{H}}$. In (4.15), the term $\|\bar{\mathbf{F}}\bar{\mathbf{h}}_R\|_1$ is used for capturing the underlying sparsity of the mmWave channel, while the term $\sum_{t=1}^{2MN_s} \log\Big(1 + e^{-\kappa[\mathbf{y}_R]_t(\bar{\mathbf{s}}_t^T\bar{\mathbf{h}}_R)}\Big)$ is the well-known objective function of the binary logistic regression added for modeling the binary outputs of one-bit ADCs. Moreover, the semi-definite relaxation (SDR) of (4.10) is imposed as a constraint analogous to the SE-TMR method. The optimization problem introduced in (4.15) is non-convex. We propose an ADMM-based solution for it in Publication VII. We use the ADMM technique twice for splitting two sets of variables. The first usage is for taking care of the SDR constraint, while the second one is for taking care of the ℓ_1 norm in the objective function. We call the former one as the outer ADMM, whereas the latter is called the inner ADMM. The scaled augmented Lagrangian of both the outer and inner ADMM, as well as the update rules can be found in Publication VII.

4.2.3 AdaBoost-Based Channel Estimation and Data Detection in One-Bit Massive MIMO

In Publication IX, AdaBoost-based algorithms for MIMO-OFDM channel estimation and data detection are proposed. The main idea is to use GDA classifier/approximate GDA classifier as weak learners in each iteration of an AdaBoost

algorithm. This approach enables us to develop such algorithms that are computationally efficient, specifically in large-scale MIMO-OFDM systems.

Binary Classification via GDA

For a training set containing m training examples with n features $\{\mathbf{x}^{(j)}\}_{j=1,\dots,m}$ and two classes $y^{(j)} \in \{1,-1\}_{j=1,\dots,m}$, GDA assumes that each $\mathbf{x}^{(j)}$ is generated from a normal distribution with the covariance matrix of $\boldsymbol{\Sigma}$ and means of $\boldsymbol{\mu}_{-1}$ and $\boldsymbol{\mu}_1$ depending on the value of $y^{(j)}$. The means and covariance matrix can be estimated using the training examples as

$$\hat{\boldsymbol{\mu}}_{-1} = \frac{\sum_{j=1}^{m} \mathbf{1}\{y^{(j)} = -1\}\mathbf{x}^{(j)}}{\sum_{j=1}^{m} \mathbf{1}\{y^{(j)} = -1\}} \tag{4.16}$$

$$\hat{\boldsymbol{\mu}}_{1} = \frac{\sum_{j=1}^{m} \mathbf{1}\{y^{(j)} = 1\}\mathbf{x}^{(j)}}{\sum_{j=1}^{m} \mathbf{1}\{y^{(j)} = 1\}} \tag{4.17}$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{m}\sum_{j=1}^{m}(\mathbf{x}^{(j)} - \hat{\boldsymbol{\mu}}_{y^{(j)}})(\mathbf{x}^{(j)} - \hat{\boldsymbol{\mu}}_{y^{(j)}})^T. \tag{4.18}$$

The decision boundary that separates two classes is then obtained as

$$\mathbf{h}_{\mathrm{GDA}} = \hat{\boldsymbol{\Sigma}}^{-1}\left(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_{-1}\right). \tag{4.19}$$

Channel Estimation

Based on the definitions and details given in Publication IX, the GDA-based weak classifier employed in the t-th iteration of the proposed AdaBoost-based channel estimator can be developed as

$$\hat{\boldsymbol{\mu}}_{-1}^{(t)} = \sum_{j=1}^{2N_{\mathrm{c}}} \mathbf{1}\{y_{i,\mathrm{R},j}^{\mathrm{TD}} = -1\} w_j^{(t)} \boldsymbol{\phi}_{\mathrm{R},j}^{\mathrm{TD}} \tag{4.20}$$

$$\hat{\boldsymbol{\mu}}_{1}^{(t)} = \sum_{j=1}^{2N_{\mathrm{c}}} \mathbf{1}\{y_{i,\mathrm{R},j}^{\mathrm{TD}} = 1\} w_j^{(t)} \boldsymbol{\phi}_{\mathrm{R},j}^{\mathrm{TD}} \tag{4.21}$$

$$\hat{\boldsymbol{\Sigma}}^{(t)} = \sum_{j=1}^{2N_{\mathrm{c}}} w_j^{(t)} (\boldsymbol{\phi}_{\mathrm{R},j}^{\mathrm{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,\mathrm{R},j}^{\mathrm{TD}}}^{(t)})(\boldsymbol{\phi}_{\mathrm{R},j}^{\mathrm{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,\mathrm{R},j}^{\mathrm{TD}}}^{(t)})^T \tag{4.22}$$

$$\hat{\mathbf{h}}_{i,\mathrm{R}}^{\mathrm{TD},(t)} = \left(\hat{\boldsymbol{\Sigma}}^{(t)}\right)^{-1}\left(\hat{\boldsymbol{\mu}}_1^{(t)} - \hat{\boldsymbol{\mu}}_{-1}^{(t)}\right) \tag{4.23}$$

where $\mathbf{h}_{i,\mathrm{R}}^{\mathrm{TD},(t)} \triangleq \left[\Re\{\mathbf{h}_i^{\mathrm{TD}}\}^T, \Im\{\mathbf{h}_i^{\mathrm{TD}}\}^T\right]^T \in \mathbb{R}^{2KL_{\mathrm{tap}}\times 1}$ and $\mathbf{h}_i^{\mathrm{TD}} \triangleq [(\mathbf{h}_{i,1}^{\mathrm{TD}})^T, (\mathbf{h}_{i,2}^{\mathrm{TD}})^T, \dots, (\mathbf{h}_{i,K}^{\mathrm{TD}})^T]^T$ for $i = 1,\dots,M$ where M is the number of

antenna at the BS (see Publication IX for more details). Moreover, $w_j^{(t)}$ is the weight allocated to the j-th training example in the t-th iteration. We name the AdaBoost-based channel estimator which employ (4.23) in its t-th iteration as one-bit GDA-Ada estimator.

