

Advances and New Applications of Spectral Analysis

Majdoddin Esfandiari

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

© 2023 Majdoddin Esfandiari

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



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

Preface	i
Contents	iii
List of Publications	v
Author's Contribution	vii
Abbreviations	ix
Symbols	xi
1. Introduction	1
1.1 Objectives	3
1.2 Contributions	3
1.3 Thesis Structure	4
2. Noisy Autoregressive (AR) Parameter Estimation	5
2.1 Signal Model	5
2.2 Proposed Methods	6
2.2.1 The first proposed method	6
2.2.2 The second proposed method	7
2.2.3 The third proposed method	8
2.2.4 The fourth proposed method	9
2.2.5 The fifth proposed method	10
2.3 Experimental Results	11
3. DOA Estimation in the Presence of Uniform, Nonuniform, and Block-diagonal Sensor Noise	13
3.1 Signal Model	13
3.2 Proposed Methods	14
3.2.1 ES ESPRIT and EU ESPRIT	14
3.2.2 NISB	16

3.2.3	Unified Approach to DOA Estimation in Unknown Noise Fields	17
4.	MIMO Channel Estimation and Data Detection with One-Bit ADCs	23
4.1	Signal Model	23
4.1.1	MmWave UL Channel	23
4.1.2	OFDM Systems With Frequency Selective Channels . .	23
4.2	Proposed Methods	24
4.2.1	SE-TMR	24
4.2.2	L1-RLR-TMR	25
4.2.3	AdaBoost-Based Channel Estimation and Data Detection in One-Bit Massive MIMO	26
4.3	Experimental Results	29
5.	Summary and Future Directions	31
5.1	Future Directions	32
	References	33
	Errata	43
	Publications	45

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

Abbreviations

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

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

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

\mathbb{R} Real numbers field

\mathbf{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

$\|\cdot\|_F$ Frobenius norm of a matrix

\odot Hadamard product

\otimes Kronecker product

\forall For all

Symbols

$\text{bdiag}\{\cdot\}$ Block-diagonal matrix with the bracketed matrices being the main block-diagonal

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

$\text{DFT}\{\cdot\}$ DFT of the bracketed argument

$\text{diag}\{\cdot\}$ Diagonal matrix with the bracketed vector being the main diagonal

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

$\log(\cdot)$ Natural logarithm operator

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

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

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

$\text{rem}(a, b)$ Remainder in the division of a by b

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

$\text{sign}\{\cdot\}$ Element-wise sign operator

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

$\text{trace}\{\cdot\}$ Trace of a square matrix

$\text{unvec}\{\cdot\}$ Unvectorization operator

$\text{vec}\{\cdot\}$ Vectorization operator

$\mathbf{\Pi} \succeq 0$ $\mathbf{\Pi}$ 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

$$x(t) = a_1x(t-1) + a_2x(t-2) + \dots + a_px(t-p) + e(t) = \mathbf{a}^T \mathbf{x}_t + e(t) \quad (1.1)$$

$$y(t) = x(t) + w(t) \quad (1.2)$$

where $e(t)$ denotes the zero mean white driving noise with variance of σ_e^2 , $\mathbf{a} = [a_1, a_2, \dots, a_p]^T$ contains coefficients of the AR model, $\mathbf{x}_t = [x(t-1), x(t-2), \dots, x(t-p)]^T$, 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. Therefore, 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

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

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

$[1, e^{-j2\pi\sin(\theta_1)d/\lambda}, \dots, e^{-j2\pi(M-1)\sin(\theta_1)d/\lambda}]^T \in \mathbb{C}^M$ being the steering vector corresponding to DOA θ_l for $l = 1, \dots, L$, $\mathbf{s}(t) \triangleq [s_1(t), s_2(t), \dots, s_L(t)]^T \in \mathbb{C}^L$ 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

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

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, $\mathbf{a}(\theta_{k,l,m}) \triangleq [1, e^{-j\pi\sin(\theta_{k,l,m})}, \dots, e^{-j(M-1)\pi\sin(\theta_{k,l,m})}]^T \in \mathbb{C}^{M \times 1}$, $\boldsymbol{\theta}_{k,l} \triangleq [\theta_{k,l,1}, \theta_{k,l,2}, \dots, \theta_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{R}^{M_{\text{path}}^{k,l} \times 1}$ for $l = 1, 2, \dots, L_k$, $\mathbf{A}(\theta_{k,l}) \triangleq [\mathbf{a}(\theta_{k,l,1}), \mathbf{a}(\theta_{k,l,2}), \dots, \mathbf{a}(\theta_{k,l,M_{\text{path}}^{k,l})}] \in \mathbb{C}^{M \times M_{\text{path}}^{k,l}}$, $\boldsymbol{\gamma}_{k,l} \triangleq [\gamma_{k,l,1}, \gamma_{k,l,2}, \dots, \gamma_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{C}^{M_{\text{path}}^{k,l} \times 1}$, $\boldsymbol{\theta}_k \triangleq [\boldsymbol{\theta}_{k,1}^T, \boldsymbol{\theta}_{k,2}^T, \dots, \boldsymbol{\theta}_{k,L_k}^T]^T$, $\mathbf{A}(\theta_k) \triangleq [\mathbf{A}(\theta_{k,1}), \mathbf{A}(\theta_{k,2}), \dots, \mathbf{A}(\theta_{k,L_k})]$, and $\boldsymbol{\gamma}_k \triangleq [\boldsymbol{\gamma}_{k,1}^T, \boldsymbol{\gamma}_{k,2}^T, \dots, \boldsymbol{\gamma}_{k,L_k}^T]^T$. 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

$$r_y(k) = r_x(k) + \sigma_w^2 \delta(k) \quad (2.1)$$

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

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

where $\mathbf{r}_x = [r_x(1), r_x(2), \dots, r_x(p)]^T$. 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 $1 \leq k \leq p$ and $p+1 \leq k \leq p+q$, where

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

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

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

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

$$\mathbf{R}_y \mathbf{a} - \sigma_w^2 \mathbf{a} = \mathbf{r}_y \quad (2.7)$$

$$\mathbf{R}'_y \mathbf{a} = \mathbf{r}'_y \quad (2.8)$$

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 $\mathbf{a}_{\text{LS}} = \mathbf{R}_y^{-1} \mathbf{r}_y$ 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 $\mathbf{c} = \mathbf{a} - \Delta$ denote the estimate of \mathbf{a} obtained in the previous iteration with $\|\Delta\|_2 \ll \|\mathbf{a}\|_2$. 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,

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

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

$$\mathbf{b}_i^T \mathbf{R}_y \mathbf{a} = \mathbf{b}_i^T \mathbf{r}_y + \sigma_w^2 \mathbf{b}_i^T \mathbf{a}, \quad i = 1, \dots, p-1. \quad (2.10)$$

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

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

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 $\mathbf{H}\mathbf{a} = \mathbf{h}$ with the following definitions:

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

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}_{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

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

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

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

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

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

$$\sigma_w^2 = \frac{\mathbf{a}^\top \mathbf{R}_y \mathbf{a} - \mathbf{a}^\top \mathbf{r}_y}{\|\mathbf{a}\|_2^2} \quad (2.16)$$

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

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

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

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

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 $m = 1, \dots, p$, we write

$$\mathbf{a} = \sum_{m=1}^p \alpha_m \mathbf{v}_m. \quad (2.19)$$

Note that we represent the eigenvalues of \mathbf{R}_y by λ_m for $m = 1, \dots, p$ where $\lambda_1 < \lambda_2 < \dots < \lambda_p$. 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 $m = 1, \dots, p$), we have

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

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

$$\alpha_m \lambda_m = \mathbf{v}_m^T \mathbf{r}_y + \sigma_w^2 \alpha_m \implies \alpha_m = \frac{\mathbf{v}_m^T \mathbf{r}_y}{\lambda_m - \sigma_w^2}, \quad m = 1, \dots, p. \quad (2.21)$$

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

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

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

$$\mathbf{H}_2 \boldsymbol{\alpha} = \mathbf{h}_2 \quad (2.23)$$

where

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

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

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

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

where

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

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

$$\mathbf{A}\mathbf{a} = \mathbf{b} + \sigma_w^2 \mathbf{c} \quad (2.27)$$

where

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

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

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

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:

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

where

$$\begin{aligned} \mathbf{G}_0 &= \mathbf{A}^T \mathbf{D}^T \mathbf{D} \mathbf{A}, \quad \mathbf{G}_1 = -(\mathbf{A}^T \mathbf{D}^T \mathbf{E} + \mathbf{E}^T \mathbf{D} \mathbf{A}), \\ \mathbf{G}_2 &= \mathbf{E}^T \mathbf{E}. \end{aligned} \quad (2.31)$$

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

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

where

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

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.

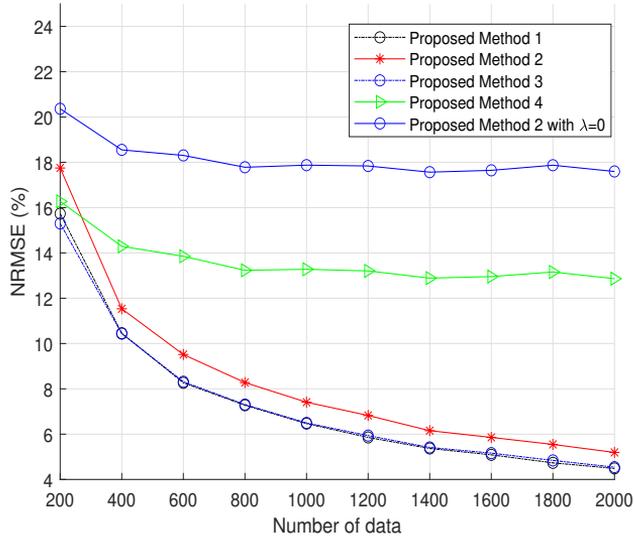


Figure 2.1. NRMSE vs. the number of data points for the first example.

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
$\sigma_a^2 = 4.6$	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
<i>NRMSE</i> (%)	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

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

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

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

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 $\mathbf{Q}_{\text{uni}} = \sigma^2 \mathbf{I}_M$, $\mathbf{Q}_{\text{nonuni}} = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2\}$, and $\mathbf{Q}_{\text{bdiag}} = \text{bdiag}\{\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_q\}$. In the latter, it is worth noting that $\mathbf{Q}_j \in \mathbb{C}^{n_j \times n_j}$ for $j = 1, \dots, q$.

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

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (3.3)$$

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

¹Note 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

$$\mathbf{J}_1 \mathbf{U}_s \Psi = \mathbf{J}_2 \mathbf{U}_s. \quad (3.4)$$

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 $\Psi \in \mathbb{C}^{L \times L}$ is a matrix whose eigenvalues λ_l 's are related to θ_l 's through $\lambda_l = e^{-j\pi \sin(\theta_l)}$ for $l = 1, \dots, L$. As a result, the aim is to first estimate Ψ , 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 $\mathbf{Z}_{\mathcal{J}} \in \mathbb{R}^{|\mathcal{J}| \times M-1}$, we have

$$\mathbf{Z}_{\mathcal{J}} \mathbf{W}_D \mathbf{J}_1 \mathbf{U}_s \Psi = \mathbf{Z}_{\mathcal{J}} \mathbf{W}_D \mathbf{J}_2 \mathbf{U}_s \quad (3.5)$$

where members of the set \mathcal{J} are the indices of the selected equations. Note that all entries of the i -th row of $\mathbf{Z}_{\mathcal{J}}$ are zero except one entry which is set to 1. The index of this nonzero entry is specified by the i -th member of \mathcal{J} . 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 $\mathbf{U}_s = \hat{\mathbf{U}}_s + \Delta \mathbf{U}_s$ into (3.5) and reorganizing the terms, we obtain

$$\mathbf{Z}_{\mathcal{J}} \mathbf{W}_D \mathbf{J}_1 \hat{\mathbf{U}}_s \Psi + \mathbf{E} = \mathbf{Z}_{\mathcal{J}} \mathbf{W}_D \mathbf{J}_2 \hat{\mathbf{U}}_s \quad (3.6)$$

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

$$\hat{\mathbf{f}}_S \triangleq \text{vec} \{ \mathbf{Z}_{\mathcal{J}} \mathbf{W}_D \mathbf{J}_2 \hat{\mathbf{U}}_s \} = \hat{\mathbf{F}}_S \boldsymbol{\psi} + \mathbf{e} = \hat{\mathbf{F}}_S \boldsymbol{\psi} + \hat{\mathbf{G}}_S \Delta \mathbf{u}_s \quad (3.7)$$

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

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

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

where as shown in [41] $\hat{\mathbf{W}}_S = \left[\hat{\mathbf{G}}_S \left(\hat{\boldsymbol{\Sigma}}_s^{-2} \otimes \mathbf{I}_M \right) \hat{\mathbf{G}}_S^H \right]^{-1}$ with $\hat{\boldsymbol{\Sigma}}_s \in \mathbb{R}^{L \times L}$ 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 $\hat{\boldsymbol{\Psi}}_{GLS} = \text{unvec} \{ \hat{\boldsymbol{\psi}}_{GLS} \}$. Due to the dependency of $\hat{\mathbf{G}}_S$ on $\boldsymbol{\Psi}$, it is natural to use an iterative method to estimate $\hat{\mathbf{W}}_S$ and $\hat{\boldsymbol{\Psi}}_{GLS}$. Based on our observations, few iterations are sufficient to reach accurate results.

The unitary extension of (3.5) is given as

$$\mathbf{Z}_{\mathcal{J}} \mathbf{W}_D \mathbf{K}_1 \mathbf{E}_s \boldsymbol{\Upsilon} = \mathbf{Z}_{\mathcal{J}} \mathbf{W}_D \mathbf{K}_2 \mathbf{E}_s \quad (3.9)$$

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 $\varphi(\bar{\mathbf{X}}) = \mathbf{Q}_M^H \bar{\mathbf{X}} \mathbf{Q}_{2N} \in \mathbb{R}^{M \times 2N}$ with \mathbf{Q}_M and \mathbf{Q}_{2N} being left Π -real matrices [35]. Similar to the steps presented for solving (3.5), a GLS-based solution of (3.9) is obtained as

$$\hat{\mathbf{v}}_{GLS} = \left(\hat{\mathbf{F}}_U^H \hat{\mathbf{W}}_U \hat{\mathbf{F}}_U \right)^{-1} \hat{\mathbf{F}}_U^H \hat{\mathbf{W}}_U \hat{\mathbf{f}}_U \quad (3.10)$$

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

In Publication IV, we propose to select members of \mathcal{J} in (3.5) as those indices associated with $|\mathcal{J}|$ 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 $|\mathcal{J}| = M - 1$ and $|\mathcal{J}| = M - 2$, 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:

$$\hat{\boldsymbol{\Theta}}_{DML} = \underset{\boldsymbol{\Theta}_i}{\text{argmin}} \text{trace} \left((\mathbf{I}_M - \mathbf{A}(\boldsymbol{\Theta}_i) (\mathbf{A}(\boldsymbol{\Theta}_i)^H \mathbf{A}(\boldsymbol{\Theta}_i))^{-1} \mathbf{A}(\boldsymbol{\Theta}_i)^H) \hat{\mathbf{R}} \right) \quad (3.11)$$

$$\forall \quad i = 1, \dots, P_{ESE(EUE)}$$

where $\boldsymbol{\Theta}_i$ 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 $l = 1, \dots, L$) DOA estimate as the member of the set of $2L - (l - 1)$ remaining DOAs that maximizes the GLR cost function [115], that is,

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

where

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

with $\mathbf{A}_{l-1} \triangleq [\mathbf{a}(\hat{\theta}_1), \mathbf{a}(\hat{\theta}_2), \dots, \mathbf{a}(\hat{\theta}_{l-1})] \in \mathbb{C}^{M \times (l-1)}$. 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:

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

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

$$\hat{\mathbf{R}} \hat{\mathbf{U}} \approx \hat{\mathbf{Q}}_{\text{nonuni}} \hat{\mathbf{U}}. \quad (3.15)$$

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 $\hat{\mathbf{R}}_{\text{RCM}} = \hat{\mathbf{R}} - \mathcal{D}(\hat{\mathbf{R}})$. We can write $\hat{\mathbf{Q}}_{\text{nonuni}}$ as

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

where σ^2 denotes the uniform part of sensor noise variances, and \mathbf{Q}_{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

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

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

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

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

$$\hat{\mathbf{Q}}_{\text{nun}} = \text{diag} \{ [\hat{\mathbf{R}}]_{1,1} - c, \dots, [\hat{\mathbf{R}}]_{M,M} - c \} \quad (3.19)$$

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 $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}_{\text{nonuni}}^{(i)}\}$. 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):

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

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]

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

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

$$\hat{\sigma}_m^{2(i+1)} = \mathbf{d}_m^T \left(\frac{1}{2} \mathcal{D} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}} \right\} \mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \right\}^{-1} \right) \mathbf{d}_m, \quad (3.22)$$

$$m = 1, 2, \dots, M.$$

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 $\hat{\mathbf{Q}}_{\text{nonuni}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$ 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

$$\hat{\mathbf{R}}\hat{\mathbf{U}} \approx \hat{\mathbf{Q}}_{\text{bdiag}}\hat{\mathbf{U}}. \quad (3.23)$$

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 $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}_{\text{bdiag}}^{(i)}\}$.

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

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

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]

$$\frac{\partial f_{\text{bdiag}}(\mathbf{Q}_{\text{bdiag}})}{\partial \mathbf{Q}_j} = \mathbf{Q}_j^* (\bar{\mathbf{V}}_{jj}^{(i)})^T + (\bar{\mathbf{V}}_{jj}^{(i)})^T \mathbf{Q}_j^T - (\bar{\mathbf{R}}_{jj}^{(i)})^T, \quad j = 1, \dots, q \quad (3.25)$$

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

$$\mathbf{V}_j^{(i)} \mathbf{q}_j = \bar{\mathbf{r}}_j^{(i)}, \quad j = 1, \dots, q \quad (3.26)$$

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

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

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

$$\bar{\mathbf{X}} \triangleq \mathbf{Q}^{-\frac{1}{2}} \mathbf{X} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A} \mathbf{S} + \mathbf{Q}^{-\frac{1}{2}} \mathbf{N} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A} \mathbf{S} + \bar{\mathbf{N}} \quad (3.28)$$

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

$$\bar{\mathbf{X}} = \mathbf{U}_s \boldsymbol{\Sigma}_s \mathbf{V}_s^H \quad (3.29)$$

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

$$\tilde{\mathbf{u}}_p + \text{diag}(\tilde{\mathbf{u}}_p) \mathbf{W}_a \mathbf{a} = \bar{\mathbf{W}} \mathbf{b}_p, \quad p = 1, \dots, L \quad (3.30)$$

where $\tilde{\mathbf{u}}_p \triangleq \text{DFT}\{\tilde{\mathbf{u}}_p\} = \mathbf{W}_D \tilde{\mathbf{u}}_p$, $\tilde{\mathbf{u}}_p$ is the p -th column of $\tilde{\mathbf{U}}_s$, $\mathbf{W}_a \triangleq [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M]^T \in \mathbb{C}^{M \times L}$, $\bar{\mathbf{W}} \triangleq [\bar{\mathbf{w}}_1, \bar{\mathbf{w}}_2, \dots, \bar{\mathbf{w}}_M]^T \in \mathbb{C}^{M \times L}$, $\mathbf{w}_k \triangleq [W_M^k, (W_M^k)^2, (W_M^k)^3, \dots, (W_M^k)^L]^T$, $\bar{\mathbf{w}}_k \triangleq [1, W_M^k, (W_M^k)^2, \dots, (W_M^k)^{L-1}]^T$, 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 $\gamma_l = e^{-j 2\pi d \sin(\theta_l)/\lambda}$ for $l = 1, \dots, L$ [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 $\mathbf{Z}_{\mathcal{S}} \in \mathbb{R}^{|\mathcal{S}| \times M}$, introduced earlier in this chapter, is used to revise (3.30) as

$$\mathbf{Z}_{\mathcal{S}} \tilde{\mathbf{u}}_p + \text{diag}(\mathbf{Z}_{\mathcal{S}} \tilde{\mathbf{u}}_p) \mathbf{Z}_{\mathcal{S}} \mathbf{W}_a \mathbf{a} = \mathbf{Z}_{\mathcal{S}} \bar{\mathbf{W}} \mathbf{b}_p, \quad p = 1, \dots, L. \quad (3.31)$$

Note that different selecting matrices choose different sets of equations in (3.31). It is shown in Publication VIII that choosing $|\mathcal{S}| = M$ and $|\mathcal{S}| = M - 1$ 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, $\mathbf{B} \in \mathbb{C}^{|\mathcal{S}| \times (|\mathcal{S}| - L)}$ is obtained such that $\mathbf{B}^H \tilde{\mathbf{Z}}_{\mathcal{S}} = \mathbf{0}_{(|\mathcal{S}| - L) \times L}$ with $\tilde{\mathbf{Z}}_{\mathcal{S}} \triangleq \mathbf{Z}_{\mathcal{S}} \bar{\mathbf{W}} \in \mathbb{C}^{|\mathcal{S}| \times L}$. Then, multiplying (3.31) by \mathbf{B}^H results in

$$\mathbf{B}^H (\mathbf{Z}_{\mathcal{S}} \tilde{\mathbf{u}}_p + \text{diag}(\mathbf{Z}_{\mathcal{S}} \tilde{\mathbf{u}}_p) \mathbf{Z}_{\mathcal{S}} \mathbf{W}_a \mathbf{a}) = \mathbf{0}_{(|\mathcal{S}| - L)}, \quad p = 1, \dots, L. \quad (3.32)$$

Reorganizing (3.32), we form $\mathbf{H}_p \mathbf{a} = \mathbf{h}_p$ for $p = 1, \dots, L$, where $\mathbf{H}_p \triangleq \mathbf{B}^H \text{diag}(\mathbf{Z}_{\mathcal{S}} \tilde{\mathbf{u}}_p) \mathbf{Z}_{\mathcal{S}} \mathbf{W}_a \in \mathbb{C}^{(|\mathcal{S}| - L) \times L}$ and $\mathbf{h}_p \triangleq -\mathbf{B}^H \mathbf{Z}_{\mathcal{S}} \tilde{\mathbf{u}}_p \in \mathbb{C}^{(|\mathcal{S}| - L)}$. 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

$$\mathbf{H} \mathbf{a} = \mathbf{h} \quad (3.33)$$

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

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

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

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

$$\hat{\mathbf{W}} = (\hat{\Sigma}_s^2 \otimes (\mathbf{C}(\mathbf{a})\mathbf{C}^H(\mathbf{a}))^{-1}), \quad (3.35)$$

where $\hat{\Sigma}_s$ is a diagonal matrix which contains the L largest singular values of the matrix $\hat{\mathbf{X}} \triangleq \hat{\mathbf{Q}}^{-\frac{1}{2}}\mathbf{X}$, and $\mathbf{C}(\mathbf{a}) \triangleq \mathbf{B}^H (\mathbf{I}_{|\mathcal{S}|} + \text{diag}\{\mathbf{Z}_{|\mathcal{S}}\mathbf{W}_\alpha\mathbf{a}\})\mathbf{Z}_{|\mathcal{S}}\mathbf{W}_D\hat{\mathbf{Q}}^{\frac{1}{2}} \in \mathbb{C}^{(|\mathcal{S}|-L) \times M}$. 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

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

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

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 $l = 1, \dots, L$ 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 $\boldsymbol{\theta}_{2L}$ 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 $L_{\text{CB}}(\theta) = \mathbf{a}(\theta)^H \hat{\mathbf{R}}\mathbf{a}(\theta)$ 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 $\boldsymbol{\theta}_{2L}$ and put those elements that have output larger than η into a new vector $\tilde{\boldsymbol{\theta}}$. Note that if certain scenarios occur, we consider the L DOAs generated by $|\mathcal{S}| = M$ 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 $\tilde{\boldsymbol{\theta}}$ becomes smaller than L or equal to $2L$.

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

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

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 $\tilde{\boldsymbol{\theta}}$ are stored in $\tilde{\boldsymbol{\theta}}$. Let \bar{L} denote the length of $\tilde{\boldsymbol{\theta}}$. Using the elements of $\tilde{\boldsymbol{\theta}}$, we construct $\tilde{G} = \frac{\bar{L}}{(L-1)!(L-L+1)!}$ DOA subsets such that each subset has $(L-1)$ DOAs. Let $\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2, \dots, \boldsymbol{\Theta}_{\tilde{G}}$ and $\mathbf{A}(\boldsymbol{\Theta}_1), \mathbf{A}(\boldsymbol{\Theta}_2), \dots, \mathbf{A}(\boldsymbol{\Theta}_{\tilde{G}})$ 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

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

where $\mathbf{P}_{\mathbf{A}(\boldsymbol{\Theta}_i)}^\perp \triangleq \mathbf{I}_M - \tilde{\mathbf{A}}(\boldsymbol{\Theta}_i) \left(\tilde{\mathbf{A}}(\boldsymbol{\Theta}_i)^H \tilde{\mathbf{A}}(\boldsymbol{\Theta}_i) \right)^{-1} \tilde{\mathbf{A}}(\boldsymbol{\Theta}_i)^H$, $\tilde{\mathbf{A}}(\boldsymbol{\Theta}_i) \triangleq \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A}(\boldsymbol{\Theta}_i)$, and $\mathbf{v}_1 \triangleq \frac{\mathbf{P}_{\mathbf{A}(\boldsymbol{\Theta}_i)}^\perp \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{a}(\hat{\theta}_1)}{\| \mathbf{P}_{\mathbf{A}(\boldsymbol{\Theta}_i)}^\perp \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{a}(\hat{\theta}_1) \|_2}$. 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 $k = 1, 2, \dots, K$ in columns of the matrix \mathbf{H} , we have

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K] = [\mathbf{A}(\theta_1)\gamma_1, \mathbf{A}(\theta_2)\gamma_2, \dots, \mathbf{A}(\theta_K)\gamma_K]. \quad (4.1)$$

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

$$\mathbf{Y} = \mathcal{Q}(\mathbf{H}\mathbf{S} + \mathbf{N}) \quad (4.2)$$

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 subcarriers employed by the MIMO-OFDM system. The frequency domain symbol of the k -th user is $\mathbf{x}_k^{\text{FD}} \in \mathbb{C}^{N_c \times 1}$. We add a cyclic prefix (CP) of length N_{cp} with N_{cp} satisfying the relation $L_{\text{tap}} - 1 \leq N_{\text{cp}} \leq N_c$. 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

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

where $\mathbf{W}_D \in \mathbb{C}^{N_c \times N_c}$ represents the normalized DFT matrix, and $\mathbf{G}_{i,k}^{\text{TD}}$ is a circulant matrix, specified by its first column $\mathbf{g}_{i,k}^{\text{TD}} = [(\mathbf{h}_{i,k}^{\text{TD}})^T, 0, \dots, 0]^T$ with $\mathbf{h}_{i,k}^{\text{TD}} \in \mathbb{C}^{L_{\text{tap}} \times 1}$ 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 $\mathbf{h}_{i,k}^{\text{TD}}$ are independent and identically distributed (i.i.d.) as $\mathcal{CN} \left(0, \frac{1}{L_{\text{tap}}} \right)$.

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

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

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

$$\mathbf{H} = \mathbf{A} \mathbf{\Gamma} \bar{\mathbf{G}} = \bar{\mathbf{H}} \bar{\mathbf{G}} \quad (4.5)$$

where

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

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

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

$$\tilde{\mathbf{H}} \triangleq \mathbf{A}\mathbf{\Gamma} \in \mathbb{C}^{M \times L} \quad (4.9)$$

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

$$\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H = \mathcal{T}(\mathbf{u}) \quad (4.10)$$

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:

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

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

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

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

$$i = 1, \dots, M, \quad j = 1, \dots, N_s$$

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

$$i = 1, \dots, M, \quad j = 1, \dots, N_s$$

$$[\mathbf{E}^R]_{i,j} \geq 0, \quad i = 1, \dots, M, \quad j = 1, \dots, N_s$$

$$[\mathbf{E}^I]_{i,j} \geq 0, \quad i = 1, \dots, M, \quad j = 1, \dots, N_s$$

where $\mathbf{W}_D \in \mathbb{C}^{M \times M}$ 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

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

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

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

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

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

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

$$\begin{aligned} \min_{\bar{\mathbf{h}}_R, \mathbf{u}} & \|\bar{\mathbf{F}}\bar{\mathbf{h}}_R\|_1 + \lambda \sum_{t=1}^{2MN_s} \log \left(1 + e^{-\kappa |y_{R,t}| (\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)} \right) \\ \text{s.t.} & \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I)^H \\ \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I & \mathcal{T}(\mathbf{u}) \end{bmatrix} \geq 0 \\ & \|\bar{\mathbf{h}}_m\|_2^2 = c, \quad m = 1, \dots, M \end{aligned} \quad (4.15)$$

where $\bar{\mathbf{F}} \triangleq \begin{bmatrix} \Re\{\bar{\mathbf{G}}^T \otimes \mathbf{W}_D\} & -\Im\{\bar{\mathbf{G}}^T \otimes \mathbf{W}_D\} \\ \Im\{\bar{\mathbf{G}}^T \otimes \mathbf{W}_D\} & \Re\{\bar{\mathbf{G}}^T \otimes \mathbf{W}_D\} \end{bmatrix}$, $\lambda > 0$ 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 $\sum_{t=1}^{2MN_s} \log \left(1 + e^{-\kappa |y_{R,t}| (\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)} \right)$ 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 Σ 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

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

$$\hat{\boldsymbol{\mu}}_1 = \frac{\sum_{j=1}^m \mathbf{1}\{y^{(j)} = 1\} \mathbf{x}^{(j)}}{\sum_{j=1}^m \mathbf{1}\{y^{(j)} = 1\}} \quad (4.17)$$

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

The decision boundary that separates two classes is then obtained as

$$\mathbf{h}_{\text{GDA}} = \hat{\Sigma}^{-1} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_{-1}). \quad (4.19)$$

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

$$\hat{\boldsymbol{\mu}}_{-1}^{(t)} = \sum_{j=1}^{2N_c} \mathbf{1}\{y_{i,\text{R},j}^{\text{TD}} = -1\} w_j^{(t)} \boldsymbol{\phi}_{\text{R},j}^{\text{TD}} \quad (4.20)$$

$$\hat{\boldsymbol{\mu}}_1^{(t)} = \sum_{j=1}^{2N_c} \mathbf{1}\{y_{i,\text{R},j}^{\text{TD}} = 1\} w_j^{(t)} \boldsymbol{\phi}_{\text{R},j}^{\text{TD}} \quad (4.21)$$

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

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

where $\mathbf{h}_{i,\text{R}}^{\text{TD},(t)} \triangleq [\Re\{\mathbf{h}_i^{\text{TD}}\}^T, \Im\{\mathbf{h}_i^{\text{TD}}\}^T]^T \in \mathbb{R}^{2KL_{\text{tap}} \times 1}$ and $\mathbf{h}_i^{\text{TD}} \triangleq [(\mathbf{h}_{i,1}^{\text{TD}})^T, (\mathbf{h}_{i,2}^{\text{TD}})^T, \dots, (\mathbf{h}_{i,K}^{\text{TD}})^T]^T$ 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

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

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

where $\hat{\boldsymbol{\Sigma}}_1^{(t)} \triangleq \text{diag} \{ \hat{\boldsymbol{\sigma}}_1^{(t)} \}$ and $\hat{\boldsymbol{\sigma}}_1^{(t)} = \sum_{j=1}^{2N_c} w_j^{(t)} \left((\boldsymbol{\phi}_{R,j}^{\text{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,R,j}^{\text{TD}}}^{(t)}) \odot (\boldsymbol{\phi}_{R,j}^{\text{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,R,j}^{\text{TD}}}^{(t)}) \right)$. 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

$$\hat{\mathbf{x}}_R^{\text{FD},(t)} = \left(\hat{\boldsymbol{\Sigma}}_d^{(t)} \right)^{-1} \left(\hat{\boldsymbol{\mu}}_{d,1}^{(t)} - \hat{\boldsymbol{\mu}}_{d,-1}^{(t)} \right) \quad (4.26)$$

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

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

where

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

$$\hat{\boldsymbol{\mu}}_{d,1}^{(t)} = \sum_{j=1}^{2MN_c} \mathbf{1}\{y_{R,j}^{\text{TD}} = 1\} w_j^{(t)} \mathbf{g}_{R,j}^{\text{FD}} \quad (4.30)$$

$$\hat{\boldsymbol{\Sigma}}_d^{(t)} = \sum_{j=1}^{2MN_c} w_j^{(t)} (\mathbf{g}_{R,j}^{\text{FD}} - \hat{\boldsymbol{\mu}}_{d,y_{R,j}^{\text{TD}}}^{(t)}) (\mathbf{g}_{R,j}^{\text{FD}} - \hat{\boldsymbol{\mu}}_{d,y_{R,j}^{\text{TD}}}^{(t)})^T \quad (4.31)$$

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

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

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., $\mathbf{S}\mathbf{S}^H = N_s \mathbf{I}_K$. The SNR and normalized mean squared error (NMSE) are respectively defined as $\text{SNR} \triangleq 10 \log_{10} \left(\frac{\|\mathbf{H}\mathbf{S}\|_F^2}{MN_s\sigma^2} \right)$ and $\text{NMSE} \triangleq \frac{1}{KN} \sum_{k=1}^K \sum_{n=1}^N \left\| \frac{\hat{\mathbf{h}}_k^{(n)}}{\|\hat{\mathbf{h}}_k^{(n)}\|_2} - \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|_2} \right\|_2^2$, where $\hat{\mathbf{h}}_k^{(n)}$ denotes the k th column of $\hat{\mathbf{H}}$ estimated in the n -th Monte Carlo run with \mathbf{h}_k being the actual k th 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 $M_{\text{path}}^{1,1} = \dots = M_{\text{path}}^{1,L_1} = \dots = M_{\text{path}}^{K,1} = \dots = M_{\text{path}}^{K,L_K} = 100$. We generate DOAs randomly once and use them for all Monte Carlo trials. The channel path gains are distributed as $\mathcal{CN}(0, 1)$. 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.

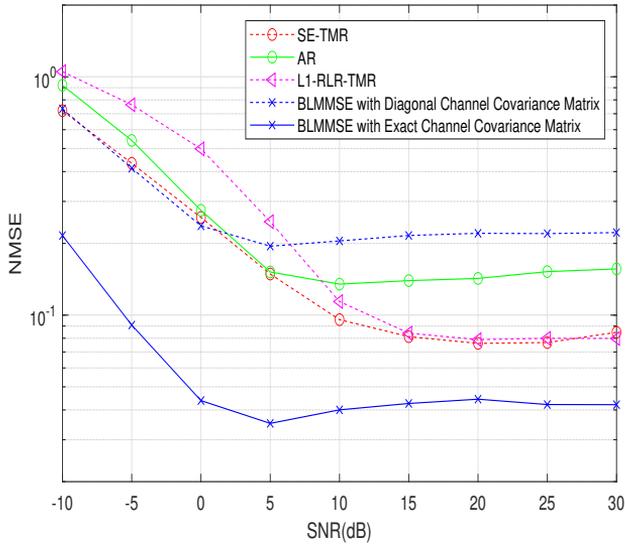


Figure 4.1. NMSE vs. SNR for $M = 16$, $N_s = 128$, and $L_k = 1$.

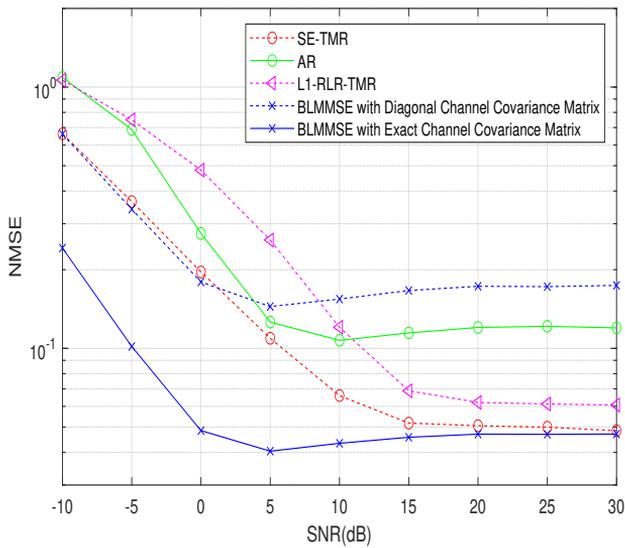


Figure 4.2. NMSE vs. SNR for $M = 16$, $N_s = 128$, and $L_k = 2$.

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 approach 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 covariance 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. Circ. 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. Circ. Syst. II*, vol. 47, no. 1, pp. 71–75, 2000.

- [23] W. X. Zheng , “Fast identification of autoregressive signals from noisy observations,” *IEEE Trans. Circ. 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. Nosssek, “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 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 Derivatives 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 $\text{SNR} = 10\log_{10} \left(\frac{\|\mathbf{HS}\|_F^2}{MN_s\sigma^2} \right)$. Therefore, the curves of both Figures should be shifted to the right for 22 dBs.

Publication 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.

© 2020

Reprinted with permission.



New estimation methods for autoregressive process in the presence of white observation noise

Majdoddin Esfandiari^{a,*}, Sergiy A. Vorobyov^a, Mahmood Karimi^b

^aDepartment of Signal Processing and Acoustics, Aalto University, Espoo, Finland

^bSchool of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

ARTICLE INFO

Article history:

Received 12 June 2019

Revised 13 December 2019

Accepted 13 January 2020

Available online 15 January 2020

Keywords:

Autoregressive signals

Noisy observations

Bias correction

Yule-Walker equations

Data analysis

ABSTRACT

This paper presents four new methods for estimating the parameters of an autoregressive (AR) process based on observations corrupted by white noise. The first three methods are iterative, while the last one is non-iterative. One method is designed to achieve an unbiased estimation of the AR parameters by undermining the destructive impact of observation noise in terms of utilizing the null space of the AR parameter vector. Another one uses both low- and high-order Yule-Walker equations to construct a constrained least squares optimization problem, in which the variance of observation noise is estimated by alternating between two equations. One more method exploits an approximation which leads to reducing the problem of estimating the AR parameters with arbitrary order p to estimating just two parameters, while the last one estimates the variance of the observation noise using the minimum eigenvalue of the enlarged autocorrelation matrix. The performance of the proposed methods is evaluated in terms of various numerical examples, which demonstrate their superiority in terms of accuracy and robustness against the observation noise compared to state-of-the-art existing methods in most simulation examples. It makes the proposed methods a good fit for practical analysis of data contaminated by observation noise, when AR modeling is applicable, and gives a range of choices of methods for different data analysis situations.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

In many signal processing applications, a random signal is modeled using autoregressive (AR) model to describe the signal of interest in a simple and effective way. Such application areas cover speech processing, noise cancellation, image processing, spectral estimation, biomedical signals modeling, digital communication [1–9], and many other data analysis problems. For example, AR modeling is used in modern data science applications such as climate and river flow forecasting [10,11], annual population assessment [12], and financial time series analysis [13–16] to name just a few. The broad employment of AR modeling is due to the simplicity of computing unknown model parameters and also excellent resolution performance. For instance, in [17–19], methods are developed for characterizing and classifying different sets of electroencephalogram (EEG) signals, as well as detecting artifacts in EEG signals using coefficients obtained from AR models. In addition to the problem of one-dimensional AR estimation, the prob-

lems of multichannel AR estimation and nonlinear AR estimation have been subject of active research (see [20–23]).

The conventional solution of the AR estimation problem is given by the standard least-squares (LS) derived using low-order Yule-Walker equations. However, in practical situations, the AR signal is contaminated by noise. Because of white observation noise corruption, the zero lag autocorrelation is biased, that is, it is leading to a biased solution for the Yule-Walker equations [24]. To estimate noisy AR parameters, three main types of techniques have been developed during past four decades. The techniques from the first category are designed to avoid zero lag autocorrelations by using high-order Yule-Walker equations. In this category, as the first step, the AR signal is modeled by the Autoregressive Moving Average (ARMA) model. Then, the AR parameters can be estimated by some method such as the Maximum Likelihood (ML) [25], the Recursive Prediction Error (RPE) [26], or by using the Modified Yule-Walker (MYW) equations [1]. Unfortunately, these methods require a lot of data for computing high autocorrelation lag estimates, which leads to errors in those autocorrelation lags and is problematic in practical data analysis applications.

On the other hand, the second type of the prevalent methods use the bias compensation principle (BCP) to estimate the noisy

* Corresponding author.

E-mail addresses: majdoddin.esfandiari@aalto.fi (M. Esfandiari), sergiy.vorobyov@aalto.fi (S.A. Vorobyov), karimi@shirazu.ac.ir (M. Karimi).

AR model. Removing bias from low-order Yule-Walker equations is the main idea of such methods. In general, the second type of methods can be divided into two subcategories as well. The methods in the first subcategory, which are known as subspace methods, are designed to model Yule-Walker equations as an eigenvalue problem and estimate both observation noise variance and the AR model parameters [5,27,34]. Although these methods provide accurate estimates in low-level noise case, their performance is poor in high-level noise situations [28].

Moreover, the methods that belong to the second subcategory attempt to find the best estimation of both observation noise variance and the AR model parameters by iterating between two sets of equations [28–32]. These methods are called improved least-squares (ILS) based methods. For instance, an ILS method with a direct structure (ILSD) has been suggested in [29], and extended to a fast convergent method in [30]. In [31], a method which uses inverse filtering equations in conjunction with Yule-Walker equations to find the desired solution is proposed. This method is called the Inverse filtering based improved least-squares (IFILS) method. Recently, a novel iterative-based method which obtains a perfect solution by solving a nonlinear equation in order to achieve an unbiased estimate of AR parameters has been introduced in [28]. The authors claimed that their method is able to reach efficient performance by picking the initial value of observation noise variance within a certain region [28]. However, our simulations showed that different initial values would lead to different estimates, which are not always the efficient ones. In fact, choosing a proper initial value is critical in order to achieve the best performance, and this can put serious difficulties due to the ambiguity in selecting the suitable initial value. We will clarify this point by means of simulations in Section 4.

The third type of methods exploit the concept named errors-in-variables (EIV) to estimate the noisy AR model [36,37]. For example in [36], the variance of observation noise is estimated by minimizing a cost function formed by high-order Yule-Walker equations, while the AR parameters are estimated via low-order Yule-Walker equations. Although, it is a non-iterative method, minimizing the proposed cost function leads to a one-dimensional search process, which turns this method to a computationally demanding one.

Based on the discussion above, it can be observed that devising new methods for estimating AR parameters in the presence of noise is vital for tackling the shortcomings of the existing methods. The desirable properties for such new methods are: 1) having certainty about choosing a proper initial value for iterative methods; 2) requiring only few iterations to find the best solution; 3) showing robustness against the presence of high-level observation noise; 4) being applicable to noisy AR processes with arbitrary poles locations; 5) showing robustness in scenarios where the number of available data is small; 6) having better performance compared to the existing state-of-the-art methods.

In this paper, four new unbiased methods for estimating AR parameters in the presence of white observation noise are proposed. Three methods are iterative and one is non-iterative. The first one is designed to reduce the destructive effect of observation noise variance in low-order Yule-Walker equations and it estimates the AR parameters in a recursive manner. The second one formulates the problem as a constrained LS optimization problem and finds the optimum value of the observation noise variance via an iterative algorithm, and then estimates the AR parameters. The third one uses an approximation which reduces the number of unknown parameters to only two parameters, and finds the AR parameters in a recursive manner. Finally, the fourth one estimates the variance of observation noise as the minimum eigenvalue of sufficiently enlarged autocorrelation matrix and uses it to estimate the AR parameters. Besides achieving on much higher accuracy in the case of high-level noise, one of the advantages of the proposed methods

is that they need far less iterations in order to achieve high accuracy in comparison with other state-of-the-art recursive methods. In the case of the fourth non-iterative method, it also can be used as a validation tool to evaluate the accuracy of other estimation methods. Moreover, the proposed methods can be used for noisy AR parameter estimation with arbitrary poles locations. Simulation results confirmed that the proposed methods have better performance compared to the other existing estimation methods.

The remainder of this paper is organized as follows. Section 2 introduces the noisy AR model and gives the problem formulation. In Section 3, the proposed methods and the corresponding algorithms in addition to a method of choosing the initial point for the proposed methods are presented. Section 4 provides numerical examples for comparing the performance of the proposed methods with that of three other state-of-the-art methods in various scenarios.

2. Data model

The noisy p th order real AR model can be represented by

$$\begin{aligned} x(t) &= a_1x(t-1) + a_2x(t-2) + \dots + a_px(t-p) + e(t) \\ &= \mathbf{a}^T \mathbf{x}_t + e(t) \end{aligned} \quad (1)$$

$$y(t) = x(t) + w(t) \quad (2)$$

where the driving noise $e(t)$ is the zero mean white stationary noise with the variance σ_e^2 , $\mathbf{a} = [a_1 a_2 \dots a_p]^T$ is the vector of coefficients of the AR model (the superscript T denotes the transpose operation), $\mathbf{x}_t = [x(t-1), \dots, x(t-p)]^T$, and $w(t)$ is the zero mean white stationary observation noise with the variance σ_w^2 . Moreover, the observation noise $w(t)$ in (2) is assumed to be uncorrelated with the driving noise $e(t)$, that is, $E\{w(t)e(n)\} = 0$ for all the N 's and T 's, where $E\{\cdot\}$ is the expectation operator.

The autocorrelation function of $y(t)$ is given by

$$\begin{aligned} r_y(k) &= E\{y(t)y(t-k)\} = E\{(x(t) + w(t))(x(t-k) + w(t-k))\} \\ &= r_x(k) + \sigma_w^2 \delta(k) \end{aligned} \quad (3)$$

where $r_x(k)$ is the autocorrelation function of the noiseless AR process $x(t)$, and $\delta(k)$ is the delta function. As it can be seen, the autocorrelation function of $y(t)$ at zero lag is biased due to the existence of observation noise. In fact, the following equation results from (3):

$$r_y(0) = E\{y(t)^2\} = E\{x(t)^2\} + \sigma_w^2. \quad (4)$$

Since the driving noise $e(t)$ is white and also independent of $x(t-i)$, $i > 0$, using (1), it is straightforward to find that

$$E\{x(t)^2\} = E\{(x(t))(\mathbf{a}^T \mathbf{x}_t + e(t))\} = \mathbf{r}_x^T \mathbf{a} + \sigma_e^2 \quad (5)$$

where $\mathbf{r}_x = [r_x(1) \ r_x(2) \ \dots \ r_x(p)]^T$.

Substituting (5) into (4) yields

$$r_y(0) = \mathbf{r}_x^T \mathbf{a} + \sigma_e^2 + \sigma_w^2. \quad (6)$$

The problem of noisy AR parameter estimation can be accomplished by estimating the noise variances σ_e^2 and σ_w^2 , and also estimating the parameters vector \mathbf{a} using N noisy observations of the process $y(t)$. According to [1], Yule-Walker equations can be written as

$$r_x(k) = \sum_{i=1}^p a_i r_x(k-i), \quad k \geq 1. \quad (7)$$

Evaluating (7) for $k = 1, \dots, p$, the following linear system of equations is obtained:

$$\mathbf{R}_x \mathbf{a} = \mathbf{r}_x \quad (8)$$

where

$$\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(-1) & \cdots & r_x(1-p) \\ r_x(1) & r_x(0) & \cdots & r_x(2-p) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(p-1) & r_x(p-2) & \cdots & r_x(0) \end{bmatrix}. \quad (9)$$

Combining (3) and (9), we obtain

$$\mathbf{R}_y = \mathbf{R}_x + \sigma_w^2 \mathbf{I}_p \quad (10)$$

$$\mathbf{r}_y = \mathbf{r}_x \quad (11)$$

where \mathbf{R}_y and \mathbf{r}_y are defined in a similar way to \mathbf{R}_x and \mathbf{r}_x , respectively, and \mathbf{I}_p is the $p \times p$ identity matrix.

The following equation follows from (8), (10), and (11):

$$\mathbf{R}_y \mathbf{a} - \sigma_w^2 \mathbf{a} = \mathbf{r}_y. \quad (12)$$

Multiplying both sides of (12) by \mathbf{R}_y^{-1} from the left and rearranging the terms, we obtain that

$$\mathbf{a} = \mathbf{R}_y^{-1} \mathbf{r}_y + \sigma_w^2 \mathbf{R}_y^{-1} \mathbf{a}. \quad (13)$$

The first term at the right-hand side of (13) is known as the LS estimate of \mathbf{a} , that is, $\mathbf{a}_{LS} = \mathbf{R}_y^{-1} \mathbf{r}_y$ [33]. Taking (13) into account one can see that \mathbf{a}_{LS} is a biased estimate of \mathbf{a} . Consequently, it can be concluded from (13) that we must have the value of observation noise variance σ_w^2 in order to remove the bias $\sigma_w^2 \mathbf{R}_y^{-1} \mathbf{a}$ from $\mathbf{a}_{LS} = \mathbf{R}_y^{-1} \mathbf{r}_y$.

Additionally, q high-order Yule-Walker equations can be obtained by writing (7), for $p+1 \leq k \leq p+q$, and using (3), as follows:

$$\mathbf{R}'_y \mathbf{a} = \mathbf{r}'_y \quad (14)$$

where

$$\mathbf{R}'_y = \begin{bmatrix} r_y(p) & r_y(p-1) & \cdots & r_y(1) \\ r_y(p+1) & r_y(p) & \cdots & r_y(2) \\ \vdots & \vdots & \ddots & \vdots \\ r_y(p+q-1) & r_y(p+q-2) & \cdots & r_y(q) \end{bmatrix} \quad (15)$$

$$\mathbf{r}'_y = [r_y(p+1) \quad r_y(p+2) \quad \cdots \quad r_y(p+q)]^T. \quad (16)$$

In the following section, new methods, for estimating the parameters of noisy AR models, which will be derived using both low- and high-order Yule-Walker equations, are introduced.