Calculating the covariance matrix estimate via (4.22) and then inverting it in (4.23) makes one-bit GDA-Ada computationally inefficient, particularly in large-scale systems. To remedy this issue, two approximate versions of (4.23) can be considered as follows

$$\hat{\mathbf{h}}_{i,\mathrm{R,app1}}^{\mathrm{TD},(t)} \triangleq \left(\hat{\boldsymbol{\Sigma}}_1^{(t)}\right)^{-1}\left(\hat{\boldsymbol{\mu}}_1^{(t)} - \hat{\boldsymbol{\mu}}_{-1}^{(t)}\right) \tag{4.24}$$

$$\hat{\mathbf{h}}_{i,\mathrm{R,app2}}^{\mathrm{TD},(t)} \triangleq \hat{\boldsymbol{\mu}}_1^{(t)} - \hat{\boldsymbol{\mu}}_{-1}^{(t)} \tag{4.25}$$

where $\hat{\boldsymbol{\Sigma}}_1^{(t)} \triangleq \mathrm{diag}\left\{\hat{\boldsymbol{\sigma}}_1^{(t)}\right\}$ and $\hat{\boldsymbol{\sigma}}_1^{(t)} = \sum_{j=1}^{2N_\mathrm{c}} w_j^{(t)}\left((\boldsymbol{\phi}_{\mathrm{R},j}^{\mathrm{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,\mathrm{R},j}^{\mathrm{TD}}}^{(t)}) \odot (\boldsymbol{\phi}_{\mathrm{R},j}^{\mathrm{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,\mathrm{R},j}^{\mathrm{TD}}}^{(t)})\right)$. We call the AdaBoost-based channel estimators which use (4.24) and (4.25) in their t-th iteration as one-bit GDA-Ada-1 and one-bit GDA-Ada-2, respectively. Noteworthy to mention that the computational complexity for implementing the one-bit GDA-Ada-1 and one-bit GDA-Ada-2 estimator is much lower than that of the one-bit GDA-Ada estimator.

Data Detection

Analogous to the channel estimation part, GDA classifier/approximate GDA classifiers can be considered as weak learners in each iteration of AdaBoost-based data detectors. Hence, the t-th weak learner corresponding to the one-bit GDA-Ada, one-bit GDA-Ada-1, and one-bit GDA-Ada-2 data detector can be respectively expressed as

$$\hat{\mathbf{x}}_{\mathrm{R}}^{\mathrm{FD},(t)} = \left(\hat{\boldsymbol{\Sigma}}_{\mathrm{d}}^{(t)}\right)^{-1}\left(\hat{\boldsymbol{\mu}}_{\mathrm{d},1}^{(t)} - \hat{\boldsymbol{\mu}}_{\mathrm{d},-1}^{(t)}\right) \tag{4.26}$$

$$\hat{\mathbf{x}}_{\mathrm{R,app1}}^{\mathrm{FD},(t)} = \left(\hat{\boldsymbol{\Sigma}}_{\mathrm{d},1}^{(t)}\right)^{-1}\left(\hat{\boldsymbol{\mu}}_{\mathrm{d},1}^{(t)} - \hat{\boldsymbol{\mu}}_{\mathrm{d},-1}^{(t)}\right) \tag{4.27}$$

$$\hat{\mathbf{x}}_{\mathrm{R,app2}}^{\mathrm{FD},(t)} = \hat{\boldsymbol{\mu}}_{\mathrm{d},1}^{(t)} - \hat{\boldsymbol{\mu}}_{\mathrm{d},-1}^{(t)} \tag{4.28}$$

where

$$\hat{\boldsymbol{\mu}}_{\mathrm{d},-1}^{(t)} = \sum_{j=1}^{2MN_\mathrm{c}} \mathbf{1}\{y_{\mathrm{R},j}^{\mathrm{TD}} = -1\} w_j^{(t)} \mathbf{g}_{\mathrm{R},j}^{\mathrm{FD}} \tag{4.29}$$

$$\hat{\boldsymbol{\mu}}_{\mathrm{d},1}^{(t)} = \sum_{j=1}^{2MN_\mathrm{c}} \mathbf{1}\{y_{\mathrm{R},j}^{\mathrm{TD}} = 1\} w_j^{(t)} \mathbf{g}_{\mathrm{R},j}^{\mathrm{FD}} \tag{4.30}$$

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{d}}^{(t)} = \sum_{j=1}^{2MN_\mathrm{c}} w_j^{(t)} (\mathbf{g}_{\mathrm{R},j}^{\mathrm{FD}} - \hat{\boldsymbol{\mu}}_{\mathrm{d},y_{\mathrm{R},j}^{\mathrm{TD}}}^{(t)})(\mathbf{g}_{\mathrm{R},j}^{\mathrm{FD}} - \hat{\boldsymbol{\mu}}_{\mathrm{d},y_{\mathrm{R},j}^{\mathrm{TD}}}^{(t)})^T \tag{4.31}$$

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{d},1}^{(t)} = \mathrm{diag}\left\{\hat{\boldsymbol{\sigma}}_{\mathrm{d},1}^{(t)}\right\} \tag{4.32}$$

$$\hat{\boldsymbol{\sigma}}_{\mathrm{d},1}^{(t)} = \sum_{j=1}^{2MN_{\mathrm{c}}} w_j^{(t)} \left((\mathbf{g}_{\mathrm{R},j}^{\mathrm{FD}} - \hat{\boldsymbol{\mu}}_{\mathrm{d},y_{\mathrm{R},j}^{\mathrm{TD}}}^{(t)}) \odot (\mathbf{g}_{\mathrm{R},j}^{\mathrm{FD}} - \hat{\boldsymbol{\mu}}_{\mathrm{d},y_{\mathrm{R},j}^{\mathrm{TD}}}^{(t)}) \right). \tag{4.33}$$

and other definitions and details can be found in Publication IX.

4.3 Experimental Results

In this section, in addition to simulation examples in Publication VII, the performance of the L1-RLR-TMR and SE-TMR methods in estimating mmWave UL channels is compared with that of the BLMMSE [84] and AR [83] methods. The pilot sequence is selected as a circularly shifted replica of a Zadoff-Chu (ZC) sequence of length N_s where each row is orthogonal to the others, i.e., $\mathbf{S}\mathbf{S}^H = N_s\mathbf{I}_{\mathrm{K}}$. The SNR and normalized mean squared error (NMSE) are respectively defined as $\mathrm{SNR} \triangleq 10\log_{10}\left(\frac{\|\mathbf{H}\mathbf{S}\|_F^2}{MN_s\sigma^2}\right)$ and $\mathrm{NMSE} \triangleq \frac{1}{KN}\sum_{k=1}^{K}\sum_{n=1}^{N}\left\|\frac{\hat{\mathbf{h}}_k^{(n)}}{\|\hat{\mathbf{h}}_k^{(n)}\|_2} - \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|_2}\right\|_2^2$, where $\hat{\mathbf{h}}_k^{(n)}$ denotes the kth column of $\hat{\mathbf{H}}$ estimated in the n-th Monte Carlo run with $\mathbf{h}_k$ being the actual kth column of $\mathbf{H}$, and N being the total number of Monte Carlo trials considered as $N = 200$. We consider $\lambda = 1$ for the SE-TMR and L1-RLR-TMR methods, and $K = 8$. The number of channel clusters and the number of the within cluster multipaths for all users are considered to be the same. The latter is set as $M_{\mathrm{path}}^{1,1} = \cdots = M_{\mathrm{path}}^{1,L_1} = \cdots = M_{\mathrm{path}}^{K,1} = \cdots = M_{\mathrm{path}}^{1,L_K} = 100$. We generate DOAs randomly once and use them for all Monte Carlo trials. The channel path gains are distributed as $\mathscr{CN}(0,1)$. Fig. 4.1 compares the NMSE of the methods tested for the scenario when $M = 16$, $N_s = 128$, $L_k = 1$ for all users, and the angle spread of 8 degrees within each cluster. It can be seen from Fig. 4.1 that the performance of BLMMSE degrades substantially when the precise estimate of the channel covariance matrix is not available. Moreover, the performance of L1-RLR-TMR is comparable to that of the SE-TMR method at high-SNR regime, although the SE-TMR method is implemented using CVX and has high complexity [121]. In Fig. 4.2, the performance of the methods tested is shown for the setup of $M = 16$, $N_s = 128$, $L_k = 2$, and the within cluster angle spreads are 8 and 10 degrees for all users. The efficiency of L1-RLR-TMR is confirmed at high-SNR regime as compared to other methods tested. Particularly, Fig. 4.2 shows that the performance of L1-RLR-TMR implemented by the ADMM is comparable with that of the SE-TMR implemented using CVX.

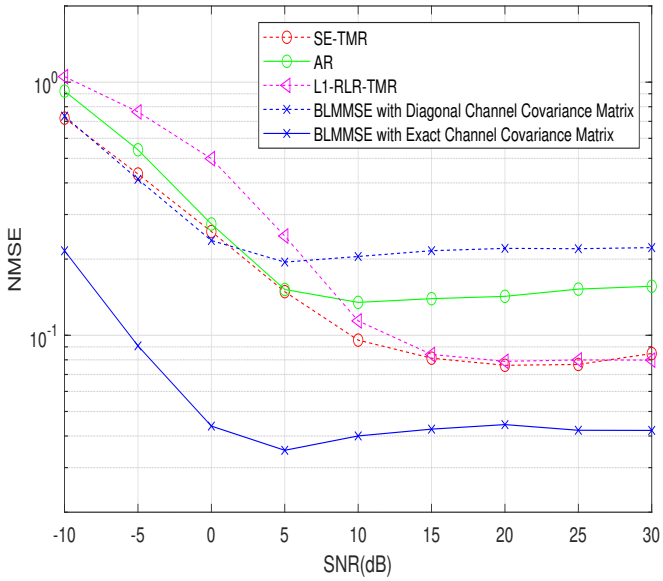

Figure 4.1. NMSE vs. SNR for $M = 16$, $N_s = 128$, and $L_k = 1$.