3. The proposed methods

3.1. The first proposed method

In this subsection, we propose a method for estimating the parameters of noisy AR model, which is based on undermining the destructive impact of σ_w^2 in (12). As it can be seen from (12), the term $\sigma_w^2 \mathbf{a}$ is the source of both non-linearity (with respect to σ_w^2 and \mathbf{a}) and bias in low-order Yule-Walker equations. Motivated by this observation, we devise an iterative scheme in which the detrimental effect of the nonlinear part ($\sigma_w^2 \mathbf{a}$) is considerably diminished over iterations such that eliminating it from the equations does not cause large errors in the modeling of the estimation problem. Thus, using this approximation, the resultant system of equations is linear with respect to \mathbf{a} , which can be solved efficiently.

To do so, consider that if a $p \times 1$ vector \mathbf{c} is a good initial estimate of \mathbf{a} , then we can assume that $\mathbf{c} = \mathbf{a} - \Delta$, where $\|\Delta\| \ll \|\mathbf{a}\|$. Like for any $p \times 1$ vector, there are $p-1$ different vectors \mathbf{b}_i , ($i = 1, \dots, p-1$), which satisfy the following conditions:

$$\mathbf{b}_i^T \mathbf{c} = 0, \quad \|\mathbf{b}_i\|^2 = 1, \quad \mathbf{b}_i^T \mathbf{b}_j = 0, \quad i, j = 1, \dots, p-1, \quad i \neq j. \quad (17)$$

These vectors \mathbf{b}_i 's construct an orthonormal basis for the null space of \mathbf{c} . Multiplying (12) from the left by the vector \mathbf{b}_i^T and rearranging the terms, the following $p-1$ equations are obtained:

$$\mathbf{b}_i^T \mathbf{R}_y \mathbf{a} = \mathbf{b}_i^T \mathbf{r}_y + \sigma_w^2 \mathbf{b}_i^T \mathbf{a}, \quad i = 1, \dots, p-1. \quad (18)$$

Using the relation $\mathbf{a} = \mathbf{c} + \Delta$ ($\|\Delta\| \ll \|\mathbf{a}\|$), and (17), we have

$$\sigma_w^2 \mathbf{b}_i^T \mathbf{a} = \sigma_w^2 \mathbf{b}_i^T (\mathbf{c} + \Delta) = \sigma_w^2 \mathbf{b}_i^T \Delta, \quad i = 1, \dots, p-1. \quad (19)$$

Needless to say that as \mathbf{c} approaches \mathbf{a} , $\sigma_w^2 \mathbf{b}_i^T \mathbf{a}$ approaches zero. In order to obtain an LS solution for \mathbf{a} , q high-order Yule-Walker equations should be added to the $p-1$ equations given by (18). It is clear that q is an arbitrary integer which should be larger than or equal to two. By adding the q high-order Yule-Walker equations presented in (14) to the $p-1$ equations derived in (18), the system of equations can be written in the following vector-matrix form:

$$\mathbf{H} \mathbf{a} = \mathbf{h} \quad (20)$$

where

$$\mathbf{H} = \begin{bmatrix} \mathbf{B}(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \\ \mathbf{R}'_y \end{bmatrix}, \quad \mathbf{B} = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_{p-1}]^T, \quad \mathbf{h} = \begin{bmatrix} \mathbf{B} \mathbf{r}_y \\ \mathbf{r}'_y \end{bmatrix}. \quad (21)$$

The LS solution of (20) with respect to \mathbf{a} is then given as

$$\mathbf{a} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{h}. \quad (22)$$

After computing \mathbf{a} , the LS estimate of σ_w^2 can be obtained using (12) as

$$\sigma_w^2 = \frac{\mathbf{a}^T \mathbf{R}_y \mathbf{a} - \mathbf{a}^T \mathbf{r}_y}{\|\mathbf{a}\|^2}. \quad (23)$$

Finally, an estimate of σ_e^2 can be obtained using (6) in the following form

$$\sigma_e^2 = r_y[0] - \mathbf{r}_y^T \mathbf{a} - \sigma_w^2. \quad (24)$$

The only issue remaining is to determine a way to calculate \mathbf{c} , which is an initial estimate of \mathbf{a} . Without \mathbf{c} , we cannot compute \mathbf{H} and \mathbf{h} . Thus, \mathbf{a} cannot be obtained from (22). Using the value obtained for \mathbf{c} , we will be able to use a recursive algorithm to estimate the AR parameters. This recursive algorithm is summarized in Algorithm 2. The important issue that should be pointed out here is that opting proper initial values has a significant effect on the performance of all iterative algorithms. Consequently, a perfect method for selecting proper initial values should be developed. In this case, contrary to most previous iterative AR estimation algorithms proposed in [29–31], we suggest to use a different initial value from \mathbf{a}_{LS} .

As mentioned in [28], we know from (10) that $\mathbf{R}_y = \mathbf{R}_x + \sigma_w^2 \mathbf{I}_p$. Therefore, we have

$$\lambda_{\min}(\mathbf{R}_y) = \lambda_{\min}(\mathbf{R}_x) + \sigma_w^2 \quad (25)$$

where $\lambda_{\min}(\mathbf{X})$ denotes the minimum eigenvalue of matrix \mathbf{X} . Since both \mathbf{R}_y and \mathbf{R}_x are symmetric and positive definite matrices, it can be concluded readily that $0 < \sigma_w^2 < \lambda_{\min}(\mathbf{R}_y)$. We propose to choose an initial value for σ_w^2 by assuming $\sigma_e^2 = 0$ in (24) and solving (24) for σ_w^2 . We aim to find an upper bound for σ_w^2 which is tighter than $\lambda_{\min}(\mathbf{R}_y)$, and use it as the initial value of σ_w^2 . Since both σ_w^2 and σ_e^2 are non-negative variables, the maximum value of σ_w^2 can be considered as the only root of (24) when σ_e^2 is set to zero (the minimum value). Using (12), (24) can be written as

$$\sigma_e^2 = f(\sigma_w^2) = r_y[0] - \sigma_w^2 - \mathbf{r}_y^T (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y \quad (26)$$

It should be pointed out that $f(\sigma_w^2)$ is a nonlinear function of σ_w^2 . The following theorem states and proves that this nonlinear function has only one root in the interval $(0, \lambda_{\min}(\mathbf{R}_y))$.

Theorem 1. The nonlinear function $f(\sigma_w^2)$ presented by (26) has exactly one real root over the interval $(0, \lambda_{\min}(\mathbf{R}_y))$.

Proof. The first statement that should be proved is that $f(\sigma_w^2)$ is strictly monotonic. The second statement is that $f(0)$ and $f(d\lambda_{\min}(\mathbf{R}_y))$ have different signs, when d approaches 1. After illustrating these two statements, the proof of the theorem will be complete. For proof of the first statement, see [28].

Now, we prove that $f(0)$ and $f(d\lambda_{\min}(\mathbf{R}_y))$ have different signs. In the case that $\sigma_w^2 = d\lambda_{\min}(\mathbf{R}_y)$, we can write:

$$\begin{aligned} \mathbf{r}_y^T (\mathbf{R}_y - d\lambda_{\min}(\mathbf{R}_y)\mathbf{I}_p)^{-1} \mathbf{r}_y &= \mathbf{r}_y^T (\mathbf{Q}(\mathbf{A} - d\lambda_{\min}(\mathbf{R}_y)\mathbf{I}_p)\mathbf{Q}^T)^{-1} \mathbf{r}_y \\ &= (\mathbf{Q}^T \mathbf{r}_y)^T (\mathbf{A} - d\lambda_{\min}(\mathbf{R}_y)\mathbf{I}_p)^{-1} \mathbf{Q}^T \mathbf{r}_y = \mathbf{v}^T (\mathbf{A} - d\lambda_{\min}(\mathbf{R}_y)\mathbf{I}_p)^{-1} \mathbf{v} \\ &= \sum_{i=1}^p \frac{1}{\lambda_i - d\lambda_{\min}(\mathbf{R}_y)} v_i^2 \end{aligned} \quad (27)$$

where \mathbf{Q} is a matrix whose columns are the orthonormal eigenvectors of matrix \mathbf{R}_y , and $\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_p)$ is a matrix whose diagonal elements are the eigenvalues of matrix \mathbf{R}_y . It can be seen from (27) that as d approaches 1, the right-hand side of (27) approaches $+\infty$, which results in $f(d\lambda_{\min}(\mathbf{R}_y))$ to approach $-\infty$. The only part that remains is to show that $f(0)$ has a positive value. Inserting $\sigma_w^2 = 0$ into (26), we have

$$f(0) = r_y(0) - \mathbf{r}_y^T \mathbf{R}_y^{-1} \mathbf{r}_y. \quad (28)$$

Taking (13) into account, we reach $\mathbf{R}_y^{-1} \mathbf{r}_y = (\mathbf{I}_p - \sigma_w^2 \mathbf{R}_y^{-1}) \mathbf{a}$. Thus, the second term in the right-hand side of (28) can be written as

$$\begin{aligned} \mathbf{r}_y^T \mathbf{R}_y^{-1} \mathbf{r}_y &= \mathbf{r}_y^T (\mathbf{I}_p - \sigma_w^2 \mathbf{R}_y^{-1}) \mathbf{a} = \mathbf{r}_y^T (\mathbf{I}_p - \sigma_w^2 \mathbf{R}_y^{-1}) (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y \\ &= \mathbf{r}_y^T (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y - \sigma_w^2 \mathbf{r}_y^T \mathbf{R}_y^{-1} (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y. \end{aligned} \quad (29)$$

Rearranging the terms in (26), we can write that

$$\mathbf{r}_y^T (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y = r_y[0] - \sigma_e^2 - \sigma_w^2. \quad (30)$$

Finally, combining (28)–(30), the following relationship is obtained for $f(0)$

$$\begin{aligned} f(0) &= r_y(0) - (r_y(0) - \sigma_e^2 - \sigma_w^2) + \sigma_w^2 \mathbf{r}_y^T \mathbf{R}_y^{-1} (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y \\ &= \sigma_e^2 + \sigma_w^2 + \sigma_w^2 \mathbf{r}_y^T \mathbf{R}_y^{-1} (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y. \end{aligned} \quad (31)$$

Because both σ_e^2 and σ_w^2 are positive, and also $\mathbf{R}_y^{-1} (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1}$ is a Positive Definite (PD) matrix over $0 < \sigma_w^2 < \lambda_{\min}(\mathbf{R}_y)$, it can be concluded that $f(0) > 0$. \square

The significant issue that should be dealt with here is to determine how to solve the nonlinear equation $f(\sigma_w^2) = 0$ and find $\hat{\sigma}_w^{2(0)}$, which is the initial value of σ_w^2 . The classic approach for finding roots of a nonlinear function is to use the Newton method. However, it is well-known that the performance of the Newton method highly depends on the closeness of the initial point to the actual root. Therefore, we develop a method to find the root of $f(\sigma_w^2)$ in the interval $(0, \lambda_{\min}(\mathbf{R}_y))$ with arbitrary precision in Section 3.1.1. It should be noted that this initial value for σ_w^2 is put into (12) to give us an initial estimate for \mathbf{a} , which will be considered as \mathbf{c} .

3.1.1. The proposed algorithm for finding the initial value of σ_w^2

In this subsection, an algorithm, which uses the fact that $f(\sigma_w^2)$ is strictly monotonic and has only one root over the interval $(0, \lambda_{\min}(\mathbf{R}_y))$, is suggested for calculating $\hat{\sigma}_w^{2(0)}$. The first step of this algorithm is to start from the lower and upper bounds of σ_w^2 , i.e., $D_1 = 0$ and $D_2 = 0.9999\lambda_{\min}(\mathbf{R}_y)$, and then calculate $f(D_1)$ and $f(D_2)$. If $f(D = (D_1 + D_2)/2) = 0$, then the initial value is found and the algorithm should be terminated. Otherwise, we can substitute either D_1 with D if $f(D) > 0$, or D_2 with D if $f(D) < 0$ based on Theorem 1. The aforementioned steps are repeated until the initial value is found. It is clear that this algorithm will always converge and find a precise solution for $f(\sigma_w^2) = 0$. The proposed algorithm is summarized in Algorithm 1.

Algorithm 1 The proposed algorithm for finding the initial value of σ_w^2 .

- 1: Compute autocorrelation estimates, using samples $\{y(1), y(2), \dots, y(N)\}$, that is, $\hat{r}_y(k) = \frac{1}{N} \sum_{t=k+1}^N y(t)y(t-k)$ ($k = 0, 1, \dots, p+q$), and then use them to form the estimates $\hat{\mathbf{R}}_y$, and $\hat{\mathbf{r}}_y$.
 - 2: Set δ_0 to be a small positive number. Here we choose the value 0.001 for δ_0 . After that, set $i = 0$, where i denotes the iteration index.
 - 3: Set $D_1 = 0$ and $D_2 = 0.9999\lambda_{\min}(\mathbf{R}_y)$, and calculate $f(D_1)$ and $f(D_2)$ via $f(D_m) = \hat{r}_y(0) - D_m - \hat{\mathbf{r}}_y^T (\hat{\mathbf{R}}_y - D_m \mathbf{I}_p)^{-1} \hat{\mathbf{r}}_y$ ($m = 1, 2$).
 - 4: Set $i = i + 1$ and calculate $D = \frac{D_1 + D_2}{2}$ and $f(D)$. If $|f(D)| \leq \delta_0$, set $\hat{\sigma}_w^{2(i)} = D$ and terminate the iterations.
- Otherwise, go to step 5.
- 5: If $f(D) > 0$, set $D_1 = D$, and if $f(D) < 0$, set $D_2 = D$ and go to step 4.

Algorithm 2 Algorithm for the first proposed method.

- 1: Compute autocorrelation estimates, using samples $\{y(1), y(2), \dots, y(N)\}$, that is, $\hat{r}_y(k) = \frac{1}{N} \sum_{t=k+1}^N y(t)y(t-k)$ ($k = 0, 1, \dots, p+q$), and then use them to form the estimates $\hat{\mathbf{R}}_y$, $\hat{\mathbf{r}}_y$, and $\hat{\mathbf{y}}_y$.
 - 2: Set $i = 0$. Let the initial variance of observation noise be $\hat{\sigma}_w^{2(0)} = D$, where D is calculated by Algorithm 1. Then compute $\mathbf{c} = \hat{\mathbf{a}}^{(0)} = (\hat{\mathbf{R}}_y - \hat{\sigma}_w^{2(0)} \mathbf{I}_p)^{-1} \hat{\mathbf{r}}_y$.
 - 3: Set $i = i + 1$ and use Singular Value Decomposition (SVD) to find $\mathbf{b}_j^{(i)}$'s ($j = 1, \dots, p-1$), which constitute an orthonormal basis for the null space of $\hat{\mathbf{a}}^{(i-1)}$, that is, they must satisfy the conditions below
 $\mathbf{b}_j^{(i)T} \hat{\mathbf{a}}^{(i-1)} = 0$, $\|\mathbf{b}_j^{(i)}\|^2 = 1$, $\mathbf{b}_j^{(i)T} \mathbf{b}_k^{(i)} = 0$ ($j, k = 1, \dots, p-1$, $j \neq k$).
 - 4: Use $\mathbf{b}_j^{(i)}$'s, $\hat{\mathbf{R}}_y$, $\hat{\sigma}_w^{2(i-1)}$, $\hat{\mathbf{r}}_y$, $\hat{\mathbf{y}}_y$ and $\hat{\mathbf{y}}_y$ to compute the estimates $\hat{\mathbf{B}}^{(i)}$, $\hat{\mathbf{H}}^{(i)}$ and $\hat{\mathbf{h}}^{(i)}$.
 - 5: Calculate $\hat{\mathbf{a}}^{(i)} = (\hat{\mathbf{H}}^{(i)T} \hat{\mathbf{H}}^{(i)})^{-1} \hat{\mathbf{H}}^{(i)T} \hat{\mathbf{h}}^{(i)}$.
 - 6: Calculate $\hat{\sigma}_w^{2(i)} = \frac{\hat{\mathbf{a}}^{(i)T} (\hat{\mathbf{R}}_y \hat{\mathbf{a}}^{(i)} - \hat{\mathbf{r}}_y)}{\|\hat{\mathbf{a}}^{(i)}\|^2}$. If $\frac{|\hat{\sigma}_w^{2(i)} - \hat{\sigma}_w^{2(i-1)}|}{|\hat{\sigma}_w^{2(i-1)}|} \leq \delta$, where δ is an appropriate small positive number, and $\|\cdot\|$ is the Euclidean norm, the iterations must be terminated, and the algorithm must be continued through step 7.
- Otherwise, go to step 3.
- 7: Calculate $\hat{\sigma}_e^2 = \hat{r}_y[0] - \hat{\mathbf{r}}_y^T \hat{\mathbf{a}}^{(i)} - \hat{\sigma}_w^{2(i)}$.

3.2. The second proposed method

Similar to the first proposed method, the second one is an iterative method. The core idea of the second proposed method is to set up a constrained optimization problem to find the best solution which leads to the minimum error in (12). In order to do that, we propose to use low-order Yule-Walker equations to construct the LS cost function, while setting the constraint to satisfy the first high-order Yule-Walker equation, as follows:

$$\begin{aligned} &\underset{\mathbf{a}}{\text{minimize}} && ((\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{a} - \mathbf{r}_y)^T ((\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{a} - \mathbf{r}_y) \\ &\text{subject to} && \bar{\mathbf{r}}^T \mathbf{a} = r_y(p+1) \end{aligned} \quad (32)$$

where $\bar{\mathbf{r}}^T$ is the first row of \mathbf{R}_y^* in (15). It is worth noting that we tested several combinations of high-order Yule-Walker equations as constraints in the second proposed approach and observed that using only the first high-order Yule-Walker equation leads to the best performance in most of the cases of AR processes. Using the Lagrangian multiplier method, the constrained problem (32) can be

converted to the problem of finding the proper \mathbf{a} and σ_w^2 which minimize the following Lagrangian function

$$\begin{aligned} L(\mathbf{a}, \sigma_w^2) &= ((\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{a} - \mathbf{r}_y)^\top ((\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{a} - \mathbf{r}_y) \\ &\quad + \lambda (\mathbf{a}^\top \bar{\mathbf{r}} - r_y(p+1)) \\ &= \mathbf{a}^\top (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2 \mathbf{a} - 2\mathbf{a}^\top (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{r}_y + \|\mathbf{r}_y\|^2 \\ &\quad + \lambda (\mathbf{a}^\top \bar{\mathbf{r}} - r_y(p+1)) \end{aligned} \quad (33)$$

where λ is a scalar Lagrangian multiplier. In order to find the parameters which minimize (33), the partial derivative of (33) should be calculated with respect to both \mathbf{a}^\top and σ_w^2 as follows:

$$\frac{\partial L(\mathbf{a}, \sigma_w^2)}{\partial \mathbf{a}^\top} = 2(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2 \mathbf{a} - 2(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{r}_y + \lambda \bar{\mathbf{r}} \quad (34)$$

$$\frac{\partial L(\mathbf{a}, \sigma_w^2)}{\partial \sigma_w^2} = -2\mathbf{a}^\top (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{a} + 2\mathbf{a}^\top \mathbf{r}_y. \quad (35)$$

Setting (34) and (35) to be zero leads us to the solution of (32). First, it is clear that setting (35) to be zero and calculating the LS estimate of σ_w^2 leads to (23). Moreover, setting (34) to be zero, we have

$$\begin{aligned} (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2 \mathbf{a} &= (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p) \mathbf{r}_y - \frac{\lambda}{2} \bar{\mathbf{r}} \\ \Rightarrow \mathbf{a} &= (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y - \frac{\lambda}{2} [(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2]^{-1} \bar{\mathbf{r}}. \end{aligned} \quad (36)$$

Since the result of (36) should satisfy the constrain of (32), substituting (36) into the constrain of (32) leads to closed form expression for λ , that is,

$$\begin{aligned} \bar{\mathbf{r}}^\top (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y - \frac{\lambda}{2} \bar{\mathbf{r}}^\top [(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2]^{-1} \bar{\mathbf{r}} &= r_y(p+1) \\ \Rightarrow \lambda &= 2 \frac{\bar{\mathbf{r}}^\top (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y - r_y(p+1)}{\bar{\mathbf{r}}^\top [(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2]^{-1} \bar{\mathbf{r}}}. \end{aligned} \quad (37)$$

Substituting (37) into (36), we obtain

$$\begin{aligned} \mathbf{a} &= (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y - \left(\frac{\bar{\mathbf{r}}^\top (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y - r_y(p+1)}{\bar{\mathbf{r}}^\top [(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2]^{-1} \bar{\mathbf{r}}} \right) \\ &\quad \times [(\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^2]^{-1} \bar{\mathbf{r}}. \end{aligned} \quad (38)$$

It can be observed that calculating \mathbf{a} and σ_w^2 is difficult using (38) and (23) directly. As a result, the use of an iterative method based on alternations between (38) and (23) is a proper choice. The output of this iterative method is a good estimate of σ_w^2 . Then, the estimated value of σ_w^2 can be used in a system of linear equations composed of both low- and high-order Yule-Walker equations for estimating \mathbf{a} as follows:

$$\mathbf{H}_1 \mathbf{a} = \mathbf{h}_1 \quad (39)$$

where

$$\mathbf{H}_1 = \begin{bmatrix} \mathbf{R}_y - \sigma_w^2 \mathbf{I}_p \\ \mathbf{R}_y \end{bmatrix}, \quad \mathbf{h}_1 = \begin{bmatrix} \mathbf{r}_y \\ \mathbf{r}_y \end{bmatrix}. \quad (40)$$

Similar to the first proposed method, the LS solution of (39) is obtained by $\mathbf{a} = (\mathbf{H}_1^\top \mathbf{H}_1)^{-1} \mathbf{H}_1^\top \mathbf{h}_1$. Finally, σ_w^2 can be estimated exactly in the same way as in the first proposed method, that is, using (24). The algorithm of the second proposed method is summarized in Algorithm 3.

Algorithm 3 Algorithm for the second proposed method.

1: Compute autocorrelation estimates, using samples $\{y(1), y(2), \dots, y(N)\}$, that is, $\hat{r}_y(k) = \frac{1}{N} \sum_{t=k+1}^N y(t)y(t-k)$ ($k = 0, 1, \dots, p+q$), and then use them to form the estimates $\hat{\mathbf{R}}_y$, $\hat{\mathbf{r}}_y$, $\hat{\mathbf{r}}$, and \hat{r}_y .

2: Set $i = 0$. Let the initial variance of observation noise be $\hat{\sigma}_w^{2(0)} = D$, where D is calculated by Algorithm 1.

3: Compute $\hat{\mathbf{a}}^{(i)} = \left(\hat{\mathbf{R}}_y - \hat{\sigma}_w^{2(i)} \mathbf{I}_p \right)^{-1} \hat{\mathbf{r}}_y - \left(\frac{\hat{\mathbf{r}}^\top (\hat{\mathbf{R}}_y - \hat{\sigma}_w^{2(i)} \mathbf{I}_p)^{-1} \hat{\mathbf{r}}_y - \hat{r}_y(p+1)}{\hat{\mathbf{r}}^\top [(\hat{\mathbf{R}}_y - \hat{\sigma}_w^{2(i)} \mathbf{I}_p)^2]^{-1} \hat{\mathbf{r}}} \right)$

$[(\hat{\mathbf{R}}_y - \hat{\sigma}_w^{2(i)} \mathbf{I}_p)^2]^{-1} \hat{\mathbf{r}}$.

4: Set $i = i + 1$ and compute $\hat{\sigma}_w^{2(i)} = \frac{\hat{\mathbf{a}}^{(i-1)\top} (\hat{\mathbf{R}}_y \hat{\mathbf{a}}^{(i-1)} - \hat{\mathbf{r}}_y)}{\|\hat{\mathbf{a}}^{(i-1)}\|^2}$. If $\frac{|\hat{\sigma}_w^{2(i)} - \hat{\sigma}_w^{2(i-1)}|}{|\hat{\sigma}_w^{2(i-1)}|} \leq \delta$, go to step 5. Otherwise, go to step 3.

5: Compute the estimates $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{h}}_1$ where $\sigma_w^2 = \hat{\sigma}_w^{2(i)}$. Then compute $\hat{\mathbf{a}} = (\hat{\mathbf{H}}_1^\top \hat{\mathbf{H}}_1)^{-1} \hat{\mathbf{H}}_1^\top \hat{\mathbf{h}}_1$.

6: Calculate $\hat{\sigma}_e^2 = \hat{r}_y[0] - \hat{\mathbf{r}}_y^\top \hat{\mathbf{a}}^{(i)} - \hat{\sigma}_w^{2(i)}$.