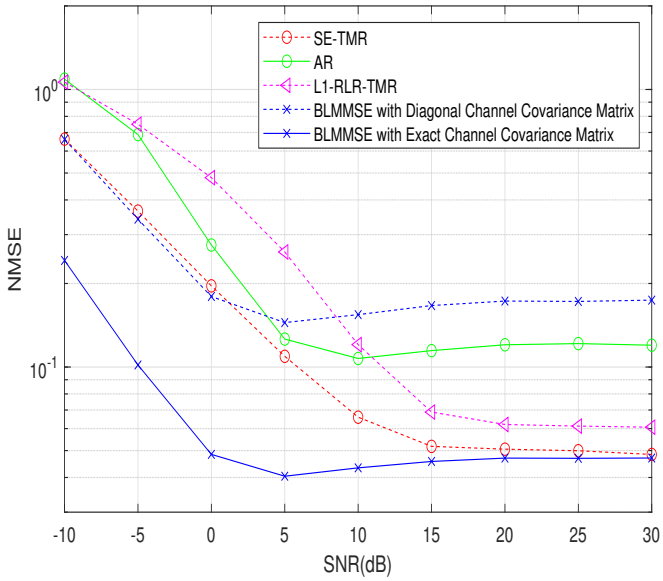

Figure 4.2. NMSE vs. SNR for $M = 16$, $N_s = 128$, and $L_k = 2$.

5. Summary and Future Directions

In this thesis, computationally efficient and yet accurate algorithms have been developed for some problems in the area of spectral analysis and its applications. Specifically, the problems of noisy AR parameter estimation, DOA estimation in the presence of unknown noise fields, and one-bit massive/mmWave MIMO channel estimation and data detection have been studied.

Five methods have been developed for noisy AR parameter estimation. The main idea of the first method is to reduce the detrimental impact of noise variance in each iteration, whereas a constrained LS optimization problem has been formulated to estimate the AR parameters in the second method. The third one uses an approximation to reduce the dimension of any arbitrary noisy AR problem to only two unknown parameters and then estimates those two parameters in an iterative manner. The fourth method estimates the observation noise variance as the minimum eigenvalue of an enlarged autocorrelation matrix. The fifth one solves a properly designed generalized eigenvalue problem to first estimate the observation noise variance, and then estimate the AR coefficients.

For the case of uniform sensor noise, two ESPRIT-based DOA estimation methods called ES ESPRIT and EU ESPRIT have been developed which use GLS-based algorithms to first generate a candidate set of DOAs and then pick up the final DOAs by either a DML-based or a GLR-based DOA selection strategies. Furthermore, a computationally efficient non-iterative method called NISB have been proposed for DOA estimation in the presence of nonuniform noise. The NISB method is composed of two phases where an initial estimate of the noise subspace is obtained in the first phase and the nonuniform noise covariance matrix as well as a refined estimate of the noise subspace are obtained in the second phase. A unified approch that contains three steps has been also developed for DOA estimation in the case of unknown sensor noise. The aim of the first step is to estimate the nonuniform or block-diagonal noise covariace matrix, while the second step has been devised for generating DOA candidates using rooting-based forward-only or FB GLS-based algorithms. The third step exploits the CB, GLR, and DML concepts to select the best final DOA estimates. This approach outperforms state-of-the-art DOA estimation methods in coping with challenging setups such as small sample size and low SNRs.

The SE-TMR and L1-RLR-TMR methods have been developed for one-bit mmWave UL channel estimation. The SE-TMR method solves a convex optimization problem that enforces the underlying sparsity and Toeplitz structure of the channel by considering the ℓ_1 norm of the channel in the DFT domain and a positive semi-definite (PSD) constraint, respectively. The aforementioned optimization problem is solved via CVX. The L1-RLR-TMR method formulates a non-convex optimization problem using the combination of ℓ_1 logistic regression and Toeplitz matrix reconstruction. A computationally efficient ADMM-based algorithm has been presented to solve the optimization problem of L1-RLR-TMR. Lastly, we have considered GDA/approximate GDA classification methods as weak learners in iterations of AdaBoost-based algorithms to develop computationally efficient channel estimators and data detectors in MIMO-OFDM systems with one-bit ADCs. It has been assumed that the fading of channels is the frequency selective fading. The proposed AdaBoost-based channel estimators and data detectors, which employ approximate versions of GDA as weak classifiers, require substantially lower computational complexity compared to other existing methods.

5.1 Future Directions

To close the loop of development in this thesis, it would be interesting to develop a one-bit DOA estimation method based on AdaBoost. Such one-bit DOA estimation method would be computationally very efficient because of using approximate GDA classifiers as weak classifiers.

Some other research directions and extensions of the development in this thesis are: (i) developing few-bit extensions of the proposed one-bit channel estimators and data detectors. (ii) developing computationally efficient algorithms for few-bit DOA estimation problem.

References

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

[2] S. M. Kay, *Modern Spectral Estimation*. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[3] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1993.

[4] R. A. Horn and C. A. Johnson, *Matrix Analysis*. Cambridge university press, Upper Saddle River, Cambridge, 1985.