3.3. The third proposed method

The main idea of the third proposed method is to reduce the dimension of the unknown parameters from p to two in each iteration, regardless of the model order p . Considering (12), it can be seen that $\mathbf{a} = (\mathbf{R}_y - \sigma_w^2 \mathbf{I}_p)^{-1} \mathbf{r}_y$, which means that \mathbf{a} inversely depends on the difference between the eigenvalues of \mathbf{R}_y and σ_w^2 . Motivated by this observation, we represent \mathbf{a} as a linear combination of the eigenvectors of \mathbf{R}_y . In fact, we convert the original unknown parameters (a_m , $m = 1, \dots, p$) to the new unknown parameters (α_m , $m = 1, \dots, p$), which are dependent on $\frac{1}{\lambda_m - \sigma_w^2}$ (λ_m 's are the eigenvalues of \mathbf{R}_y). Taking into account that $0 < \sigma_w^2 < \lambda_{\min}(\mathbf{R}_y)$, it can be concluded that different choices of σ_w^2 in that interval do not change the value of α_m as long as the corresponding eigenvalue λ_m is much larger than $\lambda_{\min}(\mathbf{R}_y)$. Consequently, we write

$$\mathbf{a} = \sum_{m=1}^p \alpha_m \mathbf{v}_m \quad (41)$$

where \mathbf{v}_m 's, $m = 1, \dots, p$ are the eigenvectors of the matrix \mathbf{R}_y which satisfy the characteristic equation $\mathbf{R}_y \mathbf{v}_m = \lambda_m \mathbf{v}_m$. It is worth noting that we assume

$$\|\mathbf{v}_m\|^2 = 1, \quad \mathbf{v}_m^\top \mathbf{v}_k = 0, \quad m, k = 1, \dots, p, \quad m \neq k \quad (42)$$

$$\lambda_1 < \lambda_2 < \dots < \lambda_p. \quad (43)$$

Thus, substituting (41) into (12) and using the fact that $\mathbf{R}_y \mathbf{v}_m = \lambda_m \mathbf{v}_m$, we obtain

$$\begin{aligned} \mathbf{R}_y \sum_{m=1}^p \alpha_m \mathbf{v}_m &= \mathbf{r}_y + \sigma_w^2 \sum_{m=1}^p \alpha_m \mathbf{v}_m \\ \Rightarrow \sum_{m=1}^p \alpha_m \lambda_m \mathbf{v}_m &= \mathbf{r}_y + \sigma_w^2 \sum_{m=1}^p \alpha_m \mathbf{v}_m. \end{aligned} \quad (44)$$

Multiplying (44) by \mathbf{v}_m^\top from the left, and using (42), we have

$$\alpha_m \lambda_m = \mathbf{v}_m^\top \mathbf{r}_y + \sigma_w^2 \alpha_m \Rightarrow \alpha_m = \frac{\mathbf{v}_m^\top \mathbf{r}_y}{\lambda_m - \sigma_w^2}, \quad m = 1, \dots, p. \quad (45)$$

As it is stated earlier in Subsection 3.1, it is known that $0 < \sigma_w^2 < \lambda_1$. Therefore, one can assume that choosing different values of σ_w^2 in the interval $(0, \lambda_1)$ does not change α_m for $m = 3, \dots, p$ dramatically since with larger m , λ_m is usually much bigger than σ_w^2 . As a result, we can assume that we are able to calculate

α_m ($m = 3, \dots, p$) by choosing a proper value for σ_w^2 in the interval $(0, \lambda_1)$ (for example choosing $\sigma_w^2 = D$ according to Algorithm 1) using (45). Thus, we can write

$$\mathbf{a} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \bar{\mathbf{x}} \quad (46)$$

where

$$\bar{\mathbf{x}} = \sum_{m=3}^p \alpha_m \mathbf{v}_m. \quad (47)$$

Substituting (46) and (47) into (14), and using (45), we can construct the following system of linear equations:

$$\mathbf{H}_2 \boldsymbol{\alpha} = \mathbf{h}_2 \quad (48)$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad \mathbf{H}_2 = \begin{bmatrix} \bar{\mathbf{A}} \\ \mathbf{R}'_y \mathbf{V} \end{bmatrix}, \quad \mathbf{h}_2 = \begin{bmatrix} \mathbf{V}^T \mathbf{r}_y \\ \mathbf{r}'_y - \mathbf{R}'_y \bar{\mathbf{x}} \end{bmatrix} \\ \bar{\mathbf{A}} = \begin{bmatrix} \lambda_1 - \sigma_w^2 & 0 \\ 0 & \lambda_2 - \sigma_w^2 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}. \quad (49)$$

The LS solution of (48) is obtained by $\boldsymbol{\alpha} = (\mathbf{H}_2^T \mathbf{H}_2)^{-1} \mathbf{H}_2^T \mathbf{h}_2$. Similar to both previously proposed methods, we cannot use (48) directly to estimate $\boldsymbol{\alpha}$ since the knowledge of the value of σ_w^2 is necessary for calculating $\bar{\mathbf{A}}$ according to (49). Thus, it is reasonable to use an iterative method to estimate the unknown parameters. As stated earlier, D (which is calculated by Algorithm 1) is a proper initial value for σ_w^2 and can be used to calculate α_m 's ($m \geq 3$), $\bar{\mathbf{x}}$, and $\bar{\mathbf{A}}$ using (45), (47), and (49), respectively. Then, an estimate of $\boldsymbol{\alpha}$ can be obtained as the LS solution of (48). Now, we have an estimate of $\boldsymbol{\alpha}$ obtained using (46), and can compute a new value for σ_w^2 using (23), and repeat the aforementioned procedure until a pre-defined stopping criteria is satisfied. In addition, σ_e^2 can be estimated via (24). The algorithm of the third proposed method is summarized in Algorithm 4.

Algorithm 4 Algorithm for the third proposed method.

- 1: Compute autocorrelation estimates, using samples $\{y(1), y(2), \dots, y(N)\}$, that is, $\hat{r}_y(k) = \frac{1}{N} \sum_{t=k+1}^N y(t)y(t-k)$ ($k = 0, 1, \dots, p+q$), and then use them to form the estimates $\hat{\mathbf{R}}_y$, $\hat{\mathbf{R}}'_y$, $\hat{\mathbf{r}}_y$, and $\hat{\mathbf{r}}'_y$. Then, apply eigendecomposition of $\hat{\mathbf{R}}_y$ and find λ_m 's and \mathbf{v}_m 's ($m = 1, \dots, p$) which satisfy (42) and (43).
- 2: Set $i = 0$. Let the initial variance of observation noise be $\hat{\sigma}_w^{2(i)} = D$, where D is calculated by Algorithm 1.
- 3: Set $i = i + 1$ and compute $\hat{\alpha}_m^{(i)} = \frac{\mathbf{v}_m^T \hat{\mathbf{r}}_y}{\lambda_m - \hat{\sigma}_w^{2(i-1)}}$ for $m = 3, \dots, p$.

Then, calculate $\hat{\mathbf{x}}^{(i)} = \sum_{m=3}^p \hat{\alpha}_m^{(i)} \mathbf{v}_m$.

- 4: Compute the estimates $\hat{\mathbf{H}}_2^{(i)}$ and $\hat{\mathbf{h}}_2^{(i)}$ where $\sigma_w^2 = \hat{\sigma}_w^{2(i)}$ and $\bar{\mathbf{x}} = \hat{\mathbf{x}}^{(i)}$. Then compute $\hat{\boldsymbol{\alpha}}^{(i)} = (\hat{\mathbf{H}}_2^{(i)T} \hat{\mathbf{H}}_2^{(i)})^{-1} \hat{\mathbf{H}}_2^{(i)T} \hat{\mathbf{h}}_2^{(i)}$.

- 5: Compute $\hat{\mathbf{a}}^{(i)} = \mathbf{V} \hat{\boldsymbol{\alpha}}^{(i)} + \bar{\mathbf{x}}$.

- 6: Compute $\hat{\sigma}_w^{2(i)} = \frac{\hat{\mathbf{a}}^{(i)T} (\hat{\mathbf{R}}_y \hat{\mathbf{a}}^{(i)} - \hat{\mathbf{r}}_y)}{\|\hat{\mathbf{a}}^{(i)}\|^2}$. If $|\frac{\hat{\sigma}_w^{2(i)} - \hat{\sigma}_w^{2(i-1)}}{\hat{\sigma}_w^{2(i-1)}}| \leq \delta$, go to step 7.

Otherwise, go to step 3.

- 7: Calculate $\hat{\sigma}_e^2 = \hat{r}_y[0] - \hat{\mathbf{a}}^{(i)T} \hat{\mathbf{a}}^{(i)} - \hat{\sigma}_w^{2(i)}$.
-

3.4. The fourth proposed method

Unlike the other proposed methods, the aim of the fourth one is to estimate AR parameters in a non-iterative manner. Towards this end, we can write

$$\mathbf{R}_y(m) = \mathbf{R}_x(m) + \sigma_w^2 \mathbf{I}_m, \quad m \geq 1 \quad (50)$$

where

$$\mathbf{R}_y(m) = E \left\{ \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-(m-1)) \end{bmatrix} \begin{bmatrix} y(t) & y(t-1) & \cdots & y(t-(m-1)) \end{bmatrix} \right\} \\ = \begin{bmatrix} r_y[0] & r_y[-1] & \cdots & r_y[1-m] \\ r_y[1] & r_y[0] & \cdots & r_y[2-m] \\ \vdots & \vdots & \ddots & \vdots \\ r_y[m-1] & r_y[m-2] & \cdots & r_y[0] \end{bmatrix} \quad (51)$$

and $\mathbf{R}_x(m)$ is built in the same way as $\mathbf{R}_y(m)$ where autocorrelation lags $r_x(\cdot)$ are calculated using noiseless data. It can be observed from (50) that minimum eigenvalue of $\mathbf{R}_y(m)$ is equal to minimum eigenvalue of $\mathbf{R}_x(m)$ plus σ_w^2 . Based on this observation, our aim is to prove that by increasing m , both minimum eigenvalues of $\mathbf{R}_y(m)$ and $\mathbf{R}_x(m)$ are reduced. Since both $\mathbf{R}_y(m)$ and $\mathbf{R}_x(m)$ are PD matrices [3], and σ_w^2 is positive, by increasing m , the minimum eigenvalue of $\mathbf{R}_y(m)$ approaches σ_w^2 . The following theorem states and proves it.

Theorem 2. *Enlarging the dimension of $\mathbf{R}_y(m)$ by increasing m will lead to reducing the minimum eigenvalue of $\mathbf{R}_y(m)$.*

Proof. We prove the theorem by showing that the minimum eigenvalue of $\mathbf{R}_y(k+1)$ is smaller than the minimum eigenvalue of $\mathbf{R}_y(k)$, where k is an arbitrary number greater than one. To do so, we write

$$\mathbf{R}_y(k) \mathbf{u} = \lambda_1 \mathbf{u}, \quad \mathbf{R}_y(k+1) \mathbf{z} = \lambda'_1 \mathbf{z}, \quad \|\mathbf{u}\| = \|\mathbf{z}\| = 1 \quad (52)$$

where λ_1 and \mathbf{u} are respectively the minimum eigenvalue of $\mathbf{R}_y(k)$ and its corresponding eigenvector. Similarly, λ'_1 and \mathbf{z} are respectively the minimum eigenvalue of $\mathbf{R}_y(k+1)$ and its corresponding eigenvector. Using the Rayleigh quotient [4], we only need to prove that

$$\mathbf{u}^T \mathbf{R}_y(k) \mathbf{u} > \mathbf{z}^T \mathbf{R}_y(k+1) \mathbf{z}. \quad (53)$$

To begin with, we can write the left hand side of (53) as

$$\mathbf{u}^T \mathbf{R}_y(k) \mathbf{u} = \begin{bmatrix} \mathbf{u}^T & 0 \end{bmatrix} \mathbf{R}_y(k+1) \begin{bmatrix} \mathbf{u} \\ 0 \end{bmatrix}. \quad (54)$$

Since it is clear from the Rayleigh quotient that the right hand side of (54) is equal to or greater than $\mathbf{z}^T \mathbf{R}_y(k+1) \mathbf{z}$, we only need to show that

$$\begin{bmatrix} \mathbf{u}^T & 0 \end{bmatrix} \mathbf{R}_y(k+1) \begin{bmatrix} \mathbf{u} \\ 0 \end{bmatrix} \neq \mathbf{z}^T \mathbf{R}_y(k+1) \mathbf{z}. \quad (55)$$

Let us assume that the left hand side of (55) is equal to the right hand side of (55), which implies that $\begin{bmatrix} \mathbf{u}^T & 0 \end{bmatrix} \mathbf{R}_y(k+1) \mathbf{z} = \mathbf{z}^T \mathbf{R}_y(k+1) \mathbf{z}$. Therefore, based on (52), it should hold that

$$\mathbf{R}_y(k+1) \begin{bmatrix} \mathbf{u} \\ 0 \end{bmatrix} = \lambda'_1 \begin{bmatrix} \mathbf{u} \\ 0 \end{bmatrix} \Rightarrow \mathbf{R}_y(k+1) \begin{bmatrix} \mathbf{u} \\ 0 \end{bmatrix} \\ = u_1 \begin{bmatrix} r_y[0] \\ r_y[1] \\ \vdots \\ r_y[k] \end{bmatrix} + u_2 \begin{bmatrix} r_y[-1] \\ r_y[0] \\ \vdots \\ r_y[k-1] \end{bmatrix} + \cdots + u_k \begin{bmatrix} r_y[-(k-1)] \\ r_y[-(k-2)] \\ \vdots \\ r_y[1] \end{bmatrix} \\ + 0 \begin{bmatrix} r_y[-k] \\ r_y[-(k-1)] \\ \vdots \\ r_y[0] \end{bmatrix} \quad (56)$$

where u_i is the i th element of \mathbf{u} . Taking into account the fact that $\mathbf{R}_y(k)\mathbf{u} = \lambda_1\mathbf{u}$, (56) can be rewritten as follows

$$\mathbf{R}_y(k+1) \begin{bmatrix} \mathbf{u} \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{u} \\ \sum_{i=1}^k u_i r_y[k+1-i] \end{bmatrix}. \quad (57)$$

Eq. (57) states that we should have

$$\sum_{i=1}^k u_i r_y[k+1-i] = 0 \quad (58)$$

which is not correct and we reached a contradiction. Therefore, $[\mathbf{u}^T \mathbf{0}]^T \neq \mathbf{z}^T$ and (55) is correct. \square

According to Theorem 2, we can choose a proper value for m (for example $m = 2p$), and calculate σ_w^2 as the minimum eigenvalue of $\mathbf{R}_y(m)$. Then, we can construct $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{h}}_1$ using (40) and calculate $\hat{\mathbf{a}} = (\hat{\mathbf{H}}_1^T \hat{\mathbf{H}}_1)^{-1} \hat{\mathbf{H}}_1^T \hat{\mathbf{h}}_1$. The main advantage of the fourth proposed method is that it is a non-iterative method and thus does not have any convergence issue. The algorithm of the fourth proposed method is summarized in Algorithm 5.

Algorithm 5 Algorithm for the fourth proposed method.

- 1: Choose a proper value for m . Then, compute autocorrelation estimates, using samples $\{y(1), y(2), \dots, y(N)\}$, that is, $\hat{r}_y(k) = \frac{1}{N} \sum_{t=k+1}^N y(t)y(t-k)$ ($k = 0, 1, \dots, \max(p+q, m)$), and then use them to form the estimate $\hat{\mathbf{R}}_y(m)$.
- 2: Apply eigendecomposition of $\hat{\mathbf{R}}_y(m)$ and set $\hat{\sigma}_w^2 = \lambda_{\min}$, where λ_{\min} is the minimum eigenvalue of $\hat{\mathbf{R}}_y(m)$.
- 3: Compute the estimates $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{h}}_1$. Then compute $\hat{\mathbf{a}} = (\hat{\mathbf{H}}_1^T \hat{\mathbf{H}}_1)^{-1} \hat{\mathbf{H}}_1^T \hat{\mathbf{h}}_1$.

3.5. Complexity analysis

In this subsection, a discussion about the required order of computational complexity for implementing each proposed method is provided. The corresponding complexity of computing the LS estimate of an unknown $p \times 1$ vector using an overdetermined system of linear equations is $O(p^3)$ (see for example [38]). Likewise, performing an SVD of a $p \times p$ matrix requires the same amount of computational complexity. Hence, for implementing the first proposed method, the overall required computational complexity is $O(p^3)$ due to calculating the SVD (step 3), and the LS estimate of $\hat{\mathbf{a}}^{(i)}$ (step 5) in i th iteration of Algorithm 2. Similarly, according to Algorithm 3, the overall needed computational complexity is also $O(p^3)$ for the second proposed method since its implementation requires computing inverse of $p \times p$ matrices (step 3) in i th iteration, and also calculating the final LS estimate (step 5). Moreover, the computations needed to implement the eigendecomposition of $\hat{\mathbf{R}}_y$ in the third proposed method (step 1), and calculate $\hat{\mathbf{a}}^{(i)}$ (step 4) is $O(p^3)$ as well. In the case of the fourth proposed method, the corresponding computational complexity is $O(p^3)$ because of the eigendecomposition of $\hat{\mathbf{R}}_y(m)$. The computations needed at each iteration of the EIV method [36], the IFILS method [31], and Xia and Zheng's method [28] are also of complexity $O(p^3)$, which means that each iteration of the proposed iterative methods needs the same order of computations as each iteration of the existing iterative methods. The difference arises from the fact that our proposed iterative methods require much smaller average number of iterations compared to the other iterative methods as it will be shown in the simulation section. In the case of the fourth proposed method, since m is greater than p (for example $m = 2p$), it needs more computations in comparison with each iteration of the above mentioned iterative methods. However, as we

pointed out earlier, the fourth proposed method is a non-iterative method, which means that we need to apply the eigendecomposition of $\hat{\mathbf{R}}_y(m)$ just once. Moreover, the constant increase by $2^3 = 8$ is negligible for a practical size p .

4. Simulation results

In this section, the performances of the proposed methods are compared with that of the EIV method [36], the IFILS method [31], and Xia and Zheng's method [28] by means of simulations. In order to evaluate the accuracy of the aforementioned methods in estimating AR parameters, four numerical examples are considered. In what follows, M stands for the number of trials and N denotes the number of data points available. These values are set to 1000 and 4000, respectively. The parameter estimation methods are evaluated in terms of the Relative Prediction Error (RPE), and Normalized Root Mean Squared Error (NRMSE), which are defined, respectively, as follows:

$$RPE = \frac{|PE - \sigma_e^2|}{\sigma_e^2} \quad (59)$$

$$PE = \frac{1}{M(N-p)} \sum_{m=1}^M \sum_{t=p+1}^N (y(t) - \hat{\mathbf{a}}_m^T \mathbf{y}_t) \quad (60)$$

$$NRMSE = \frac{\sqrt{((\sum_{m=1}^M \|\hat{\mathbf{a}}_m - \mathbf{a}\|^2)/M)}}{\|\mathbf{a}\|} \quad (61)$$

where $\mathbf{y}_t = [y(t-1), y(t-2), \dots, y(t-p)]^T$, and $\hat{\mathbf{a}}_m$ denotes the estimate of \mathbf{a} in the m th trial. Among these two, RPE, which shows how well each method predicts the noiseless AR signal, is more informative and reliable. To calculate RPE, we use 1000 sets of data which are independent from the test data whose AR parameters are estimated.

It should be mentioned that $\sigma_w^{2(0)}$ is chosen to be $\eta \lambda_{\min}(\mathbf{R}_y)$ in the Xia-Zheng method, and η is picked within the interval [0.55, 0.99] (see [28]). As mentioned before in the introduction section, the main drawback of the Xia-Zheng's method is that there is an uncertainty about choosing the value of η . The initial parameters of each simulated method tested are the following. In the IFILS method, the parameter q is set to 2, and the value of parameter δ , which determines when to terminate the iteration process, is set to 0.001 in all the examples. As mentioned before, choosing different values for parameters η , δ_1 , and δ_2 in the Xia-Zheng method [28] leads to different performances of the method. Therefore, in each example, we choose appropriate values for parameters η , δ_1 , and δ_2 , which lead to the best performance of the the Xia-Zheng's method. In addition, δ is set to 0.1 in all proposed methods which are iterative. Moreover, q is respectively set to 3, 3, 4, and 3 for the first to fourth proposed methods. Moreover, m is set to 8 for the fourth proposed method.

Example 1. Consider a fourth-order noisy AR process with $\mathbf{a} = [1.6771, -1.6875, 0.9433, -0.3164]^T$ and $\sigma_e^2 = 1$. It should be noted that $e(t)$ is assumed to be a zero mean white Gaussian process in all the examples. Two scenarios are considered. In the first one, $w(t)$ is assumed to be a zero mean white Gaussian process with $\sigma_w^2 = 0.056$, making signal-to-noise ratio (SNR) to be

$$SNR = 10 \log_{10} \frac{E\{x(t)^2\}}{\sigma_w^2} \approx 20 \text{ dB}.$$

Moreover, to show the reliability and robustness of the proposed methods, the situation in which the AR signal is corrupted by high-level noise with the variance $\sigma_w^2 = 4.6$ is studied. In this case, the SNR is about 1 dB.

Table 1 displays the means and standard deviations of the estimates of the AR parameters obtained from 1000 independent trials together with NRMSE, RPE, and Number of Iterations Per Trial

Table 1
Computed results of estimated parameters for SNR = 20 dB in Example 1.

True value	EIV	IFILS	Xia-Zheng method	Proposed method I	Proposed method II	Proposed method III	Proposed method IV
$a_1 = 1.6771$	1.6749 ± 0.0358	1.6729 ± 0.0293	1.6771 ± 0.0283	1.6729 ± 0.0436	1.6847 ± 0.0366	1.6846 ± 0.0484	1.7336 ± 0.0321
$a_2 = -1.6875$	-1.6828 ± 0.0680	-1.6791 ± 0.0542	-1.6874 ± 0.0525	-1.6804 ± 0.0793	-1.6986 ± 0.0690	-1.6986 ± 0.0854	-1.7735 ± 0.0617
$a_3 = 0.9433$	0.9385 ± 0.0649	0.9352 ± 0.0510	0.9431 ± 0.0459	0.9371 ± 0.0722	0.9518 ± 0.0648	0.9522 ± 0.0753	1.0082 ± 0.0589
$a_4 = -0.3164$	-0.3143 ± 0.0305	-0.3130 ± 0.0242	-0.3165 ± 0.0273	-0.3140 ± 0.0306	-0.3185 ± 0.0290	-0.3187 ± 0.0313	-0.3352 ± 0.0270
$\sigma_w^2 = 0.0560$	0.0552 ± 0.0074	0.0559 ± 0.0065	0.0576 ± 0.0062	0.0559 ± 0.0098	0.0635 ± 0.0064	0.589 ± 0.0114	0.0772 ± 0.0059
$\sigma_e^2 = 1$	1.0029 ± 0.0543	1.0053 ± 0.0449	0.9989 ± 0.0435	1.0066 ± 0.0726	0.9761 ± 0.0513	0.9813 ± 0.0845	0.8760 ± 0.0448
NRMSE(%)	4.0852	3.2780	3.1409	4.6593	4.1397	4.9902	6.0307
RPE(%)	0.1142	0.0958	0.0873	0.1748	0.1329	0.2442	0.4497
NIPT	-	15.9630	46.9330	2.8600	3.8080	3.3840	-

Table 2
Computed results of estimated parameters for SNR = 1 dB in Example 1.

True value	EIV	IFILS	Xia-Zheng method	Proposed method I	Proposed method II	Proposed method III	Proposed method IV
$a_1 = 1.6771$	1.1552 ± 0.6019	0.3981 ± 0.0829	1.5039 ± 0.1547	1.5879 ± 0.1589	1.6298 ± 0.1604	1.6442 ± 0.1446	1.5224 ± 0.1777
$a_2 = -1.6875$	-0.9966 ± 0.7858	-0.0481 ± 0.0321	-1.3542 ± 0.3379	-1.5241 ± 0.2721	-1.5967 ± 0.2682	-1.6138 ± 0.2548	-1.5094 ± 0.2967
$a_3 = 0.9433$	0.4726 ± 0.5280	-0.1416 ± 0.0288	0.6325 ± 0.3721	0.8030 ± 0.2258	0.8621 ± 0.2180	0.8698 ± 0.2175	0.7967 ± 0.2412
$a_4 = -0.3164$	-0.1932 ± 0.1340	-0.0425 ± 0.0189	-0.1840 ± 0.2103	-0.2654 ± 0.0816	-0.2850 ± 0.0770	-0.2842 ± 0.0795	-0.2654 ± 0.0848
$\sigma_w^2 = 4.6$	2.8867 ± 2.1611	-0.6246 ± 2.4832	4.5279 ± 2.342	4.5666 ± 0.1992	4.6360 ± 0.2017	4.5986 ± 0.1818	4.5652 ± 0.1921
$\sigma_e^2 = 1$	3.8640 ± 3.5005	9.1475 ± 2.7775	1.3014 ± 0.3862	1.1722 ± 0.2934	1.0502 ± 0.2740	1.0487 ± 0.2498	1.2207 ± 0.3478
NRMSE(%)	58.3309	91.6373	29.4308	17.9332	15.9501	15.1377	19.4546
RPE(%)	78.9535	208.8564	10.4519	4.2038	3.5929	2.9560	5.2138
NIPT	-	78.6010	1.4500	1	1	1	-

(NIPT) for this case. As it can be seen, the performance of all methods is satisfactory. Note that η is set to 0.96, and δ_1 and δ_2 are set to 0.001 and 0.01, respectively, for the Xia-Zheng's method.