[5] M. D. Srinath, P. K. Rajasekaran, and R. Viswanathan, *Introduction to Statistical Signal Processing with Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1996.

[6] R. Sameni, "Online filtering using piecewise smoothness priors: application to normal and abnormal electrocardiogram denoising," *Signal Process.*, vol. 133, pp. 52–63, 2017.

[7] L. Legrand and E. Grivel, "Jeffrey's divergence between autoregressive processes distributed by additive white noise," *Signal Process.*, vol. 149, pp. 162–178, 2018.

[8] S. Dutta, M. Singh, and A. Kumar, "Automated classification of non-motor mental task in electroencephalogram based brain-computer interface using multivariate autoregressive model in the intrinsic mode function domain," *Biomedical Signal Process. & Control*, vol. 43, pp. 174–182, 2018.

[9] S. Ganapathy, "Multirate autoregressive spectrogram modeling for noisy speech recognition," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1373–1377, 2017.

[10] N. Tolimieri, E. E. Holmes, G. D. Williams, R. Pacunski, and D. Lowry, "Population assessment using multivariate time-series analysis: A case

study of rockfishes in Puget Sound," *Ecology and Evolution*, vol. 7, no. 8, pp. 2846–2860, 2017.

[11] O. Terzi and G. Ergin, "Forecasting of monthly river flow with autoregressive modeling and data-driven techniques," *Neural Computing and Application*, vol. 25, no. 1, pp. 179–188, 2014.

[12] J. J. Musa, "Stochastic modeling of Shiroro river stream flow process," *American Journal of Eng. Research* vol. 2, no. 6, pp. 49–54, 2013.

[13] M. Marcellino, J. H. Stock, and M. W. Watson, "A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series," *Journal of Econometrics*, vol. 135, no. 1, pp. 499–526, 2006.

[14] R. Weron and A. Misiorek, "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models," *Int. Journal of Forecasting*, vol. 24, no. 4, pp. 744–763, 2008.

[15] A. K. Lohani, R. Kumar, and R. D. Singh, "Hydrological time series modeling: A comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques," *Journal of Hydrology*, vol. 442, pp. 23–35, 2012.

[16] S. Feng, W. Ren, M. Han, and Y. W. Chen, "Robust manifold broad learning system for large-scale noisy chaotic time series prediction: A perturbation perspective," *Neural networks*, vol. 117, pp. 179–190, 2019.

[17] S. Maanan, B. Dumitrescu, and C. D. Giurcaneanu, "Conditional independence graphs for multivariate autoregressive models by convex optimization: Efficient algorithms," *Signal Process.*, vol. 133, pp. 122–134, 2017.

[18] M. Kallas, P. Honeine, C. Francis, and H. Amond, "Kernel autoregressive models using Yule-Walker equations," *Multidim. Signal Process.*, vol. 93, no. 11, pp. 3053–3061, 2013.

[19] S. M. Kay, "Noise compensation for autoregressive spectral estimates," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 3, pp. 292–303, 1980.

[20] W. X. Zheng , "On estimation of autoregressive signals in the presence of noise," *IEEE Trans. Circt. Syst. II*, vol. 53, no. 12, pp. 1471–1475, 2006.

[21] Y. Xia and W. X. Zheng , "Novel parameter estimation of autoregressive signals in the presence of noise," *Automatica*, vol. 62, pp. 98–105, 2015.

[22] W. X. Zheng , "Autoregressive parameter estimation from noisy data," *IEEE Trans. Circt. Syst. II*, vol. 47, no. 1, pp. 71–75, 2000.

[23] W. X. Zheng , “Fast identification of autoregressive signals from noisy observations,” *IEEE Trans. Circt. Syst. II*, vol. 52, no. 1, pp. 43–48, 2005.

[24] C. E. Davila , “A subspace approach to estimation of autoregressive parameters from noisy measurements,” *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 531–534, 1988.

[25] S.A. Vorobyov, A.B. Gershman, and K.M. Wong, “Maximum likelihood direction of arrival estimation in unknown noise fields using sparse sensor arrays,” *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 34–43, Jan. 2005.

[26] H.L. Van Trees, *Optimum Array Processing. Part IV: Detection, Estimation, and Modulation Theory.* John Wiley & Sons, 2004.

[27] P.-J. Chung, M. Viberg, and J. Yu, “DOA estimation methods and algorithms,” in *Academic Press Library in Signal Processing.* Elsevier, 2014, vol. 3, pp. 599–650.

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

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

[30] J.M. Kim, O.K. Lee, and J.C. Ye, “Compressive MUSIC: Revisiting the link between compressive sensing and array signal processing,” *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 278–301, Jan. 2012.

[31] A. Barabell, “Improving the resolution performance of eigenstructure-based direction-finding algorithms,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 8, pp. 336–339, 1983.

[32] M. Pesavento, A.B. Gershman, and M. Haardt, “Unitary root-MUSIC with a real-valued eigendecomposition: A theoretical and experimental performance study,” *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1306–1314, May 2000.

[33] C. Qian, L. Huang, and H.-C. So, “Improved unitary root-MUSIC for DOA estimation based on pseudo-noise resampling,” *IEEE Signal Process. Lett.*, vol. 21, no. 2, pp. 140–144, Feb. 2013.

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

[35] 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, May 1995.

[36] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[37] P. Stoica and K.C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1132–1143, Jul. 1990.