Table 2 displays the means and standard deviations of the estimates of the AR parameters obtained in the presence of high-level observation noise together with NRMSE, RPE, and NIPT. The parameter η is set to 0.96, and both δ_1 and δ_2 are set to 0.01 for the Xia-Zheng's method. As it can be seen, the proposed methods have much better performances than that of the EIV, the IFILS, and the Xia-Zheng's method in terms of parameters estimation accuracy and also NRMSE, and RPE. Note that selecting any value other than 0.96 for η from the interval [0.55, 0.99] leads to worse performance for the Xia-Zheng's method. As we stated, Table 2 confirms that the RPE describes the performance of the methods in a more complete way. As it can be observed in Table 2, the NRMSE for the IFILS method is about 90 percent, which is about six times, larger than that of the third proposed method. However, the RPE is about 208 percent, which is about 70 times, larger for the IFILS method than that of the third proposed method. The interesting fact is that the proposed iterative methods require much smaller number of iterations compared to IFILS and Xia-Zheng's method. The first reason is that the proposed iterative methods are initialized properly. Secondly, as the actual value of σ_w^2 increases, the ratio $\frac{|\hat{\sigma}_w^{2(0)} - \sigma_w^2|}{\sigma_w^2}$ decreases. Thus, in high SNR the proposed methods need about 3 iterations on average to find their best solution while they need 1 iteration on average in low SNR scenario in Example 1. As it can be seen, $\sigma_w^2 = 0.056$ in Table 1, and the average value of $\hat{\sigma}_w^{2(0)}$ in 1000 trials is about 0.15 which makes the average value of $\frac{|\hat{\sigma}_w^{2(0)} - \sigma_w^2|}{\sigma_w^2}$ be about 3. However, in the case of Table 2, the average value of $\hat{\sigma}_w^{2(0)}$ in 1000 trials is about 5.2 which makes the average value of $\frac{|\hat{\sigma}_w^{2(0)} - \sigma_w^2|}{\sigma_w^2}$ be less than 0.2. Thirdly, we choose the termination parameter δ to be 0.1 which avoids our proposed method to overfit to the noisy data in low SNR scenarios.

Example 2. In this example, the performances of the aforementioned methods are compared for a more general simulation setup to provide more insights into accuracy of parameter estimation by these methods. The parameters N and σ_e^2 are set to 4000 and 1,

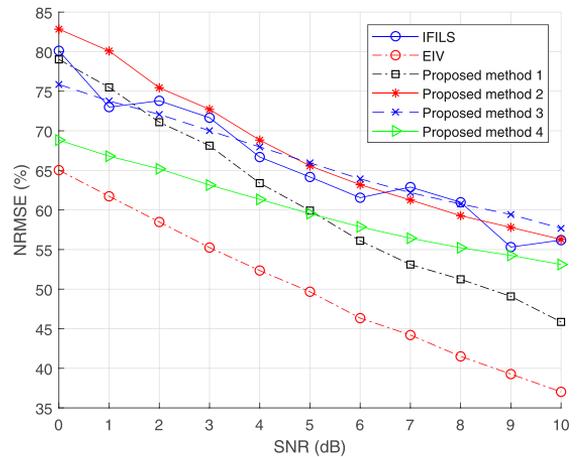


Fig. 1. Total NRMSE versus SNR for Example 2.

respectively. Additionally, η is set to 0.94, and both δ_1 and δ_2 are set to 0.05 for the Xia-Zheng's method. A hundred sets of poles, each set containing four poles selected randomly inside the unit circle, are considered as the poles of a hundred different AR processes. Then, for each AR process, the AR parameters are estimated via the aforementioned methods for 1000 trials, and NRMSE is calculated using (61). The mean of the resultant NRMSEs of a hundred different AR processes is then taken as the total NRMSE in this simulation example. We repeat this scenario for different values of SNR that varies from 0 dB to 10 dB. Fig. 1 shows the resultant total NRMSE versus SNR. Due to poor performance of Xia-Zheng's method, the resultant curve for this method is not shown. Moreover, Fig. 2 plots mean RPE versus SNR for the aforementioned setup. Again, the resultant curves of Xia-Zheng's method and IFILS are not shown because of their very poor performance.

It can be observed from Fig. 1 that the performance of all methods (except Xia-Zheng's method) is almost same in the terms of

Table 3
Computed results of estimated parameters for SNR= 10 dB in Example 3.

True value	EIV	IFILS	Xia-Zheng method	Proposed method I	Proposed method II	Proposed method III	Proposed method IV
$a_1 = 0.5500$	0.5494 ± 0.0187	0.5490 ± 0.0206	0.5503 ± 0.0269	0.5488 ± 0.0186	0.5813 ± 0.0258	0.5629 ± 0.0185	0.6210 ± 0.0154
$a_2 = 0.1550$	0.1543 ± 0.0139	0.1540 ± 0.0135	0.1534 ± 0.0140	0.1545 ± 0.0144	0.1503 ± 0.0141	0.1526 ± 0.0147	0.1500 ± 0.0137
$a_3 = -0.5495$	-0.5484 ± 0.0190	-0.5474 ± 0.0219	-0.5488 ± 0.0292	-0.5476 ± 0.0181	-0.5852 ± 0.0268	-0.5635 ± 0.0188	-0.6342 ± 0.0147
$a_4 = 0.6241$	0.6236 ± 0.0192	0.6222 ± 0.0206	0.6234 ± 0.0267	0.6225 ± 0.0179	0.6519 ± 0.0219	0.6383 ± 0.0181	0.6873 ± 0.0145
$\sigma_w^2 = 0.24$	0.2398 ± 0.0286	0.2384 ± 0.0330	0.2395 ± 0.0462	0.2389 ± 0.0277	0.3453 ± 0.0593	0.2663 ± 0.0286	0.3583 ± 0.0206
$\sigma_z^2 = 1$	1.0015 ± 0.0544	1.0050 ± 0.0630	1.0018 ± 0.0896	1.0014 ± 0.0523	0.8359 ± 0.0954	0.9450 ± 0.0535	0.7456 ± 0.0350
NRMSE(%)	3.5348	3.8647	4.9497	3.4483	7.0851	4.2181	12.9570
RPE(%)	0.1781	0.0267	0.0632	0.1135	0.5013	0.1742	1.7148
NIPT	-	11.1090	20.6180	2.0060	4.6000	3.9620	-

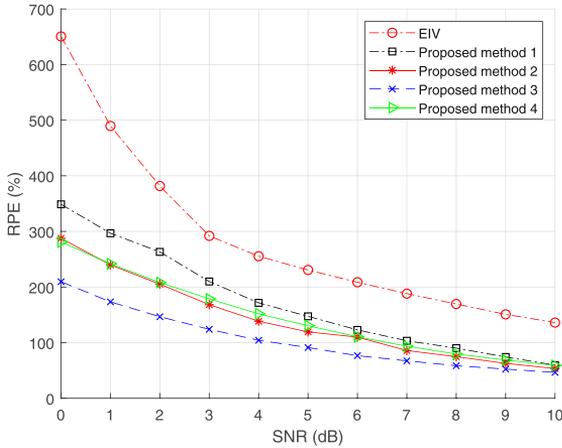


Fig. 2. Total RPE versus SNR for Example 2.

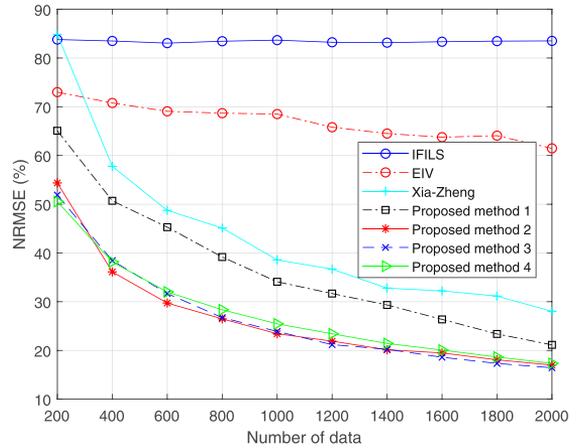


Fig. 3. NRMSE versus the number of data points for Example 4.

NRMSE. For example, for SNR=0dB, the difference between the worst performing method and the best performing method is less than 20%. Fig. 1 confirms that NRMSE does not give a complete picture about the performance of different methods of noisy AR parameter estimation. On the other hand, it can be seen from Fig. 2 that the third proposed method has the best performance, and EIV has the worst performance in terms of the RPE. The performance difference between the third proposed method and EIV is more than 400 percent for SNR=0dB. Thus, Fig. 2 shows that the proposed methods are more robust in dealing with various kind of AR processes in comparison to the EIV, the IFILS, and the Xia-Zheng’s methods for all SNRs. Moreover, since the selected value of η is not appropriate for all of AR processes generated in this example, the Xia-Zheng’s method does not demonstrate a good performance. Thus, this simulation example confirms the dependency of the performance of Xia-Zheng’s method on the choice of initial values, that is, η , δ_1 , and δ_2 .

Example 3. Consider a fourth-order noisy AR process with $\mathbf{a} = [0.5500, 0.1550, -0.5495, 0.6241]^T$ and $\sigma_w^2 = 1$. The poles of the AR model are located at $z = 0.3 \pm 0.8i$ and $z = -0.95$ and $z = 0.9$ in this example. Unlike the two previous examples, the poles of the AR model are comparatively close to the unit circle. Two situations are studied here. In the first one, SNR is set to 10 dB, leading to $\sigma_w^2 = 0.24$, and in the second one, SNR is set to -5 dB, resulting in $\sigma_w^2 = 7.3$.

Table 3 shows the means and standard deviations of the estimates of the AR parameters obtained from 1000 independent trials together with the NRMSE, RPE, and NIPT for SNR= 10 dB. The parameter η is set to 0.7, and both δ_1 and δ_2 are set to 0.001 for the Xia-Zheng’s method. As illustrated in Table 3, all the methods ex-

cept the second proposed method and the fourth proposed method perform well in terms of parameters estimation accuracy and also the NRMSE and RPE in the low-level noise scenario. It should be noted that the proposed methods need much less average number of iterations to achieve their best solution.

Table 4 lists the means and standard deviations of the estimates of the AR parameters obtained in the presence of high level observation noise together with the NRMSE, RPE, and NIPT. The parameter η is set to 0.885, and δ_1 and δ_2 are set to 0.05 and 0.02, respectively, for the Xia-Zheng’s method. The results presented in Table 4 show that the fourth proposed method have the best performance in the presence of high-level observation noise.

Example 4. Consider a fourth-order noisy AR process with $\mathbf{a} = [1.352, -1.338, 0.662, -0.24]^T$ and $\sigma_w^2 = 1$. The aim of the Example is to evaluate the performance of the aforementioned methods in the scenario that the number of observation data points varies from 200 to 2000 for SNR=2dB. Figs. 3 and 4, respectively, plot the NRMSE and RPE versus number of data points. As it can be observed, the proposed methods demonstrate much better performance compared to the three other methods. Therefore, it can be concluded that the proposed methods are more robust than the other methods in the scenarios with small number of data points available.

Generally, the convergence of the iterative methods is not guaranteed [31,35]. However, it is clear that choosing large values like 0.1 for δ will increase the chance of convergence. As it can be seen from the examples presented in this paper, the proper value of δ for the iterative proposed methods is equal to 0.1. This fact is an advantage for the proposed methods in comparison with other iterative methods like the IFILS and the Xia-Zheng’s methods.

Table 4
Computed results of estimated parameters for SNR= -5 dB in Example 3.

True value	EIV	IFILS	Xia-Zheng method	Proposed method I	Proposed method II	Proposed method III	Proposed method IV
$a_1 = 0.5500$	0.5298 ± 0.1044	0.5421 ± 0.1145	0.5194 ± 0.1051	0.5262 ± 0.1040	0.6195 ± 0.0865	0.6097 ± 0.0933	0.5951 ± 0.0859
$a_2 = 0.1550$	0.1598 ± 0.0521	0.1521 ± 0.0435	0.1549 ± 0.0547	0.1759 ± 0.0893	0.1503 ± 0.0740	0.1701 ± 0.0697	0.1491 ± 0.0621
$a_3 = -0.5495$	-0.5239 ± 0.1130	-0.5360 ± 0.1284	-0.5116 ± 0.1171	-0.5355 ± 0.1256	-0.6369 ± 0.0978	-0.6313 ± 0.1025	-0.6052 ± 0.0939
$a_4 = 0.6241$	0.5958 ± 0.1023	0.6064 ± 0.0964	0.5876 ± 0.1026	0.6169 ± 0.1085	0.6957 ± 0.0846	0.7084 ± 0.0783	0.6710 ± 0.0772
$\sigma_w^2 = 7.3$	7.2138 ± 0.4058	7.2113 ± 0.4022	7.1921 ± 0.4688	7.2683 ± 0.3985	7.5950 ± 0.2694	7.4433 ± 0.3112	7.3777 ± 0.3144
$\sigma_e^2 = 1$	1.1312 ± 0.4494	1.1048 ± 0.4490	1.1383 ± 0.5415	1.0403 ± 0.4234	0.5691 ± 0.1726	0.6909 ± 0.2632	0.8347 ± 0.2649
NRMSE(%)	19.5188	20.1514	20.3137	21.6196	21.5501	21.6643	18.1112
RPE(%)	4.1017	4.6911	4.5581	5.4780	5.3238	5.7754	3.7692
NIPT	-	27.3280	1.4250	1.0560	1	1	-

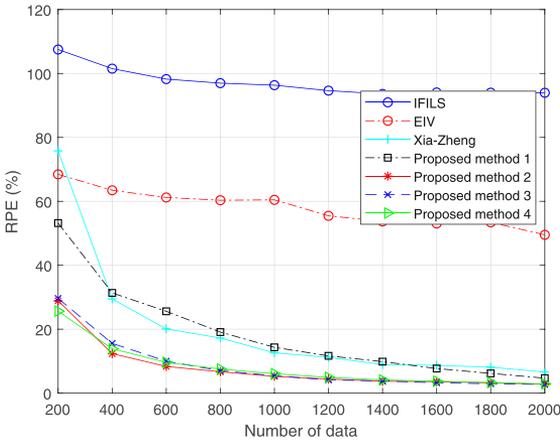


Fig. 4. RPE versus the number of data points for Example 4.

5. Discussion and conclusion

Four new methods have been proposed for estimating the parameters of AR processes in the presence of white observation noise. The first one has been based on undermining the destructive impact of observation noise variance, while the second method has estimated the AR parameters using a constrained LS optimization approach. The other one exploits an approximation which leads to reducing the problem of estimating the AR parameters with arbitrary order p to estimating just two parameters, while the last one estimates the variance of the observation noise using the minimum eigenvalue of the enlarged autocorrelation matrix. Based on the simulations, the performance of the proposed methods have been evaluated and compared with three other state-of-the-art methods. The simulation results have demonstrated that the proposed methods typically had better performances than the other methods in terms of requiring a smaller average number of iterations (for iterative methods), having smaller NRMSE and RPE in estimation, and also showing better robustness in the presence of high level observation noise in various situations with different locations of the AR model poles. Unlike the Xia-Zheng's method, the iterative proposed methods have enjoyed the advantage of starting from an initial point that makes them more reliable. In addition, the fourth proposed method performs well in estimating AR parameters, despite the fact that it is a non-iterative method. Therefore, the proposed methods give a range of choices for different data analysis situations, when AR modeling is applicable.

Based on the simulations results, it is recommended to use the first and the last proposed methods in the scenarios in which at least one of the AR poles is close to the unit circle. Additionally,

it is seen from the figures that the second and the third proposed method provide better estimates, and show more robustness against observation noise in most of the cases. Therefore, we suggest to use the second and especially the third proposed methods in the cases when there is no prior information about the poles locations. Moreover, it is desirable to use the third proposed method in the cases when the size of data available is small. Although the fourth proposed method usually provides a little worse performance for small data size, since it is non-iterative, it can be used in any scenarios to find a rough and valid estimates of the unknown parameters for validation of the result of the other applicable methods.

Declaration of Competing Interest

None.

Acknowledgments

This work was supported in part by Academy of Finland under Research Grant 319822.

References

- [1] S.M. Kay, *Modern Spectral Estimation*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
- [2] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [3] P. Stoica, R.L. Moses, *Spectral Analysis of Signals*, Prentice-Hall, Upper Saddle River, NJ, 2005.
- [4] R.A. Horn, C.A. Johnson, *Matrix Analysis*, Cambridge university press, Upper Saddle River, Cambridge, 1985.
- [5] M.D. Srinath, P.K. Rajasekaran, 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.* 133 (2017) 52–63.
- [7] L. Legrand, E. Grivel, Jeffrey's divergence between autoregressive processes distributed by additive white noise, *Signal Process.* 149 (2018) 162–178.
- [8] S. Dutta, M. Singh, 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, *Biomed. Signal Process. Control* 43 (2018) 174–182.
- [9] S. Ganapathy, Multivariate autoregressive spectrogram modeling for noisy speech recognition, *IEEE Signal Process. Lett.* 24 (9) (2017) 1373–1377.
- [10] O. Terzi, G. Ergin, Forecasting of monthly river flow with autoregressive modeling and data-driven techniques, *Neural Comput. Appl.* 25 (1) (2014) 179–188.
- [11] J.J. Musa, Stochastic modeling of Shiroro river stream flow process, *Am. J. Eng. Res.* 2 (6) (2013) 49–54.
- [12] N. Tolimieri, E.E. Holmes, G.D. Williams, R. Pacunski, D. Lowry, Population assessment using multivariate time-series analysis: a case study of rockfishes in puget sound, *Ecol. Evol.* 7 (8) (2017) 2846–2860.
- [13] M. Marcellino, J.H. Stock, M.W. Watson, A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series, *J. Econ.* 135 (1) (2006) 499–526.
- [14] R. Weron, A. Misiorek, Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models, *Int. J. Forecast.* 24 (4) (2008) 744–763.
- [15] A.K. Lohani, R. Kumar, R.D. Singh, Hydrological time series modeling: a comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques, *J. Hydrol.* 442 (2012) 23–35.
- [16] S. Feng, W. Ren, M. Han, Y.W. Chen, Robust manifold broad learning system for large-scale noisy chaotic time series prediction: a perturbation perspective, *Neural Netw.* 117 (2019) 179–190.

- [17] E.D. Ubeyli, Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals, *Expert Syst. Appl.* 37 (2010) 233–239.
- [18] V. Lawhern, W.D. Hairston, K. McDowel, M. Westerfield, K. Robbins, Detection and classification of subject-generated artifacts in EEG signals using autoregressive models, *J. Neurosci. Methods* 208 (2012) 181–189.
- [19] S.A. Vorobyov, A. Cichocki, Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis, *Biol. Cybern.* 86 (2002) 293–303.
- [20] A. Mahmoudi, M. Karimi, Estimation of the parameters of multichannel autoregressive signals from noisy observations, *Signal Process.* 88 (11) (2008) 2777–2783.
- [21] S. Maanan, B. Dumitrescu, C.D. Giurcaneanu, Conditional independence graphs for multivariate autoregressive models by convex optimization: Efficient algorithms, *Signal Process.* 133 (2017) 122–134.
- [22] A. Amanat, A. Mahmoudi, M. Hatam, Two-dimensional noisy autoregressive estimation with application to joint frequency and direction of arrival estimation, *Multidimens. Syst. Signal Process.* 29 (2) (2018) 671–685.
- [23] M. Kallias, P. Honeine, C. Francis, H. Amond, Kernel autoregressive models using Yule-Walker equations, *Multidimens. Syst. Signal Process.* 93 (11) (2013) 3053–3061.
- [24] S.M. Kay, Noise compensation for autoregressive spectral estimates, *IEEE Trans. Acoust. Speech Signal Process.* 28 (3) (1980) 292–303.
- [25] H. Tong, Autoregressive model fitting with noisy data by Akaike's information criterion, *IEEE Trans. Inf. Theory* 21 (4) (1975) 476–480.
- [26] A. Nehorai, P. Stoica, Adaptive algorithms for constrained ARMA signals in the presence of noise, *IEEE Trans. Acoust. Speech Signal Process.* 36 (8) (1988) 1282–1291.
- [27] W.X. Zheng, On estimation of autoregressive signals in the presence of noise, *IEEE Trans. Circ. Syst. II* 53 (12) (2006) 1471–1475.
- [28] Y. Xia, W.X. Zheng, Novel parameter estimation of autoregressive signals in the presence of noise, *Automatica* 62 (2015) 98–105.
- [29] W.X. Zheng, Autoregressive parameter estimation from noisy data, *IEEE Trans. Circ. Syst. II* 47 (1) (2000) 71–75.
- [30] W.X. Zheng, Fast identification of autoregressive signals from noisy observations, *IEEE Trans. Circ. Syst. II* 52 (1) (2005) 43–48.
- [31] A. Mahmoudi, M. Karimi, Inverse filtering based method for estimation of noisy autoregressive signals, *Signal Process.* 91 (7) (2011) 1659–1664.
- [32] A. Mahmoudi, M. Karimi, Parameter estimation of autoregressive signals from observations corrupted with colored noise, *Signal Process.* 90 (2010) 157–164.
- [33] M.H.A. Davis, R.B. Vinter, *Stochastic Modeling and Control*, Chapman and Hall, London, U.K., 1985.
- [34] C.E. Davila, A subspace approach to estimation of autoregressive parameters from noisy measurements, *IEEE Trans. Signal Process.* 46 (2) (1988) 531–534.
- [35] C.Z. Jin, L.J. Jia, Z.J. Yang, K. Wada, On convergence of a BCLS algorithms for noisy autoregressive process estimation, in: *Proceedings of the 41st IEEE Conference on Decision and Control*, 4, 2002, pp. 4252–4257.
- [36] R. Diversi, R. Guidorzi, U. Soverini, Identification of autoregressive models in the presence of additive noise, *Int. J. Adapt. Control Signal Process.* 22 (2008) 465–481.
- [37] R. Diversi, R. Guidorzi, U. Soverini, A noise-compensated estimation scheme for AR processes, in: *Proc. 44th IEEE Conf. Decision and Control and 8-th European Control Conf. Melia Sevilla, Seville, Spain, 2005*, pp. 4146–4151.
- [38] G.H. Golub, C.F.V. Loan, *Matrix Computations*, John Hopkins university press, Baltimore, 1996.

Publication 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.

© 2019

Reprinted with permission.

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

Majdoddin Esfandiari*, Sergiy A. Vorobyov*, and Mahmood Karimi†

*Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

†School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

Email: {majdoddin.esfandiari, sergiy.vorobyov}@aalto.fi, karimi@shirazu.ac.ir

Abstract—We consider the problem of estimating the parameters of autoregressive (AR) processes in the presence of white observation noise with unknown variance, which appears in many signal processing applications such as spectral estimation, and speech processing. A new non-iterative subspace-based method named extended subspace (ESS) method is developed. The basic idea of the ESS is to estimate the variance of the observation noise via solving a generalized eigenvalue problem, and then estimate the AR parameters using the estimated variance. The major advantages of the ESS method include excellent reliability and robustness against high-level noise, and also estimating the AR parameters in a non-iterative manner. Simulation results help to evaluate the performance of the ESS method, and demonstrate its robustness.

Index Terms—Autoregressive signals, Noisy observations, Yule-Walker equations, Subspace-based method

I. INTRODUCTION

In many signal processing applications, the autoregressive (AR) modeling of random signals is used to describe the signal of interest (SoI) in a simple and effective way. The application areas cover array processing, spectral estimation, speech processing, noise cancellation, image processing, biomedical signal processing, and communication channel estimation, especially array processing for pilot decontamination in massive multiple-input multiple-output systems and mmWave channel estimation [1]–[9]. The broad usage of AR modeling is due to its simplicity of computing unknown model parameters and its excellent resolution performance. In addition to the problem of one-dimensional AR estimation, the problems of multichannel AR estimation and nonlinear AR estimation have been subject of active research [10]–[13]. The conventional solution of the AR estimation problem is the standard least-squares (LS) derived by low-order Yule-Walker equations. However, in practical situations, the AR signal may be contaminated by noise. Because of white observation noise corruption of the measurements, the zero lag autocorrelation is biased, leading to a biased solution of Yule-Walker equations [14].

To estimate noisy AR parameters, three main types of techniques have been developed in the past decades. Techniques belonging to the first type aim to avoid zero lag autocorrelations by using high-order Yule-Walker equations. As the first step, the AR signal is modeled by the AR moving

average (ARMA) model. Then, the AR parameters can be estimated by some methods such as the maximum likelihood (ML) method [15], the recursive prediction error (RPE) [16], as well as using the modified Yule-Walker (MYW) equations [1]. Unfortunately, these methods suffer from the lack of data for computing high autocorrelation lags estimates, which causes the presence of error in those autocorrelation lags.

Methods of the second type use the bias compensation principle to estimate the noisy AR model. Removing bias from low-order Yule-Walker equations is the key then. In general, such methods can be divided into two subcategories. Methods in the first subcategory, which are known as subspace methods, model Yule-Walker equations as an eigenvalue problem and estimate both the observation noise variance and the AR model parameters [3], [17], [24]. Methods in the second subcategory attempt to find the best estimation of both observation noise variance and the AR model parameters by iterating between two sets of equations [18]–[22]. Such methods are called improved least-squares (ILS) based methods. In [19], an ILS method with a direct structure (ILSD) has been suggested, while in [20] it has been extended to achieve fast convergence. In [21], the inverse filtering based improved least-squares (IFILS) method, which uses inverse filtering equations in conjunction with Yule-Walker equations to find the desired solution, has been proposed. Recently, a novel iterative-based method which obtains a perfect solution by solving a nonlinear equation in order to achieve an unbiased estimate of AR parameters has been developed in [18]. It is claimed that this method is able to achieve efficient performance by picking the initial value of observation noise variance within a certain region [18]. However, our simulations showed that different initial values would lead to different estimates, which are not always the efficient ones.

The third type of methods exploit the concept named errors-in-variables (EIV) to estimate the noisy AR model [27]–[28]. For example in [27], the variance of observation noise is estimated by minimizing a cost function formed by high-order Yule-Walker equations, while the AR parameters are estimated via low-order Yule-Walker equations. Although, it is a non-iterative method, minimizing the proposed cost function leads to a one-dimensional search process, which turns this method

to a computationally demanding one in comparison with the subspace-based methods.

In this paper, a novel subspace-based method for estimating the AR parameters contaminated by white noise is developed. This method tries to estimate the variance of the observation noise by solving a generalized eigenvalue problem, followed by using the estimated variance to estimate the AR parameters. This method is non-iterative, computationally efficient, and it demonstrates excellent robustness against high-level observation noise. In addition, unlike the iterative methods, the convergence problem is not existing here. Simulation results confirm our claims.

II. DATA MODEL

The noisy p th order real AR model can be represented by

$$x(t) = a_1x(t-1) + a_2x(t-2) + \dots + a_px(t-p) + e(t) \\ = \mathbf{a}^T \mathbf{x}_t + e(t) \quad (1)$$

$$y(t) = x(t) + w(t) \quad (2)$$

where $e(t)$ is zero mean white stationary noise with variance σ_e^2 ; $\mathbf{a} = [a_1, a_2, \dots, a_p]^T$ is the vector of coefficients of the AR model (T denotes the transpose operation); $\mathbf{x}_t = [x(t-1), x(t-2), \dots, x(t-p)]^T$; and $w(t)$ is zero mean white stationary observation noise with variance σ_w^2 . Moreover, $w(t)$ in (2) is assumed to be uncorrelated with the driving noise $e(t)$, that is, $E\{w(t)e(n)\} = 0$ for all the N s and T s, where $E\{\cdot\}$ is the expectation operator.

The autocorrelation function of $y(t)$ can be presented by :

$$r_y[k] = E\{y(t)y(t-k)\} \\ = E\{(x(t) + w(t))(x(t-k) + w(t-k))\} \\ = r_x[k] + \sigma_w^2 \delta[k] \quad (3)$$

where $r_x[k]$ is the autocorrelation function of the noiseless AR process $x(t)$, and $\delta[k]$ is the delta function. The result obtained in (3) clearly indicates that the presence of the observation noise causes zero lag of the autocorrelation function of $y(t)$ to be biased. By taking into account the fact that $e(t)$ is white and also independent of $x(t-i)$, $i > 0$, (3) can be written for $k = 0$ as

$$r_y[0] = E\{y(t)^2\} = E\{x(t)^2\} + \sigma_w^2 \\ = E\{(x(t))(\mathbf{x}_t^T \mathbf{a} + e(t))\} + \sigma_w^2 = \mathbf{r}_x^T \mathbf{a} + \sigma_e^2 + \sigma_w^2 \quad (4)$$

where $\mathbf{r}_x = [r_x[1], r_x[2], \dots, r_x[p]]^T$. In addition, according to [1], the well-known Yule-Walker equations can be written as

$$r_x[k] = \sum_{i=1}^p a_i r_x[k-i], \quad k \geq 1. \quad (5)$$

As a result, by evaluating (5) for $k = 1, \dots, p$, the following linear system of equations is obtained:

$$\mathbf{R}_x \mathbf{a} = \mathbf{r}_x \quad (6)$$

where

$$\mathbf{R}_x = \begin{bmatrix} r_x[0] & r_x[-1] & \cdots & r_x[1-p] \\ r_x[1] & r_x[0] & \cdots & r_x[2-p] \\ \vdots & \vdots & \ddots & \vdots \\ r_x[p-1] & r_x[p-2] & \cdots & r_x[0] \end{bmatrix}. \quad (7)$$

Combining (3), (6) and (7), we obtain

$$\mathbf{R}_y \mathbf{a} = \mathbf{r}_y + \sigma_w^2 \mathbf{a} \quad (8)$$

where \mathbf{R}_y and \mathbf{r}_y are defined in a similar way as \mathbf{R}_x and \mathbf{r}_x , and also $\mathbf{r}_y = \mathbf{r}_x$. Note that p equations (8) are called the low-order Yule-Walker equations. By multiplying both sides of (8) by \mathbf{R}_y^{-1} from the left, we obtain

$$\mathbf{a} = \mathbf{R}_y^{-1} \mathbf{r}_y + \sigma_w^2 \mathbf{R}_y^{-1} \mathbf{a} \quad (9)$$

Additionally, q high-order Yule-Walker equations can be obtained by evaluating (5), for $p+1 \leq k \leq p+q$, and using (3) as follows:

$$\mathbf{R}_q \mathbf{a} = \mathbf{r}_q \quad (10)$$

where $\mathbf{r}_q = [r_y[p+1], r_y[p+2], \dots, r_y[p+q]]^T$, and

$$\mathbf{R}_q = \begin{bmatrix} r_y[p] & r_y[p-1] & \cdots & r_y[1] \\ r_y[p+1] & r_y[p] & \cdots & r_y[2] \\ \vdots & \vdots & \ddots & \vdots \\ r_y[p+q-1] & r_y[p+q-2] & \cdots & r_y[q] \end{bmatrix} \quad (11)$$

III. EXTENDED SUBSPACE METHOD

In this section, we propose a novel non-iterative subspace-based method for the problem above. This method is primarily based on combining the low- and high-order Yule-Walker equations, which are respectively given in (8) and (10), and estimating σ_w^2 via converting the resultant linear system of equations to a generalized eigendecomposition problem. After obtaining σ_w^2 , \mathbf{a} can be computed as the LS solution of the aforementioned system of equations.

To begin with, by combining (8) and (10), we obtain

$$\mathbf{A} \mathbf{a} = \mathbf{b} + \sigma_w^2 \mathbf{c} \quad (12)$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{R}_y \\ \mathbf{R}_q \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{r}_y \\ \mathbf{r}_q \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} \mathbf{a} \\ \mathbf{0}_q \end{bmatrix} \quad (13)$$

where $\mathbf{0}_q$ denotes a $q \times 1$ vector whose all entries are zero. Since \mathbf{b} is a $(p+q) \times 1$ vector, there exist $p+q-1$ vectors \mathbf{d}_i ($i = 1, 2, \dots, p+q-1$) which satisfy the following conditions:

$$\mathbf{d}_i^T \mathbf{b} = 0 \quad i = 1, \dots, p+q-1 \quad (14)$$

$$\mathbf{d}_i^T \mathbf{d}_j = 0 \quad i \neq j. \quad (15)$$

Consequently, we can construct a $(p+q-1) \times (p+q)$ matrix \mathbf{D} whose rows are \mathbf{d}_i^T . By premultiplying (12) by \mathbf{D} , and taking advantage of (14), we obtain

$$\mathbf{D} \mathbf{A} \mathbf{a} = \sigma_w^2 \mathbf{D} \mathbf{c}. \quad (16)$$

In addition, by taking the definition of \mathbf{c} in (13) into account, (16) can be rewritten as

$$(\mathbf{D}\mathbf{A} - \sigma_w^2 \mathbf{E})\mathbf{a} = \mathbf{0}_{p+q-1} \quad (17)$$

where \mathbf{E} denotes the matrix constructed by the first p columns of \mathbf{D} . As it can be observed, (17) appears to have the form of the generalized eigendecomposition problem.

Multiplying both sides of (17) by $(\mathbf{D}\mathbf{A} - \sigma_w^2 \mathbf{E})^T$, we obtain a quadratic eigenvalue problem as follows:

$$(\mathbf{G}_0 + \sigma_w^2 \mathbf{G}_1 + (\sigma_w^2)^2 \mathbf{G}_2)\mathbf{a} = \mathbf{0}_p \quad (18)$$

where

$$\begin{aligned} \mathbf{G}_0 &= \mathbf{A}^T \mathbf{D}^T \mathbf{D} \mathbf{A}, \quad \mathbf{G}_1 = -(\mathbf{A}^T \mathbf{D}^T \mathbf{E} + \mathbf{E}^T \mathbf{D} \mathbf{A}), \\ \mathbf{G}_2 &= \mathbf{E}^T \mathbf{E}. \end{aligned} \quad (19)$$

Several approaches have been presented in the literature in order to solve (18) and find σ_w^2 [26]. We can rewrite (18) as a generalized eigenvalue problem in the following way [26]:

$$(\mathbf{P} - \sigma_w^2 \mathbf{Q})\bar{\mathbf{a}} = \mathbf{0}_{2p} \quad (20)$$

where

$$\mathbf{P} = \begin{bmatrix} \mathbf{G}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_p \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} -\mathbf{G}_1 & -\mathbf{G}_2 \\ \mathbf{I}_p & \mathbf{0} \end{bmatrix}, \quad \bar{\mathbf{a}} = \begin{bmatrix} \mathbf{a} \\ \sigma_w^2 \mathbf{a} \end{bmatrix} \quad (21)$$

where \mathbf{I}_p is the $p \times p$ identity matrix. After solving (20), since the resultant $2p$ eigenvalues are real or complex conjugate [26], the real eigenvalue with the smallest modulus should be chosen as σ_w^2 . However, in practical scenarios, due to the finite number of samples utilized to estimate the autocorrelation matrix, all of the eigenvalues obtained by solving (20) may be complex. Therefore, it is reasonable to choose the modulus of the eigenvalue whose absolute value of the imaginary part is minimum, as the estimated σ_w^2 .

The only issue remaining is to define a method for estimating \mathbf{a} . Making use of the estimated σ_w^2 , (12) can be rearranged as

$$\mathbf{H}\mathbf{a} = \mathbf{b} \quad (22)$$

where

$$\mathbf{H} = \begin{bmatrix} \mathbf{R}_y - \sigma_w^2 \mathbf{I}_p \\ \mathbf{R}_q \end{bmatrix}. \quad (23)$$

Thus, the LS solution of (23) with respect to \mathbf{a} is given by

$$\mathbf{a} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{b}. \quad (24)$$

After calculating σ_w^2 and \mathbf{a} , σ_e^2 can be obtained via (4).

IV. SIMULATION RESULTS

In this section, the performances of the proposed ESS algorithm is compared with that of the subspace method (SS) [24], the IFILS method [21], and Xia and Zheng's method [18] by means of simulations. In order to evaluate the accuracy of the aforementioned methods in estimating AR parameters, two numerical examples are considered. The number of trials is set to $M = 1000$. Moreover, the parameter estimation methods are evaluated in terms of relative error (RE), and normalized root

mean squared error (NRMSE), which are defined through the upcoming procedure:

$$RE = \frac{\|m(\hat{\mathbf{a}}) - \mathbf{a}\|}{\|\mathbf{a}\|} \quad (25)$$

$$m(\hat{\mathbf{a}}) = \frac{1}{M} \sum_{m=1}^M \hat{\mathbf{a}}_m \quad (26)$$

$$NRMSE = \frac{\sqrt{((\sum_{m=1}^M \|\hat{\mathbf{a}}_m - \mathbf{a}\|^2)/M)}}{\|\mathbf{a}\|} \quad (27)$$

where $\hat{\mathbf{a}}_m$ denotes the estimate of \mathbf{a} in the m th trial.

It should be mentioned that $\sigma_w^{(0)}$ is chosen to be $\eta \lambda_{\min}(\mathbf{R}_y)$ in the Xia-Zheng method; η is picked within the interval $[0.55, 0.99]$ (see [18]). As mentioned before, the main drawback of the Xia-Zheng method is that there is an uncertainty about choosing the value of η . Therefore, in each scenario, we choose the appropriate values for parameters η , δ_1 and δ_2 , which lead to the best performance of the Xia-Zheng method.

The initial parameters of each simulated method are as follows. In the IFILS method, the parameter q is set to 2, and the value of parameter δ , which determines when to terminate the iteration process, is set to 0.001 in all the examples. The parameter q is set to 8 for the SS method. In addition, q is set to 3 for the proposed ESS method.

Example 1. Consider a fourth-order noisy AR process with $\mathbf{a} = [1.6771, -1.6875, 0.9433, -0.3164]^T$ and $\sigma_e^2 = 1$. It should be noted that $e(t)$ is assumed to be a zero mean, white Gaussian process in all the examples. Two scenarios are considered. In the first one, $w(t)$ is assumed to be a zero mean, white Gaussian process with $\sigma_w^2 = 0.056$, making signal-to-noise ratio (SNR) as follows:

$$SNR = 10 \log_{10} \frac{m(x(t)^2)}{\sigma_w^2} \approx 20 \text{ dB}$$

The sample size is set to $N = 100$ in this situation. Table I displays the means and standard deviations of the estimates of the AR parameters obtained from 1000 independent trials together with RE, and NRMSE for this case. As it can be noticed, the Xia-Zheng method and the ESS show better performance in comparison with two other methods. Note that η is set to 0.96, and δ_1 and δ_2 are set to 0.001 and 0.01, respectively, for the Xia-Zheng method.

In the second case, in order to demonstrate the reliability and robustness of the ESS, we increase σ_w^2 to 4.6, which leads SNR to be 1 dB. In this case, N is set to be 4000. Table II shows the means and standard deviations of the estimates of the AR parameters obtained from 1000 independent trials together with RE, and NRMSE for this case. As it can be observed, the performance of the ESS is much better than that of the other three methods. The IFILS has very bad performance in this scenario. In this case, both δ_1 and δ_2 are considered to be 0.01 for the Xia-Zheng method. Moreover, η is set to 0.96.

TABLE I
COMPUTED RESULTS OF ESTIMATED PARAMETERS FOR $SNR = 20$ dB
WITH $N=100$ IN EXAMPLE 1

True value	SS	IFILS	Xia-Zheng method	ESS
$a_1 = 1.6771$	2.0546 ± 0.2567	1.5771 ± 0.2289	1.6538 ± 0.1962	1.7045 ± 0.1873
$a_2 = -1.6875$	-2.3323 ± 0.4017	-1.5266 ± 0.4073	-1.6668 ± 0.3521	-1.7219 ± 0.3522
$a_3 = 0.9433$	1.4755 ± 0.3053	0.8154 ± 0.3733	0.9429 ± 0.3232	0.9650 ± 0.3388
$a_4 = -0.3164$	-0.4996 ± 0.1011	-0.2723 ± 0.1677	-0.3259 ± 0.1453	-0.3183 ± 0.1592
$\sigma_w^2 = 0.0560$	0.2218 ± 0.0652	0.0563 ± 0.1093	0.0863 ± 0.0598	0.1103 ± 0.0520
$\sigma_e^2 = 1$	0.2273 ± 0.4176	1.1292 ± 0.4088	1.0023 ± 0.3043	0.8542 ± 0.2798
$RE(\%)$	36.2746	9.0271	1.2636	1.9031
$NRMSE(\%)$	42.5794	25.7180	20.8500	21.2980

TABLE II
COMPUTED RESULTS OF ESTIMATED PARAMETERS FOR $SNR = 1$ dB
WITH $N=4000$ IN EXAMPLE 1

True value	SS	IFILS	Xia-Zheng method	ESS
$a_1 = 1.6771$	1.9133 ± 0.3722	0.3981 ± 0.0829	1.5039 ± 0.1547	1.6088 ± 0.1445
$a_2 = -1.6875$	-2.0425 ± 0.6860	-0.0481 ± 0.0321	-1.3542 ± 0.3379	-1.5638 ± 0.2456
$a_3 = 0.9433$	1.2087 ± 0.6175	-0.1416 ± 0.0288	0.6325 ± 0.3721	0.8369 ± 0.2035
$a_4 = -0.3164$	-0.3966 ± 0.2508	-0.0425 ± 0.0189	-0.1841 ± 0.2103	-0.2772 ± 0.0745
$\sigma_w^2 = 4.6$	4.6255 ± 0.1901	-0.6246 ± 2.4832	4.5279 ± 0.2342	4.5783 ± 0.1892
$\sigma_e^2 = 1$	0.5190 ± 0.6857	9.1475 ± 2.7775	1.3014 ± 0.3862	1.1369 ± 0.2546
$RE(\%)$	19.7227	91.5637	19.5889	7.0255
$NRMSE(\%)$	44.4177	91.6373	29.4308	15.5602

TABLE III
COMPUTED RESULTS OF ESTIMATED PARAMETERS FOR $SNR = 10$ dB
WITH $N=100$ IN EXAMPLE 2

True value	SS	IFILS	Xia-Zheng method	ESS
$a_1 = 1.352$	2.0411 ± 1.1262	0.7839 ± 0.3288	1.2366 ± 0.3637	1.3933 ± 0.3017
$a_2 = -1.338$	-2.4366 ± 1.8283	-0.5626 ± 0.4276	-1.1973 ± 0.5554	-1.4045 ± 0.3836
$a_3 = 0.662$	1.6026 ± 1.5322	0.0645 ± 0.3514	0.5722 ± 0.4811	0.7010 ± 0.3884
$a_4 = -0.24$	-0.5992 ± 0.4939	-0.0428 ± 0.1616	-0.2224 ± 0.2127	-0.2553 ± 0.1787
$\sigma_w^2 = 0.38$	0.5524 ± 0.1246	-0.7137 ± 2.3966	0.2020 ± 0.8515	0.4554 ± 0.3231
$\sigma_e^2 = 1$	-0.0158 ± 1.2487	2.9012 ± 2.8150	1.3353 ± 1.1903	0.8411 ± 0.2947
$RE(\%)$	80.9451	56.6415	10.0421	4.3774
$NRMSE(\%)$	155.1096	65.4127	42.9537	32.2851

TABLE IV
COMPUTED RESULTS OF ESTIMATED PARAMETERS FOR $SNR = 2$ dB
WITH $N=4000$ IN EXAMPLE 2

True value	SS	IFILS	Xia-Zheng method	ESS
$a_1 = 1.352$	1.3481 ± 0.4356	0.4662 ± 0.1229	1.4521 ± 0.0928	1.3497 ± 0.1194
$a_2 = -1.338$	-1.3364 ± 0.6445	-0.2083 ± 0.1071	-1.5117 ± 0.2185	-1.3325 ± 0.1802
$a_3 = 0.662$	0.6559 ± 0.5383	-0.1644 ± 0.0511	0.8263 ± 0.2466	0.6562 ± 0.1503
$a_4 = -0.24$	-0.2413 ± 0.2016	0.0046 ± 0.0295	-0.3173 ± 0.1557	-0.2365 ± 0.0596
$\sigma_w^2 = 2.4$	2.3480 ± 0.2361	0.3662 ± 0.8487	2.4215 ± 0.0870	2.3942 ± 0.1348
$\sigma_e^2 = 1$	1.0540 ± 0.7885	4.4457 ± 1.0818	0.8827 ± 0.0466	1.0161 ± 0.2278
$RE(\%)$	0.3711	82.5522	13.3357	0.4449
$NRMSE(\%)$	47.6875	82.9949	22.8363	13.3178

Example 2. Consider a fourth-order noisy AR process with $\mathbf{a} = [1.352, -1.338, 0.662, -0.24]^T$ and $\sigma_e^2 = 1$. Two scenarios are studied here. In the first one, σ_w^2 is set to 0.38, giving rise to $SNR=10$ dB. Moreover, N is set to 100. Table III shows the means and standard deviations of the estimates of the AR parameters. Similar to the previous examples, the ESS has much better performance in this case. For the Xia-Zheng method, η , δ_1 , and δ_2 are respectively set to 0.7, 0.001, and 0.01.

In the second scenario, σ_w^2 is assumed to be 2.4, yielding $SNR=2$ dB. In addition, N is set to 4000. Table IV shows the means and standard deviations of the estimates of the AR parameters. Analogous to the previous scenarios, the performance of the ESS is much better than the other three methods. Note that η is set to 0.96, and both δ_1 and δ_2 are set to 0.05 here for the Xia-Zheng method.

Example 3. In this example, the performance of the aforementioned methods are compared for a more general

simulation setup to provide more insights into accuracy of parameter estimation by these methods. The parameters N and σ_e^2 are set to 4000 and 1, respectively. Additionally, η is set to 0.94, and both δ_1 and δ_2 are set to 0.05 here for the Xia-Zheng method. A hundred sets of poles, each set containing four poles selected randomly inside the unit circle, are considered as the poles of a hundred different AR processes. Then, for each AR process, the AR parameters are estimated via the aforementioned methods for 1000 trials, and NRMSE is calculated using (27). The mean of resultant NRMSEs of a hundred different AR signals is then taken as the total NRMSE in this simulation example. We repeat this scenario for different values of SNR that varies from 0 dB to 10 dB and plot the results in Fig. 1.

It can be observed from Fig. 1 that the proposed method shows more robust overall performance in dealing with various kind of AR processes in comparison with the IFILS and Xia and Zheng's methods for all SNRs. Note that since the performance of the SS method is very bad in this example, we omitted it. Moreover, since the selected value of η is not

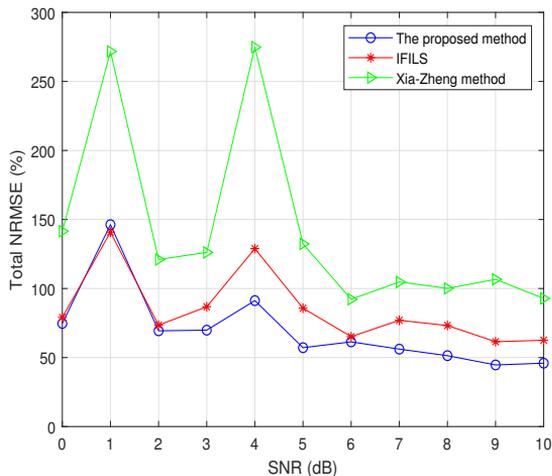


Fig. 1. Total RMSE versus SNR for Example 3.

appropriate for all of AR signals generated in this example, the Xia and Zheng's method does not have a good performance. Thus, this simulation example confirms the dependency of Xia and Zheng method's performance to the choice of initial value.

V. CONCLUSION

In this paper, a novel non-iterative subspace-based method was proposed for estimating the parameters of AR processes in the presence of white observation noise with unknown variance. The main notion of this method is to estimate the variance of the observation noise by solving a generalized eigenvalue problem as the first step, and then estimate the AR vector of parameters by finding the LS solution of a linear system of equations. The performance of the proposed method has been evaluated and compared with that of three other methods presented in the literature by means of simulations. The simulation results have demonstrated the superiority of the proposed method in terms of having smaller NRMSE, and also shown better robustness against high level of the observation noise. Moreover, the proposed method is non-iterative, and thus there are no convergence issues, unlike it is with the iterative-base methods.