[38] P. Stoica and A. Nehorai, "Performance comparison of subspace rotation and MUSIC methods for direction estimation," *IEEE Trans. Signal Process.*, vol. 39, no. 2, pp. 446–453, Feb. 1991.

[39] M. Shaghaghi and S.A. Vorobyov, "Subspace leakage analysis and improved DOA estimation with small sample size," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3251–3265, Jun. 2015.

[40] C. Qian, L. Huang, N.D. Sidiropoulos, and H. C. So, "Enhanced PUMA for direction-of-arrival estimation and its performance analysis," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4127–4137, Aug. 2016.

[41] J. Steinwandt, F. Roemer, and M. Haardt, "Generalized least squares for ESPRIT-type direction of arrival estimation," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1681–1685, Nov. 2017.

[42] M. Trinh-Hoang, M. Viberg, and M. Pesavento, "Partial relaxation approach: An eigenvalue-based DOA estimator framework," *IEEE Trans. Signal Process.*, vol. 66, no. 23, pp. 6190–6203, Dec. 2018.

[43] M. Trinh-Hoang, M. Viberg and M. Pesavento, "Cramer-Rao bound for DOA estimators under the partial relaxation framework: Derivation and comparison," *IEEE Trans. Signal Process.*, vol. 68, pp. 3194–3208, May 2020.

[44] M. W. Morency, S.A. Vorobyov, and G. Leus, "Joint detection and localization of an unknown number of sources using the algebraic structure of the noise subspace," *IEEE Trans. Signal Process.*, vol. 66, no. 17, pp. 4685–4700, Sep. 2018.

[45] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[46] M. Esfandiari and S.A. Vorobyov, "Enhanced standard ESPRIT for overcoming imperfections in DOA estimation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Toronto, Canada*, Jun. 2021, pp. 4375–4379.

[47] D.A. Linebarger, R.D. DeGroat, and E.M. Dowling, "Efficient direction-finding methods employing forward/backward averaging," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2136–2145, Aug. 1994.

[48] S.U. Pillai and B.H. Kwon, “Forward/backward spatial smoothing techniques for coherent signal identification,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 1, pp. 8–15, Jan. 1989.

[49] M. Bilodeau and D. Brenner, *Theory of Multivariate Statistics*. Springer Science & Business Media, 2008.

[50] B. Ottersten, P. Stoica, and R. Roy, “Covariance matching estimation techniques for array signal processing applications,” *Digit. Signal Process.*, vol. 8, no. 3, pp. 185–210, Jul. 1998.

[51] K.-C. Huarng and C.-C. Yeh, “A unitary transformation method for angle-of-arrival estimation,” *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 975–977, Apr. 1991.

[52] A.B. Gershman and P. Stoica, “On unitary and forward–backward MODE,” *Digit. Signal Process.*, vol. 9, no. 2, pp. 67–75, Apr. 1999.

[53] B.D. Rao and K.V.S. Hari, “Weighted subspace methods and spatial smoothing: Analysis and comparison,” *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 788–803, Feb. 1993.

[54] 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, Jul. 2001.

[55] C.E. Chen, F. Lorenzelli, R.E. Hudson, and K. Yao, “Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise,” *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3038–3044, Jul. 2008.

[56] D. Madurasinghe, “A new DOA estimator in nonuniform noise,” *IEEE Signal Process. Lett.*, vol. 12, no. 4, pp. 337–339, Apr. 2005.

[57] B. Liao, S.-C. Chan, L. Huang, and C. Guo, “Iterative methods for subspace and DOA estimation in nonuniform noise,” *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3008–3020, Jun. 2016.

[58] B. Liao, L. Huang, C. Guo, and H.C. So, “New approaches to direction-of-arrival estimation with sensor arrays in unknown nonuniform noise,” *IEEE J. Sensors*, vol. 16, no. 24, pp. 8982–8989, Dec. 2016.

[59] J. Wen, B. Liao, and C. Guo, “Spatial smoothing based methods for direction-of-arrival estimation of coherent signals in nonuniform noise,” *Digit. Signal Process.*, vol. 67, pp. 116–122, Aug. 2017.

[60] C. Qi, Z. Chen, Y. Wang, and Y. Zhang, “DOA estimation for coherent sources in unknown nonuniform noise fields,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 3, pp. 1195–1204, Jul. 2007.

[61] 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, Feb. 2011.

[62] P. Stoica and P. Babu, "SPICE and LIKES: Two hyperparameter-free methods for sparse-parameter estimation," *Signal Process.*, vol. 92, no. 7, pp. 1580–1590, Jul. 2012.

[63] P. Stoica, D. Zachariah, and J. Li, "Weighted SPICE: A unifying approach for hyperparameter-free sparse estimation," *Digit. Signal Process.*, vol. 33, pp. 1–12, Oct. 2014.

[64] Z.-Q. He, Z.-P. Shi, and L. Huang, "Covariance sparsity-aware DOA estimation for nonuniform noise," *Digit. Signal Process.*, vol. 28, pp. 75–81, May 2014.

[65] H. Wang, X. Wang, L. Wan, and M. Huang, "Robust sparse Bayesian learning for off-grid DOA estimation with non-uniform noise," *IEEE Access*, vol. 6, pp. 64688–64697, Nov. 2018.