REFERENCES

- [1] S. M. Kay, *Modern Spectral Estimation*. Prentice-Hall, Englewood Cliffs, NJ, 1988.
- [2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [3] M. D. Srinath, P. K. Rajasekaran, and R. Viswanathan, *Introduction to Statistical Signal Processing with Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [4] R. Sameni, "Online filtering using piecewise smoothness priors: Application to normal and abnormal electrocardiogram denoising", *Signal Processing*, vol. 133, pp. 52–63, 2017.
- [5] L. Legrand and E. Grivel, "Jeffrey's divergence between autoregressive processes distributed by additive white noise", *Signal Processing*, vol. 149, pp. 162–178, 2018.
- [6] 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 Processing & Control*, vol. 43, pp. 174–182, 2018.
- [7] S. Ganapathy, "Multirate autoregressive spectrogram modeling for noisy speech recognition", *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1373–1377, 2017.
- [8] K. Upadhyaya and S. A. Vorobyov, "An array processing approach to pilot decontamination for massive MIMO", in *Proc. 6th IEEE Int. Workshop Computational Advances in Multi-Sensor Adaptive Processing, CAMSAP'15*, Cancun, Mexico, Dec. 2015, pp. 465–468.
- [9] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems", *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 7, pp. 831–846, 2014.
- [10] A. Mahmoudi and M. Karimi, "Estimation of the parameters of multichannel autoregressive signals from noisy observations", *Signal Processing*, vol. 88, no. 11, pp. 2777–2783, 2008.
- [11] S. Maanan, B. Dumitrescu, and C. D. Giurcaneanu, "Conditional independence graphs for multivariate autoregressive models by convex optimization: Efficient algorithms", *Signal Processing*, vol. 133, pp. 122–134, 2017.
- [12] A. Amanat, A. Mahmoudi, and M. Hatam, "Two-dimensional noisy autoregressive estimation with application to joint frequency and direction of arrival estimation", *Multidim. Syst. Sign. Process.*, vol. 29, no. 2, pp. 671–685, 2018.
- [13] M. Kallas, P. Honeine, C. Francis, and H. Amond, "Kernel autoregressive models using Yule-Walker equations", *Multidim. Signal Processing*, vol. 93, no. 11, pp. 3053–3061, 2013.
- [14] S. M. Kay, "Noise compensation for autoregressive spectral estimates", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 3, pp. 292–303, 1980.
- [15] H. Tong, "Autoregressive model fitting with noisy data by Akaike's information criterion", *IEEE Trans. Inf. Theory*, vol. 21, no. 4, pp. 476–480, 1975.
- [16] A. Nehorai and P. Stoica, "Adaptive algorithms for constrained ARMA signals in the presence of noise", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 8, pp. 1282–1291, 1988.
- [17] W. X. Zheng, "On estimation of autoregressive signals in the presence of noise", *IEEE Trans. Circuits Syst. II*, vol. 53, no. 12, pp. 1471–1475, 2006.
- [18] Y. Xia and W. X. Zheng, "Novel parameter estimation of autoregressive signals in the presence of noise", *Automatica*, vol. 62, pp. 98–105, 2015.
- [19] W. X. Zheng, "Autoregressive parameter estimation from noisy data", *IEEE Trans. Circuits Syst. II: Analog Digital Signal Process.*, vol. 47, no. 1, pp. 71–75, 2000.
- [20] W. X. Zheng, "Fast identification of autoregressive signals from noisy observations", *IEEE Trans. Circuits Syst. II*, vol. 52, no. 1, pp. 43–48, 2005.
- [21] A. Mahmoudi and M. Karimi, "Inverse filtering based method for estimation of noisy autoregressive signals", *Signal Processing*, vol. 91, no. 7, pp. 1659–1664, 2011.
- [22] A. Mahmoudi and M. Karimi, "Parameter estimation of autoregressive signals from observations corrupted with colored noise", *Signal Processing*, vol. 90, pp. 157–164, 2010.
- [23] M. H. A. Davis and R. B. Vinter, *Stochastic Modeling and Control*. London, U.K.: Chapman and Hall, 1985.
- [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] C. Z. Jin, L. J. Jia, Z. J. Yang, and K. Wada, "On convergence of a BCLS algorithms for noisy autoregressive process estimation", in *Proc. 41st IEEE Conf. Decision and Control*, Las Vegas, NV, USA, Dec. 2002, pp. 4252–4257.
- [26] F. Tisseur and K. Meerbergen, "The quadratic eigenvalue problem", *SIAM Review*, vol. 43, no. 2, pp. 235–286, 2001.
- [27] R. Diversi, R. Guidorzi, and U. Soverini, "Identification of autoregressive models in the presence of additive noise", *Int. Journal of Adaptive Control and Signal Processing*, vol. 22, pp. 465–481, 2008.
- [28] R. Diversi, R. Guidorzi, and U. Soverini, "A noise-compensated estimation scheme for AR processes", in *Proc. 44th IEEE Conf. Decision and Control and 8-th European Control Conf.*, Melia Sevilla, Seville, Spain, Dec. 2005, pp. 4146–4151.

Publication 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.

© 2019

Reprinted with permission.

Non-Iterative Subspace-Based DOA Estimation in the Presence of Nonuniform Noise

Majdoddin Esfandiari ¹, Sergiy A. Vorobyov ², *Fellow, IEEE*, Simin Alibani ³, and Mahmood Karimi

Abstract—The uniform white noise assumption is one of the basic assumptions in most of the existing direction-of-arrival (DOA) estimation methods. In many applications, however, the nonuniform white noise model is more adequate. Then, the noise variances at different sensors have to be also estimated as nuisance parameters while estimating DOAs. In this letter, different from the existing iterative methods that address the problem of nonuniform noise, a non-iterative two-phase subspace-based DOA estimation method is proposed. The first phase of the method is based on estimating the noise subspace via eigendecomposition (ED) of some properly designed matrix and it avoids estimating the noise covariance matrix. In the second phase, the results achieved in the first phase are used to estimate the noise covariance matrix, followed by estimating the noise subspace via generalized ED. Since the proposed method estimates DOAs in a non-iterative manner, it is computationally more efficient and has no convergence issues as compared to the existing methods. Simulation results demonstrate better performance of the proposed method as compared to other existing state-of-the-art methods.

Index Terms—Array processing, direction-of-arrival (DOA) estimation, subspace-based methods, nonuniform noise, spectral analysis.

I. INTRODUCTION

DIRECTION of Arrival (DOA) and spectral estimation are the fundamental problems in array processing and spectral analysis with many applications in radar, sonar, navigation and communication systems, as well as acoustic tracking to mention just a few [1]–[5]. There exist several DOA estimation approaches. Among the most notable are the maximum likelihood (ML), beamforming-based, parametric subspace-based, and sparse representation-based approaches [6]–[10]. Parametric subspace-based DOA estimation methods, such as multiple signal classification (MUSIC) [6], [8] and estimation of signal parameters via rotational invariance technique (ESPRIT) [9] are well known to provide high accuracy and high resolution for DOA estimation with low computational complexity in comparison to the methods such as ML [6]. In addition to the traditional far-field narrowband signal assumption, the assumption

of uncorrelated sources is also critical for the former methods. A fundamental assumption that applies to all the aforementioned methods is however the presence of spatially uniform white noise. Under this assumption, the analytic concentration of the ML function with respect to the noise variance single parameter becomes possible, while parametric subspace-based methods are just built on this assumption since it enables separation of signal and noise subspaces [11], [12].

In diverse practical scenarios, the spatially uniform white noise assumption may be violated. Indeed, the sensor noise may be nonuniform [7], [13]–[17], spatially correlated [18], [19], or block-correlated [20]. Spatially white nonuniform noise arises when sensor noise powers are non-identical across the array, and leads to diagonal noise covariance matrix with non-identical entries. To overcome the problem of performance degradation in the presence of nonuniform noise, a variety of DOA estimation algorithms and techniques have been proposed. In [7] and [13], two deterministic and stochastic ML estimators are respectively proposed based on iterative procedures. These two estimators suffer from high computational complexity. Thus, two iterative subspace estimation algorithms with lower complexity, called iterative maximum likelihood subspace estimation (IMLSE) and iterative least squares subspace estimation (ILSSE), respectively, based on ML and least squares (LS) have been proposed in [14] for estimating signal subspace and noise covariance matrix. These algorithms then use spectral MUSIC method for performing the DOA estimation.

In this letter, we propose a new subspace-based method for DOA estimation in spatially nonuniform noise, which is non-iterative, thus leading to lower computational complexity and avoiding any convergence issues. The method has two phases. In the first phase, the noise subspace is initially estimated via eigendecomposition (ED) of some properly designed matrix without knowing the noise covariance matrix. In the second phase, the noise covariance matrix is then estimated by exploiting the results of the first phase and then the generalized ED is applied to the output array covariance matrix and noise covariance matrix. Simulation results demonstrate the efficiency and superiority of the proposed method in terms of both performance and complexity over the existing methods.

II. SIGNAL MODEL

Consider an array of M sensors receiving L ($L < M$ is known) independent narrowband signals impinging from the sources in far-field. The signal observed at time instance t by the array is given as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{A}(\boldsymbol{\theta}) \triangleq [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_L)]$ is the $M \times L$ matrix whose columns are the signal steering vectors $\mathbf{a}(\theta_i)$, $i = 1, \dots, L$, $\boldsymbol{\theta} \triangleq$

Manuscript received March 4, 2019; revised March 27, 2019; accepted April 1, 2019. Date of publication April 9, 2019; date of current version April 25, 2019. This work was supported in part by Academy of Finland under Research Grant 319822. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Monica F. Bugallo. (*Corresponding author: Sergiy A. Vorobyov.*)

M. Esfandiari and S. A. Vorobyov are with the Department Signal Processing and Acoustics, Aalto University, Espoo 00076, Finland (e-mail: majdoddin.esfandiari@aalto.fi; sergiy.vorobyov@aalto.fi).

S. Alibani and M. Karimi are with the School of Electrical and Computer Engineering, Shiraz University, Shiraz 71946-84471, Iran (e-mail: alibani.simin@gmail.com; karimi@shirazu.ac.ir).

Digital Object Identifier 10.1109/LSP.2019.2909587

1070-9908 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

$[\theta_1, \dots, \theta_L]^T$ is the vector of unknown source DOA's, $\mathbf{s}(t)$ is the $L \times 1$ vector of source signals, $\mathbf{n}(t)$ is the $M \times 1$ vector of zero-mean spatially and temporally white Gaussian noise, N is the number of snapshots, t is the discrete time index, and $[\cdot]^T$ denotes the transpose.

Using (1), the array output covariance matrix can be expressed as

$$\mathbf{R} \triangleq E\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q} \quad (2)$$

where $\mathbf{P} \triangleq E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$ and $\mathbf{Q} \triangleq E\{\mathbf{n}(t)\mathbf{n}^H(t)\}$ are respectively the $L \times L$ signal and $M \times M$ noise covariance matrices, and $E\{\cdot\}$ and $(\cdot)^H$ denote the expectation and Hermitian transpose operators, respectively. For uncorrelated sources, \mathbf{P} is a diagonal matrix. Under the nonuniform uncorrelated noise assumption, \mathbf{Q} is also a diagonal matrix of the form

$$\mathbf{Q} = \text{diag}\{\sigma_1^2, \dots, \sigma_M^2\} \quad (3)$$

where σ_m^2 , $m = 1, \dots, M$ are the sensor noise variances and $\text{diag}\{\cdot\}$ denotes a diagonal matrix.

III. NEW PROPOSED METHOD

In parametric subspace-based DOA estimation, the noise subspace needs to be first estimated. Under the nonuniform noise, also the noise covariance matrix possibly needs to be estimated. Then the orthogonality of the noise subspace basis vectors and the source steering vectors can be used to estimate the source DOA's.

To estimate the noise subspace, recall that in (2), $\mathbf{A}^H(\boldsymbol{\theta})$ is an $L \times M$ full row-rank matrix, and there are $M - L$ orthogonal vectors \mathbf{u}_l , $l = 1, \dots, M - L$ satisfying the following homogeneous equation

$$\mathbf{A}^H(\boldsymbol{\theta})\mathbf{u}_l = \mathbf{0} \quad (4)$$

where $\mathbf{0}$ is the vector of zeros. Multiplying both sides of (2) on the right by \mathbf{u}_l and using (4), we obtain

$$\mathbf{R}\mathbf{u}_l = \mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{u}_l + \mathbf{Q}\mathbf{u}_l = \mathbf{Q}\mathbf{u}_l, \quad l = 1, \dots, M - L. \quad (5)$$

According to (5), $M - L$ vectors \mathbf{u}_l , $l = 1, \dots, M - L$ which span the range space of the noise subspace are the solutions of the generalized ED problem for the matrices \mathbf{R} and \mathbf{Q} , while all $M - L$ eigenvalues being equal to one. However, since the matrix \mathbf{Q} is unknown, \mathbf{u}_l , $l = 1, \dots, M - L$ cannot be computed as simple as in the uniform noise case.

We observe, however, that the array output covariance matrix \mathbf{R} can be written as the following sum of two matrices

$$\mathbf{R} = \mathbf{R}_1 + \mathbf{R}_2 \quad (6)$$

where

$$[\mathbf{R}_1]_{i,j} = \begin{cases} [\mathbf{R}]_{i,j}, & i \neq j \\ 0, & i = j \end{cases} \quad (7)$$

and

$$\begin{aligned} \mathbf{R}_2 &= \text{diag}\left\{[\mathbf{R}]_{1,1}, \dots, [\mathbf{R}]_{M,M}\right\} \\ &= \text{diag}\left\{\sum_{k=1}^L s_k + \sigma_1^2, \dots, \sum_{k=1}^L s_k + \sigma_M^2\right\} \end{aligned} \quad (8)$$

with s_k being the received power of the k th source.

Substituting (6), (7), and (8) into (5), we obtain

$$\mathbf{R}_1\mathbf{u}_l = (\mathbf{Q} - \mathbf{R}_2)\mathbf{u}_l = -\left(\sum_{k=1}^L s_k\right)\mathbf{u}_l. \quad (9)$$

It can be seen from (9) that \mathbf{u}_l , $l = 1, \dots, M - L$ can be obtained by applying ED to the matrix \mathbf{R}_1 only. As a matter of fact, since adding the scaled identity matrix of the form $b \cdot \mathbf{I}$ to \mathbf{R}_1 does not alter eigenvectors and only shifts eigenvalues of \mathbf{R}_1 , the noise subspace basis vectors \mathbf{u}_l , $l = 1, \dots, M - L$ can be computed by applying ED to every matrix whose diagonal elements are identical and off-diagonal elements are equal to the off-diagonal elements of \mathbf{R} . Thus, after applying ED to \mathbf{R}_1 , the noise subspace basis vectors \mathbf{u}_l , $l = 1, \dots, M - L$, or in matrix form $\mathbf{U} \triangleq [\mathbf{u}_1, \dots, \mathbf{u}_{M-L}]$, can be obtained even without the need to estimate \mathbf{Q} . This novel result is stated and proved in the following lemma.

Lemma 1: The noise subspace basis vectors \mathbf{u}_l , $l = 1, \dots, M - L$ are the $M - L$ eigenvectors of \mathbf{R}_1 whose corresponding eigenvalues are the smallest.

Proof: Assume that there exists an $M \times 1$ vector \mathbf{d} that satisfies the following conditions

$$\mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{d} \neq \mathbf{0} \quad (10)$$

$$\mathbf{R}_1\mathbf{d} = \lambda\mathbf{d}. \quad (11)$$

Adding $\mathbf{R}_2\mathbf{d}$ to both sides of (11) and using (6), we obtain

$$\mathbf{R}_1\mathbf{d} + \mathbf{R}_2\mathbf{d} = \mathbf{R}\mathbf{d} = \lambda\mathbf{d} + \mathbf{R}_2\mathbf{d}. \quad (12)$$

Inserting (2) into (12), and rearranging the terms, we have

$$\mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{d} = \lambda\mathbf{d} + (\mathbf{R}_2 - \mathbf{Q})\mathbf{d}. \quad (13)$$

Moreover, substituting (3) and (8) into (13), we obtain

$$\mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{d} = \left(\lambda + \sum_{k=1}^L s_k\right)\mathbf{d}. \quad (14)$$

It can be seen from (14) that \mathbf{d} is an eigenvector of the matrix $\mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})$ while its corresponding eigenvalue is equal to $\lambda + \sum_{k=1}^L s_k$. Since the matrix $\mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})$ has L positive eigenvalues and the condition (10) has to be satisfied, it can be concluded that

$$\lambda + \sum_{k=1}^L s_k > 0 \quad \Rightarrow \quad \lambda > -\sum_{k=1}^L s_k. \quad (15)$$

In other words, (15) indicates that $-\sum_{k=1}^L s_k$ is the lower bound on the smallest eigenvalue of \mathbf{R}_1 . Thus, the noise subspace basis \mathbf{u}_l , $l = 1, \dots, M - L$ is composed of $M - L$ eigenvectors of \mathbf{R}_1 with the smallest eigenvalues. ■

Knowing \mathbf{U} , the spectral-MUSIC method, for example, can be used for the source DOA's estimation by finding the locations of L peaks in the pseudo-spectrum

$$\mathbf{S}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\mathbf{U}\mathbf{U}^H\mathbf{a}(\theta)}. \quad (16)$$

However, the estimate of \mathbf{U} can be further improved using the results of the initial \mathbf{U} estimation and exploiting (5). Indeed, if there exists an estimate of \mathbf{Q} , (5) can be solved by applying generalized ED to the matrices \mathbf{R} and \mathbf{Q} . Then more accurate noise subspace basis vectors \mathbf{u}_l , $l = 1, \dots, M - L$ can be found as stated in the following lemma.

Lemma 2: The noise subspace basis vectors \mathbf{u}_l , $l = 1, \dots, M - L$ are the $M - L$ eigenvectors, obtained by applying generalized eigendecomposition to the matrices \mathbf{R} and \mathbf{Q} whose corresponding eigenvalues are the smallest.

Proof: Similar to the proof of Lemma 1, assume that there exists an $M \times 1$ vector \mathbf{d} that satisfies (10) and $\mathbf{R}\mathbf{d} = \lambda\mathbf{Q}\mathbf{d}$.

Inserting (2) into $\mathbf{R}\mathbf{d} = \lambda\mathbf{Q}\mathbf{d}$, and rearranging the terms, we have

$$\mathbf{A}(\theta)\mathbf{P}\mathbf{A}^H(\theta)\mathbf{d} = (\lambda - 1)\mathbf{Q}\mathbf{d}. \quad (17)$$

Multiplying both sides of (17) on the left by \mathbf{d}^H , we obtain

$$\mathbf{d}^H\mathbf{A}(\theta)\mathbf{P}\mathbf{A}^H(\theta)\mathbf{d} = \bar{\mathbf{d}}^H\bar{\mathbf{P}}\bar{\mathbf{d}} = (\lambda - 1)\mathbf{d}^H\mathbf{Q}\mathbf{d} \quad (18)$$

where $\bar{\mathbf{d}} = \mathbf{A}^H(\theta)\mathbf{d}$. Since both \mathbf{P} and \mathbf{Q} are positive definite matrices and the condition (10) has to be satisfied, it can be concluded that

$$\lambda - 1 > 0 \Rightarrow \lambda > 1. \quad (19)$$

In other words, (19) indicates that 1 is the lower bound on the smallest eigenvalue of generalized ED of \mathbf{R} and \mathbf{Q} . Thus, the noise subspace basis \mathbf{u}_l , $l = 1, \dots, M - L$ is composed of $M - L$ eigenvectors with the smallest eigenvalues. ■

To apply Lemma 2, we first need to find an estimate of \mathbf{Q} . Towards this end, first we write \mathbf{Q} as the following sum of two diagonal matrices

$$\mathbf{Q} = \sigma^2\mathbf{I} + \mathbf{Q}_{\text{num}} \quad (20)$$

where σ^2 represents the common part of sensor noise powers, which is computed later, and \mathbf{Q}_{num} is a diagonal matrix whose diagonal elements, except for one of them, are nonzero. The place of this zero element is the place of the smallest diagonal element of \mathbf{R} . As a result, the rank of \mathbf{Q}_{num} is $M - 1$, and

$$\mathbf{e}_k^T\mathbf{Q}_{\text{num}} = \mathbf{0} \quad (21)$$

where \mathbf{e}_k is the $M \times 1$ unit vector such that

$$[\mathbf{e}_k]_i = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases} \quad (22)$$

and k is the index of the smallest diagonal element of \mathbf{R} .

Multiplying both sides of (5) by \mathbf{e}_k^T on the left and using (20) and (21), we obtain

$$\mathbf{e}_k^T\mathbf{R}\mathbf{u}_l = \mathbf{e}_k^T(\sigma^2\mathbf{I} + \mathbf{Q}_{\text{num}})\mathbf{u}_l = \sigma^2\mathbf{e}_k^T\mathbf{u}_l. \quad (23)$$

Equation (23) can be written for all vectors \mathbf{u}_l , $l = 1, \dots, M - L$ in the following matrix-vector form

$$\mathbf{e}_k^T\mathbf{R}\mathbf{U} = \sigma^2\mathbf{e}_k^T\mathbf{U} \quad (24)$$

where \mathbf{U} is composed of \mathbf{u}_l , $l = 1, \dots, M - L$ obtained using (9). Consequently, σ^2 can be computed as

$$\sigma^2 = \left| \frac{\mathbf{e}_k^T\mathbf{R}\mathbf{U}\mathbf{U}^H\mathbf{e}_k}{\mathbf{e}_k^T\mathbf{U}\mathbf{U}^H\mathbf{e}_k} \right| \quad (25)$$

where $|\cdot|$ denotes the absolute value operator. The only issue remaining is the construction of the matrix \mathbf{Q}_{num} . Let us set the nonzero diagonal elements of \mathbf{Q}_{num} as the differences of the corresponding elements in \mathbf{R} with the smallest diagonal element of \mathbf{R} , that is,

$$\mathbf{Q}_{\text{num}} = \text{diag}\left\{[\mathbf{R}]_{1,1} - c, \dots, [\mathbf{R}]_{M,M} - c\right\} \quad (26)$$

where c is the smallest diagonal element of \mathbf{R} .

Finally, the matrix \mathbf{Q} can be formed by utilizing (20), (25), and (26). With the matrices \mathbf{R} and \mathbf{Q} , the noise subspace basis \mathbf{u}_l , $l = 1, \dots, M - L$ can be re-estimated as stated in (5) and Lemma 2. Then the new \mathbf{U} can be formed and the source DOA's can be estimated by finding, for example, the locations of L peaks in (16). The corresponding algorithm for DOA estimation in nonuniform noise is summarized in Algorithm 1, where the sample data covariance matrix $\hat{\mathbf{R}}$ is used as an estimate of the array output covariance matrix \mathbf{R} . Steps 1 and 2 represent the

Algorithm 1: The proposed method.

- 1: Compute the sample covariance matrix $\hat{\mathbf{R}} = \frac{1}{N} \sum_{t=1}^N \mathbf{X}(t)\mathbf{X}^H(t)$.
 - 2: Form $\hat{\mathbf{R}}_1$ from $\hat{\mathbf{R}}$ as in (7), carry out the ED of $\hat{\mathbf{R}}_1$ to obtain the noise subspace basis $\hat{\mathbf{u}}_l$, $l = 1, \dots, M - L$, and construct the matrix $\hat{\mathbf{U}}$.
 - 3: Construct \mathbf{e}_k and $\hat{\mathbf{Q}}_{\text{num}}$ according to (21), (22), and (26).
 - 4: Using the data sample covariance matrix and $\hat{\mathbf{U}}$ obtained in step 2, estimate $\hat{\sigma}^2$ according to (25).
 - 5: Using, $\hat{\sigma}^2$ and $\hat{\mathbf{Q}}_{\text{num}}$, estimate $\hat{\mathbf{Q}}$ according to (20).
 - 6: Apply generalized ED to $\hat{\mathbf{R}}$ and $\hat{\mathbf{Q}}$, and obtain the new estimate of the noise subspace basis $\hat{\mathbf{u}}_l$, $l = 1, \dots, M - L$, i.e., the new estimate of $\hat{\mathbf{U}}$.
 - 7: Use spectral-MUSIC, i.e., find the locations of L peaks in (16), where \mathbf{U} is substituted its estimate $\hat{\mathbf{U}}$.
-

first phase of the algorithm that can be followed by step 7 directly. Steps 3–6 represent the second *correction* phase of the algorithm.

Complexity analysis: For the proposed method, the ED of \mathbf{R}_1 or the generalized ED of \mathbf{R} and \mathbf{Q} are involved. The corresponding complexity is $O(M^3)$ [21]. It is equivalent to the complexity in each iteration of IMLSE or ILSSE [14]. The difference arises from the fact that IMLSE and ILSSE are iterative methods (the number of iterations can be comparable to M in many scenarios to converge to their best result).

IV. SIMULATION RESULTS

A ULA with $M = 8$ omnidirectional sensors, which are separated by half wavelength, is considered. Two far-field uncorrelated narrowband signals (with equal powers) impinge on the array simultaneously from $\theta_1 = -3^\circ$ and $\theta_2 = 6^\circ$, respectively.

The worst noise power ratio (WNPR), and the signal-to-noise ratio (SNR) are defined as $\text{WNPR} = \sigma_{\text{max}}^2/\sigma_{\text{min}}^2$ and $\text{SNR} = \sigma_s^2/M \sum_{m=1}^M (\sigma_m^2)^{-1}$, respectively, where σ_{max}^2 and σ_{min}^2 are the maximum and minimum sensors noise powers, respectively, and σ_s^2 is the signal power. The number of snapshots (N) and the number of Monte Carlo trials (K) are set to 500 and 5000, respectively. The root mean squared error (RMSE) of DOA estimation is defined as

$$\text{RMSE} = \sqrt{\frac{1}{KL} \sum_{k=1}^K \sum_{l=1}^L (\hat{\theta}_{k,l} - \theta_l)^2}$$

where $\hat{\theta}_{k,l}$ denotes the l th DOA estimate in the k th trial. To validate the performance of the proposed methods, two examples are considered, and the results are compared to the performance of the standard spectral-MUSIC as well as IMLSE and ILSSE methods both after the first iteration only and also after convergence is achieved.¹ The number of sources is assumed to be known for all methods tested.

Example 1: The nonuniform noise covariance matrix is fixed in all simulation runs and is given as $\mathbf{Q} = \text{diag}\{1, 1, 1, 1, 1, 20, 30, 50\}$, resulting in $\text{WNPR} = 50$. Fig. 1 shows the RMSEs for the methods tested versus SNR. The Cramer-Rao bound (CRB) [7] is also shown. It can be seen from

¹The performance of ILSSE is plotted only for the SNRs when ILSSE converges for all trials.

TABLE I
COMPARISON OF AVERAGE RUN TIME OF ONE TRIAL (IN *ms*)

Method	IMLSE	ILSSE	IMLSE after first iteration	ILSSE after first iteration	Proposed method (phase 1)	Proposed method (phase 2)
SNR=-5 dB	13.108	Not converged	0.573	0.478	0.371	0.891
SNR=0 dB	14.395	21.739	0.552	0.470	0.368	0.877
SNR=5 dB	15.133	23.870	0.546	0.462	0.359	0.861
SNR=10 dB	15.470	26.667	0.545	0.459	0.357	0.860
SNR=15 dB	15.604	26.385	0.547	0.459	0.357	0.861

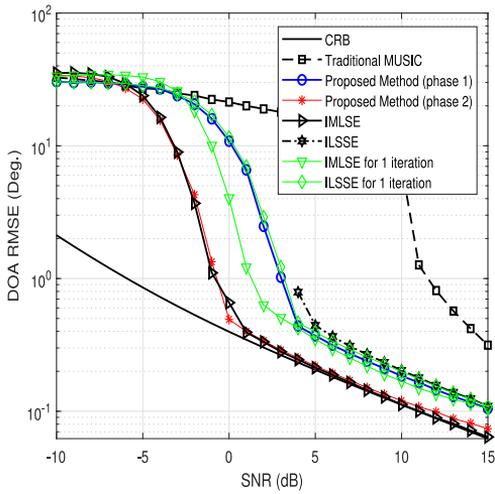


Fig. 1. The RMSEs of DOA estimation versus SNR in Example 1.