[66] M. Esfandiari, S.A. Vorobyov, S. Alibani, and M. Karimi, "Non-iterative subspace-based DOA estimation in the presence of nonuniform noise," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 848–852, Jun. 2019.

[67] J. LeCadre, "Parametric methods for spatial signal processing in the presence of unknown colored noise field," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 965–983, Jul. 1989.

[68] H. Ye and R.D. DeGroat, "Maximum likelihood DOA estimation and asymptotic Cramér-Rao bounds for additive unknown colored noise," *IEEE Trans. Signal Process.*, vol. 43, no. 4, pp. 938–949, Apr. 1995.

[69] B. Friedlander and A.J. Weiss, "Direction finding using noise covariance modeling," *IEEE Trans. Signal Process.*, vol. 43, no. 7, pp. 1557–1567, Jul. 1995.

[70] B. Göransson and B. Ottersten, "Direction estimation in partially unknown noise fields," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2375–2385, Sep. 1999.

[71] M. Agrawal and S. Prasad, "A modified likelihood function approach to DOA estimation in the presence of unknown spatially correlated Gaussian noise using a uniform linear array," *IEEE Tran. Signal Process.*, vol. 48, no. 10, pp. 2743–2749, Oct. 2000.

[72] A.B. Gershman, P. Stoica, M. Pesavento, and E. Larsson, "The stochastic Cramer–Rao bound for direction estimation in unknown noise fields," *IEE Radar, Sonar, Navigat.*, vol. 149, no. 1, pp. 2–8, Feb. 2002.

[73] "5G channel model for bands up to 100 GHz (v2.3)," Tech. Rep., 2016. [Online]. Available: http://www.5gworkshops.com/5gcm.html

[74] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[75] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[76] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[77] B. Murmann, "Energy limits in A/D converters," *Proc. IEEE Faible Tension Faible Consomation*, 2013, pp. 1–4.

[78] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath, "Uplink performance of wideband massive MIMO with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Jan. 2017.

[79] H. Kim and J. Choi, "Channel estimation for spatially/temporally correlated massive MIMO systems with one-bit ADCs," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, Dec. 2019, Art. no. 267.

[80] J. Liu, Z. Luo, and X. Xiong, "Low-resolution ADCs for wireless communication: A comprehensive survey," *IEEE Access*, vol. 7, pp. 91291–91324, 2019.

[81] J. Choi, J. Mo, and R. W. Heath, "Near maximum-likelihood detector and channel estimator for for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, May 2016.

[82] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, Jun. 2016.

[83] C. Qian, X. Fu, and N. D. Sidiropoulos, "Amplitude retrieval for channel estimation of MIMO systems with one-bit ADCs," *IEEE Signal Process. Lett.*, vol. 26, no. 11, pp. 1698–1702, Nov. 2019.

[84] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[85] Q. Wan, J. Fang, H. Duan, Z. Chen, and H. Li, "Generalized bussgang LMMSE channel estimation for one-bit massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4234–4246, Jun. 2020.

[86] C.-K. Wen, C.-J. Wang, S. Jin, K.-K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, May 2016.

[87] P. Sun, Z. Wang, R. W. Heather, and P. Schniter, "Joint channel-estimation/decoding with frequency-selective channels and few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 899–914, Feb. 2019.

[88] X. Cheng, B. Xia, K. Xu, and S. Li, "Bayesian channel estimation and data detection in oversampled OFDM receiver with low-resolution ADC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5558–5571, Sep. 2021.

[89] L. V. Nguyen, A. L. Swindlehurst, and D. H. N. Nguyen, "SVM-based channel estimation and data detection for one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2086–2099, Mar. 2021.

[90] S. Gao, X. Cheng, L. Fang, and L. Yang, "Model enhanced learning based detectors (Me-LeaD) for wideband multi-user 1-bit mmWave communications," *IEEE Trans. Wireless Commun.*, vol. 20, pp. 4646–4656, Jul. 2021.

[91] D. H. N. Nguyen, "Neural network-optimized channel estimator and training signal design for MIMO systems with few-bit ADCs," *IEEE Signal Process. Lett.*, vol. 27, pp. 1370–1374, Jul. 2020.

[92] Y. Zhang, M. Alrabeiah, and A. Alkhateeb, "Deep learning for massive MIMO with 1-bit ADCs: When more antennas need fewer pilots," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1273–1277, Aug. 2020.

[93] Y. Dong, H. Wang, and Y. D. Yao, "Channel estimation for one-bit multiuser massive MIMO using conditional GAN," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 854–858, Mar. 2021.

[94] Y. Jeon, S. Hong, and N. Lee, "Supervised-learning-aided communication framework for MIMO systems with low-resolution ADCs," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7299–7313, Aug. 2018.

[95] L. V. Nguyen, D. T. Ngo, N. H. Tran, A. L. Swindlehurst, and D. H. N. Nguyen, "Supervised and semi-supervised learning for MIMO blind detection with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2427–2442, Apr. 2019.

[96] Y.-S. Jeon, N. Lee, and H. V. Poor, "Robust data detection for MIMO systems with one-bit ADCs: A reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1663–1676, Mar. 2020.

[97] C. Rusu, R. Mendez-Rial, N. Gonzalez-Prelcic, and R. W. Heath, "Adaptive one-bit compressive sensing with application to low-precision receivers

at mmWave," in *Proc. IEEE Glob. Commun. Conf.*, San Diego, CA, USA, Dec. 2015, pp. 1–6.

[98] J. Mo, P. Schniter, and R. W. Heath, "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, Mar. 2018.

[99] H. Kim and J. Choi, "Channel AoA estimation for massive MIMO systems using one-bit ADCs," *J. Commun. Netw.*, vol. 20, no. 4, pp. 374–382, Aug. 2018

[100] S. Rao, A. Mezghani, and A. L. Swindlehurst, "Channel estimation in one-bit massive MIMO systems: Angular versus unstructured models," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1017–1031, Sep. 2019.

[101] F. Liu, H. Zhu, C. Li, J. Li, P. Wang, and P. Orlik, "Angular-domain channel estimation for one-bit massive MIMO systems: Performance bounds and algorithms," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2928–2942, Mar. 2020.

[102] L. Xu, C. Qian, F. Gao, W. Zhang, and S. Ma, "Angular domain channel estimation for mmWave massive MIMO with one-bit ADCs/DACs," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 969–982, Feb. 2021.

[103] M. Esfandiari, S. A. Vorobyov, and R. W. Heath Jr., "Sparsity enforcing with Toeplitz matrix reconstruction method for mmWave UL channel estimation with one-bit ADCs," in *Proc. IEEE 12th Sensor Array multichan. Signal Process. Workshop*, Trondheim, Norway, Jun. 2022, pp. 141–145.

[104] M. Esfandiari, S.A. Vorobyov, and R. W. Heath Jr., "ADMM-based solution for mmWave UL channel estimation with one-bit ADCs via sparsity enforcing and Toeplitz matrix reconstruction," in *Proc. 57th IEEE Int. Conf. Communications, IEEE ICC'23, Rome, Italy, May 28–June 1, 2023.*

[105] Y. Jeon, N. Lee, S. Hong, and R. W. Heath, "One-bit sphere decoding for uplink massive MIMO systems with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4509–4521, Jul. 2018.

[106] A. Hjørungnes, *Complex-Valued Matrix Werivatives With Applications in Signal Processing and Communications.* Cambridge University Press, 2011.

[107] S. Provencher, "Parameters estimation of complex multitone signal in the DFT domain," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3001–3012, Jul. 2011.

[108] Y. Tan, K. Wang, L. Wang, and H. Wen, "Efficient FFT based multi source DOA estimation for ULA," *Signal Process.*, vol. 189, Dec. 2021, Art. no. 108284.

[109] T. Amemiya, “Generalized least squares theory,” *Advanced Econometrics*, Cambridge, MA, USA: Harvard Univ. Press, 1985.

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

[111] A.B. Gershman, “Pseudo-randomly generated estimator banks: A new tool for improving the threshold performance of direction finding,” *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1351–1364, May 1998.

[112] V. Vasylyshyn, “Improved beamspace ESPRIT-based DOA estimation via pseudo-noise resampling,” in *Proc. IEEE 9th Europ. Radar Conf.*, Amsterdam, Netherlands, 2012, pp. 238–241.

[113] V. Vasylyshyn, “Removing the outliers in root-MUSIC via pseudo-noise resampling and conventional beamformer,” *Signal Process.*, vol. 93, no. 12, pp. 3423–3429, Dec. 2013.

[114] A.B. Gershman and P. Stoica, “New MODE-based techniques for direction finding with an improved threshold performance,” *Signal Process.*, vol. 76, no. 3, pp. 221–235, Aug. 1999.

[115] F. Izedi, M. Karimi, and M. Derakhtian, “Joint DOA estimation and source number detection for arrays with arbitrary geometry,” *Signal Process.*, vol. 140, pp. 149–160, Nov. 2017.

[116] K. Petersen and M. Pedersen, *The Matrix Cookbook*, Technical Univ. Denmark, Kongens Lyngby, Denmark, Tech. Rep., vol. 3274, 2012.

[117] [online]: `https://sgfin.github.io/files/notes/CS229_Lecture_Notes.pdf`

[118] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: Data mining, inference, and prediction*. Springer, 2009.

[119] T. Hastie, R. Saharon, Z. Ji, and Z. Hui, “Multi-class AdaBoost,” *Statistics and its Interface*, vol. 2, no. 3, pp. 349–360, 2009.

[120] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via alternating direction method of multipliers,” *Found. Trends. Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[121] M. Grant, S. Boyd, and Y. Ye, CVX: Matlab software for disciplined convex programming. 2008 [Online]. Available: http://stanford.edu/~boyd/cvx

[122] J. Li and P. Stoica, “Efficient mixed-spectrum estimation with applications to target feature extraction,” *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 281–295, Feb. 1996.

Errata

Publication VI

The definition of SNR should be revised as $\mathrm{SNR} = 10\log_{10}\left(\frac{\|\mathbf{HS}\|_{\mathrm{F}}^2}{MN_s\sigma^2}\right)$. Therefore, the curves of both Figures should be shifted to the right for 22 dBs.

Aalto-DT 142/2023

ISBN 978-952-64-1420-1 (printed)
ISBN 978-952-64-1421-8 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Electrical Engineering
Department of Information and Communications Engineering
www.aalto.fi

BUSINESS + ECONOMY

ART + DESIGN + ARCHITECTURE

SCIENCE + TECHNOLOGY

CROSSOVER

DOCTORAL THESES