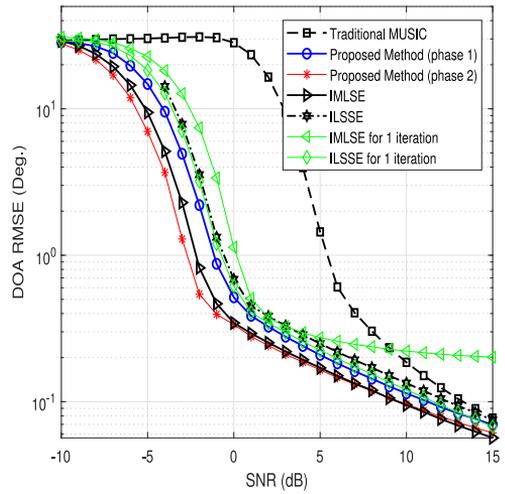


Fig. 2. The RMSEs of DOA estimation versus SNR in Example 2.

the figure that the proposed method after both phases demonstrates the best threshold behavior, despite it is non-iterative and has lower computational complexity. Although the IMLSE method after the first iteration only possesses as low computational complexity as the proposed method, its DOA estimation accuracy is very poor. Thus, multiple iterations have to run for it to achieve its best result.

Example 2: To provide more comprehensive insights into the performance of the methods tested, the noise covariance matrix is chosen in this example multiple times randomly with maximum WNPR of 30. Then the RMSE results are also averaged over 50 different realizations of the noise covariance matrix \mathbf{Q} for which of each 5000 Monte Carlo trials are averaged. The average RMSEs for the methods tested are shown in Fig. 2.² It can be seen from the figure that the second phase of the proposed method shows the best performance and improves the threshold behavior by about 2 dB as compared to the next best performing method that is the IMLSE method.

Furthermore, Table I shows the average run time of the methods tested for SNRs = -5, 0, 5, 10, and 15 dB for the setup of Example 2. The simulation is performed on a PC running an Intel(R) Xeon(R) 3.40GHz CPU. It can be observed that the proposed method is superior in terms of the required time which is

reduced by orders of magnitude compared to the existing methods. The average number of iterations for IMLSE and ILSSE are about 27 and 52, respectively, for SNR = 5 dB, for example. Inspecting Fig. 2 and Table I together, it can be seen that even the first phase of the proposed method leads to superior performance with lower complexity compared to the competitive methods after the first iteration.

V. CONCLUSION

A novel computationally efficient non-iterative two-phase subspace-based parametric method for DOA estimation in the presence of unknown spatially nonuniform noise has been proposed. The noise subspace estimation problem is converted in the first phase of the proposed method to the problem of finding eigenvectors of a properly designed matrix so that the noise covariance matrix estimation is avoided. In the second phase, the covariance matrix of the nonuniform noise is first estimated based on the results of the first phase and then it is used for finding the noise subspace more accurately by means of generalized ED of this matrix and the data covariance matrix. It is of importance that the proposed method has low computational complexity and is non-iterative, and thus, has no issues with convergence. Moreover, it is superior to the existing iterative state-of-the-art methods in both the performance and especially the computational cost.

²Since the curves in Fig. 2 are resulted also from averaging over different realization of the nuisance parameters, i.e., \mathbf{Q} 's with different WNPRs, the CRB is not applicable in the setup and is not shown. Indeed, fixed parameters of interest as well as nuisance parameters have to be assumed for CRB.

REFERENCES

- [1] H. L. Van Trees, *Optimum Array Processing*. New York, NY, USA: Wiley, 2002.
- [2] A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds., *Array and Statistical Signal Processing* (Academic Press Library in Signal Processing), vol. 3. New York, NY, USA: Academic, 2014.
- [3] B. Friedlander and T. Tuncer, *Classical and Modern Direction-of-Arrival Estimation*. New York, NY, USA: Academic, 2009.
- [4] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1997.
- [5] Y. Wang, J. Li, and P. Stoica, *Spectral Analysis of Signals*. San Rafael, CA, USA: Morgan and Claypool, 2005.
- [6] 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.
- [7] M. Pesavento and A. 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.
- [8] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.
- [9] 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.
- [10] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.
- [11] M. Shaghghi 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.
- [12] M. W. Morency, S. A. Vorobyov, and G. Leus, "Joint detection and localization of an unknown number of sources using algebraic structure of the noise subspace," *IEEE Trans. Signal Process.*, vol. 66, no. 17, pp. 4685–4700, Sep. 2018.
- [13] C. Chen, F. Lorenzelli, R. 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.
- [14] B. Liao, S. 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.
- [15] 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 Sensors J.*, vol. 16, no. 24, pp. 8982–8989, Dec. 2016.
- [16] B. Liao, G. S. Liao, and J. Wen, "A method for DOA estimation in the presence of unknown nonuniform noise," *J. Electromagn. Waves Appl.*, vol. 22, no. 14/15, pp. 2113–2123, 2008.
- [17] B. Liao and S. C. Chan, "A simple method for DOA estimation in the presence of unknown nonuniform noise," in *Proc. IEEE Conf. Acoust., Speech, Signal Process.*, 2015, pp. 2789–2793.
- [18] 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 Trans. Signal Process.*, vol. 48, no. 10, pp. 2743–2749, Oct. 2000.
- [19] M. Malek-Mohammadi, M. Jansson, A. Owrang, A. Koochakzadeh, and M. Babaie-Zadeh, "DOA estimation in partially correlated noise using low-rank/sparse matrix decomposition," in *Proc. IEEE Workshop Sensor Array Multichannel Signal Process.*, 2014, pp. 373–376.
- [20] 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.
- [21] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.

Publication 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.

© 2021

Reprinted with permission.

ENHANCED STANDARD ESPRIT FOR OVERCOMING IMPERFECTIONS IN DOA ESTIMATION

Majdoddin Esfandiari and Sergiy A. Vorobyov

Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

ABSTRACT

Direction-of-arrival (DOA) estimation problem is a challenging one in the presence of coherent sources, when the sample size is small, and the signal-to-noise ratio is low. We address this problem by developing a new method called *enhanced standard ESPRIT* (ES ESPRIT), and also its unitary extension called *enhanced unitary ESPRIT* (EU ESPRIT). The proposed methods use statistics of the subspace perturbation. First, they generate $2K$ DOA candidates for K sources, and then discreetly select K of them. Numerical results show the superiority of EU ESPRIT over other existing methods especially in improving threshold performance and separating closely located sources with a small sample size.

Index Terms— DOA estimation, ESPRIT, generalized least squares (GLS), small sample size

1. INTRODUCTION

Direction-of-arrival (DOA) and spectral estimation are problems of significant importance in many applications including radar, sonar, navigation and wireless communication, to mention just a few [1]–[2]. The most notable DOA estimation methods are the maximum likelihood, subspace-based, and sparse representation-based approaches [3]–[4]. Due to such advantages as low complexity and high estimation precision, ESPRIT-type algorithms [5], [6] have been wildly exploited. The essence of ESPRIT-type algorithms is to find the solution of a linear system of equations, known as shift invariance equation (SIE), which is a function of the signal subspace. For solving the SIE, the methods of least squares (LS) [5] or total least squares (TLS) [7] are usually used. However, such solutions of the SIE suffer from not considering the statistics of the subspace perturbation, resulting in being suboptimal. To alleviate this problem, generalized least squares (GLS) [8], [9] has been employed in a number of ESPRIT-based methods [10]–[11], where the covariance matrix of the subspace estimation error is included as the weighting matrix. In [10] and [12], a first-order performance analysis [13] has been used for optimizing the weighting matrix. The disadvantage of [10] is that the weighting matrix is optimized only for one of the DOAs, while in the case of [12], no

proof was presented for the resulting formulation. In [14], an iterative GLS-based solution to the SIE of the standard and unitary ESPRIT that takes into account the statistics of the signal subspace perturbation has been also proposed.

Continuing on the developments in [14], we propose in this paper a new ESPRIT-type method by first transforming the SIE of the standard and unitary ESPRIT to the discrete Fourier transform (DFT) domain, and then multiplying the resulting equation by a selection matrix. The aforementioned procedure is performed twice for two selection matrices to generate $2K$ DOA candidates for K sources. Afterwards, two DOA selection strategies are employed for extracting the final K DOAs. Simulation results demonstrate that the proposed methods outperform the existing methods.

2. SIGNAL MODEL

Consider K narrowband signals emitted by K sources located in the far field and received by a uniform linear array (ULA) of M ($M > K$) sensors. The received signal after collecting N time snapshots can be expressed as

$$\mathbf{X} = \mathbf{A}(\boldsymbol{\mu})\mathbf{S} + \mathbf{N} \in \mathbb{C}^{M \times N} \quad (1)$$

where $\mathbf{A}(\boldsymbol{\mu}) \triangleq [\mathbf{a}(\mu_1), \dots, \mathbf{a}(\mu_K)] \in \mathbb{C}^{M \times K}$ is the array steering matrix whose columns are the corresponding array response vectors for the spatial frequencies $\boldsymbol{\mu} \triangleq [\mu_1, \dots, \mu_K]^T$. Note that each $\mu_i, i = 1, \dots, K$ is related to the corresponding θ_i . In addition, $\mathbf{S} \in \mathbb{C}^{K \times N}$ is the source signals matrix and $\mathbf{N} \in \mathbb{C}^{M \times N}$ is the sensor noise matrix.

The essence of the ESPRIT-type algorithms is to estimate the desired spatial frequency by exploiting a property of the array steering matrix \mathbf{A} known as *the shift invariance property*. The SIE is given as $\mathbf{J}_1 \mathbf{A} \boldsymbol{\Phi} = \mathbf{J}_2 \mathbf{A}$, where $\mathbf{J}_1 = [\mathbf{I}_{M-1} \ \mathbf{0}_M]$ and $\mathbf{J}_2 = [\mathbf{0}_M \ \mathbf{I}_{M-1}]$ pick the first and the last $M - 1$ rows of \mathbf{A} , respectively. Furthermore, the diagonal matrix $\boldsymbol{\Phi}$ is related to the spatial frequencies as $\boldsymbol{\Phi} = \text{diag}\{e^{j\mu_i}\}_{i=1}^K$. Conducting the truncated singular value decomposition (TSVD), \mathbf{X} can be decomposed as $\mathbf{X} = \hat{\mathbf{U}}_s \hat{\boldsymbol{\Sigma}}_s \hat{\mathbf{V}}_s^H$, where the columns of $\hat{\mathbf{U}}_s \in \mathbb{C}^{M \times K}$ and $\hat{\mathbf{V}}_s \in \mathbb{C}^{N \times K}$ are respectively the left and the right singular vectors. Moreover, $\hat{\boldsymbol{\Sigma}}_s \in \mathbb{R}^{K \times K}$ is a diagonal matrix whose diagonal elements are K principal singular values of \mathbf{X} .

This work was supported in part by Academy of Finland under Research Grant 319822.

Since \mathbf{A} is unknown, instead of working with the aforementioned direct expression of the SIE, the fact that both \mathbf{A} and $\hat{\mathbf{U}}_s$ share the same range space can be exploited to write $\mathbf{A} \approx \hat{\mathbf{U}}_s \mathbf{T}$, where $\mathbf{T} \in \mathbb{C}^{K \times K}$ is a full-rank matrix. Thus, the SIE can be written as

$$\mathbf{J}_1 \hat{\mathbf{U}}_s \Psi \approx \mathbf{J}_2 \hat{\mathbf{U}}_s \quad (2)$$

where $\Psi \approx \mathbf{T} \Phi \mathbf{T}^{-1}$ is a diagonalisable matrix such that its eigenvalues λ_i , $i = 1, \dots, K$ are related to the spatial frequencies via $\lambda_i = e^{j\mu_i}$. Based on this observation, the main objective of the ESPRIT-type algorithms boils down to estimating Ψ , and then extracting μ_i 's from its eigenvalues.

3. ENHANCED GLS FOR ESPRIT

We first transform the SIE into the DFT domain to generate multiple estimates of Ψ by selecting different equations of transformed SIE (TSIE). In doing so, the statistics of the subspace estimation errors are taken into account that leads to finding more precise solution of TSIE. Considering the noiseless case, (2) with respect to the actual signal subspace \mathbf{U}_s is represented by

$$\mathbf{J}_1 \mathbf{U}_s \Psi = \mathbf{J}_2 \mathbf{U}_s. \quad (3)$$

Applying DFT to (3), we obtain

$$\mathbf{W}_D \mathbf{J}_1 \mathbf{U}_s \Psi = \mathbf{W}_D \mathbf{J}_2 \mathbf{U}_s \quad (4)$$

where $\mathbf{W}_D \in \mathbb{C}^{M-1 \times M-1}$ is the DFT matrix, i.e.,

$$\mathbf{W}_D = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{M-2} & \dots & \omega^{(M-2)(M-2)} \end{bmatrix} \quad (5)$$

with $\omega = e^{\frac{j2\pi}{M-1}}$. Multiplying both sides of (4) by the selecting matrices $\mathbf{Z}_{\mathcal{I}} \in \mathbb{R}^{|\mathcal{I}| \times M-1}$, which enables us to exploit diverse subsets of TSIE for estimating Ψ , we obtain

$$\mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_1 \mathbf{U}_s \Psi = \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_2 \mathbf{U}_s \quad (6)$$

where \mathcal{I} denotes the set containing the indices of the selected equations and $|\mathcal{I}|$ denotes the cardinality of this set, i.e., the number of its members. Thus, the entries of the i th row of $\mathbf{Z}_{\mathcal{I}}$ are all zero except one whose index is determined by the i th member of \mathcal{I} . Since $\mathbf{Z}_{\mathcal{I}}$ is a selecting matrix, the values of its nonzero elements are set to be 1.

The important problem that should be addressed now is determining the method of selecting equations in (6). We propose to pick the equations based on the absolute value of DFT of $\mathbf{J}_1 \mathbf{u}_1$, where \mathbf{u}_1 is the left singular vector of \mathbf{X} which corresponds to the largest singular value. Thus, the members of \mathcal{I} are opted to be the corresponding indices of those elements of $\mathbf{W}_D \mathbf{J}_1 \mathbf{u}_1$ with $|\mathcal{I}|$ largest absolute values. Using diverse

\mathcal{I} 's with varied number of members leads to finding multiple estimates for Ψ , and consequently for μ_i 's as well. Albeit it will be elaborated later, it is worth mentioning here that we propose to generate two estimates of Ψ by selecting $M-1$ and $M-2$ equations of (6), i.e., setting $|\mathcal{I}|$ to $M-1$ and $M-2$. Note that $\mathbf{Z}_{\mathcal{I}} = \mathbf{I}_{M-1}$ in the former case.

Taking into account the estimation error, \mathbf{U}_s can be expressed as $\mathbf{U}_s = \hat{\mathbf{U}}_s + \Delta \mathbf{U}_s$, where $\Delta \mathbf{U}_s$ is the subspace estimation error. Subsequently, (6) can be rewritten as

$$\mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_1 (\hat{\mathbf{U}}_s + \Delta \mathbf{U}_s) \Psi = \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_2 (\hat{\mathbf{U}}_s + \Delta \mathbf{U}_s). \quad (7)$$

Expanding and rearranging (7), we obtain

$$\mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_1 \hat{\mathbf{U}}_s \Psi + \mathbf{E} = \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_2 \hat{\mathbf{U}}_s \quad (8)$$

where $\mathbf{E} \triangleq \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_1 \Delta \mathbf{U}_s \Psi - \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_2 \Delta \mathbf{U}_s$ denotes the error matrix imposed by the observation noise as a function of $\Delta \mathbf{U}_s$. Applying the vectorization operator to (8), we have

$$\hat{\mathbf{f}} \triangleq \text{vec}\{\mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_2 \hat{\mathbf{U}}_s\} = \hat{\mathbf{F}} \psi + \mathbf{e} = \hat{\mathbf{F}} \psi + \hat{\mathbf{G}} \Delta \mathbf{u}_s \quad (9)$$

where $\hat{\mathbf{F}} \triangleq \mathbf{I}_K \otimes \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_1 \hat{\mathbf{U}}_s \in \mathbb{C}^{|\mathcal{I}|K \times K^2}$, and $\hat{\mathbf{G}} \triangleq (\Psi^T \otimes \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_1) - (\mathbf{I}_K \otimes \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \mathbf{J}_2) \in \mathbb{C}^{|\mathcal{I}|K \times MK}$ thanks to the property that $\text{vec}\{\mathbf{ABC}\} = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}\{\mathbf{B}\}$. In addition, $\psi \triangleq \text{vec}\{\Psi\} \in \mathbb{C}^{K^2 \times 1}$ and $\Delta \mathbf{u}_s \triangleq \text{vec}\{\Delta \mathbf{U}_s\} \in \mathbb{C}^{MK \times 1}$. Afterwards, the GLS [8], [9], [14] can be used to find the optimal solution of (9), by taking the second-order statistic of $\hat{\mathbf{U}}_s$ into account. In doing so, the GLS optimization problem is expressed as

$$\hat{\psi}_{GLS} = \arg \min_{\psi} (\hat{\mathbf{f}} - \hat{\mathbf{F}} \psi)^H \mathbf{W} (\hat{\mathbf{f}} - \hat{\mathbf{F}} \psi) \quad (10)$$

where

$$\begin{aligned} \mathbf{W} &= [\mathbb{E}\{\mathbf{e}\mathbf{e}^H\}]^{-1} = [\mathbb{E}\{\hat{\mathbf{G}} \Delta \mathbf{u}_s \Delta \mathbf{u}_s^H \hat{\mathbf{G}}^H\}]^{-1} \\ &= [\hat{\mathbf{G}} \mathbb{E}\{\Delta \mathbf{u}_s \Delta \mathbf{u}_s^H\} \hat{\mathbf{G}}^H]^{-1} = [\hat{\mathbf{G}} \mathbf{Q} \hat{\mathbf{G}}^H]^{-1} \end{aligned} \quad (11)$$

with $\mathbf{Q} \triangleq \mathbb{E}\{\Delta \mathbf{u}_s \Delta \mathbf{u}_s^H\} \in \mathbb{C}^{MK \times MK}$ being the covariance matrix of signal subspace perturbation. Subsequently, the optimal solution of (10) is given by

$$\hat{\psi}_{GLS} = (\hat{\mathbf{F}}^H \mathbf{W} \hat{\mathbf{F}})^{-1} \hat{\mathbf{F}}^H \mathbf{W} \hat{\mathbf{f}}. \quad (12)$$

As it can be seen from (11), we need second-order statistic of $\Delta \mathbf{u}_s$ for computing \mathbf{Q} . Using the first-order perturbation expansion of SVD [15], [16] of \mathbf{X} in (1) (when the noise is absent), it can be written that

$$\Delta \mathbf{U}_s \approx (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H) \mathbf{N} \mathbf{V}_s \Sigma_s^{-1}. \quad (13)$$

Therefore, by exploiting (13), \mathbf{Q} can be computed as

$$\begin{aligned} \mathbf{Q} &= \mathbb{E}\{\text{vec}\{\Delta \mathbf{U}_s\} \text{vec}\{\Delta \mathbf{U}_s\}^H\} \\ &\approx (\Sigma_s^{-1} \mathbf{V}_s^T \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H)) \mathbb{E}\{\mathbf{nn}^H\} \\ &\quad \times (\mathbf{V}_s^* \Sigma_s^{-1} \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H)) \end{aligned} \quad (14)$$

where $\mathbf{n} \triangleq \text{vec}\{\mathbf{N}\} \in \mathbb{C}^{MK \times 1}$ is assumed to be zero-mean white random noise vector, resulting in $\mathbb{E}\{\mathbf{nn}^H\} = \sigma^2 \mathbf{I}_{MN}$. Inserting this result into (14), we have

$$\mathbf{Q} = \sigma^2 (\boldsymbol{\Sigma}_s^{-2} \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H)) \quad (15)$$

thanks to the property $(\mathbf{A} \otimes \mathbf{B}) \cdot (\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$. It is worth noting that (15) exactly matches the result reported in [14] since the same procedure for computing \mathbf{Q} has been used. It is not surprising though, since the dependency of $\Delta \mathbf{u}_s$ to \mathbf{n} is exploited to compute \mathbf{Q} in both GLS ESPRIT [14] and the proposed method. The differences arise from choosing various matrices and vectors for $\hat{\mathbf{F}}$, $\hat{\mathbf{G}}$, and $\hat{\mathbf{f}}$, yielding different $\hat{\mathbf{W}}$ as well as different $\hat{\boldsymbol{\psi}}$.

From (11) and (15), it is clear that $\hat{\mathbf{G}}\mathbf{Q}\hat{\mathbf{G}}^H$ is a singular matrix for $|\mathcal{I}| > M - K$. Thus, $\hat{\mathbf{Q}} = \sigma^2 (\boldsymbol{\Sigma}_s^{-2} \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H + \lambda \mathbf{I}_M))$ can be used in (11) instead of \mathbf{Q} with $\lambda \neq \{0, 1\}$ being the regularization parameter. Hence, \mathbf{W} can be replaced by $\tilde{\mathbf{W}} = (\hat{\mathbf{G}}\hat{\mathbf{Q}}\hat{\mathbf{G}}^H)^{-1}$ in (10) and (12), which is a full rank matrix. Based on the same discussion as in [14] (see Theorem 1), it is straight-forward to show that any choice of $\lambda \in \mathbb{R} - \{0, 1\}$ leads to the same solution for $\hat{\boldsymbol{\psi}}_{GLS}$, if $\tilde{\mathbf{W}}$ is used instead of \mathbf{W} . Therefore, for the sake of convenience, $\tilde{\mathbf{W}}$ can be replaced by $\hat{\mathbf{W}}$ as

$$\begin{aligned} \tilde{\mathbf{W}} &= \lim_{\lambda \rightarrow \infty} \left[\hat{\mathbf{G}} (\boldsymbol{\Sigma}_s^{-2} \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H + \lambda \mathbf{I}_M)) \hat{\mathbf{G}}^H \right]^{-1} \\ &= \left[\hat{\mathbf{G}} (\boldsymbol{\Sigma}_s^{-2} \otimes \mathbf{I}_M) \hat{\mathbf{G}}^H \right]^{-1}. \end{aligned} \quad (16)$$

Note that σ^2 is omitted in (16) as it cancels out in computing $\hat{\boldsymbol{\psi}}_{GLS}$. After initializing the algorithm by $\hat{\boldsymbol{\Psi}}_{LS} = (\mathbf{Z}_T \mathbf{W}_D \mathbf{J}_1 \hat{\mathbf{U}}_s)^\dagger \mathbf{Z}_T \mathbf{W}_D \mathbf{J}_2 \hat{\mathbf{U}}_s$, $\boldsymbol{\Sigma}_s$ is replaced by its estimate $\hat{\boldsymbol{\Sigma}}_s$ to construct $\hat{\tilde{\mathbf{W}}} = \left[\hat{\mathbf{G}} (\hat{\boldsymbol{\Sigma}}_s^{-2} \otimes \mathbf{I}_M) \hat{\mathbf{G}}^H \right]^{-1}$. Inserting this result into (12), we have

$$\hat{\boldsymbol{\psi}}_{GLS} = \left(\hat{\mathbf{F}}^H \hat{\tilde{\mathbf{W}}} \hat{\mathbf{F}} \right)^{-1} \hat{\mathbf{F}}^H \hat{\tilde{\mathbf{W}}} \hat{\mathbf{f}} \quad (17)$$

and consequently, $\hat{\boldsymbol{\Psi}}_{GLS} = \text{unvec}\{\hat{\boldsymbol{\psi}}_{GLS}\} \in \mathbb{C}^{K \times K}$, where $\text{unvec}\{\cdot\}$ forms a matrix from a vector, i.e., it is an inverse of $\text{vec}\{\cdot\}$. Finally, by computing the K eigenvalues $\hat{\lambda}_i$ of $\hat{\boldsymbol{\Psi}}_{GLS}$, the desired special frequencies are obtained as $\hat{\mu}_i = \arg\{\hat{\lambda}_i\}$.

Remark 1: To obtain sufficiently precise estimate for $\hat{\boldsymbol{\psi}}_{GLS}$, the use of an iterative estimation procedure, where $\hat{\tilde{\mathbf{W}}}$ is initialized as in (17) with $\hat{\boldsymbol{\Psi}}_{LS}$ followed by updating $\hat{\boldsymbol{\Psi}}_{GLS}$ and $\hat{\tilde{\mathbf{W}}}$ via alternation, is natural. Based on our simulation results, the sufficient number of iterations for achieving the accurate estimate is 5.

Remark 2: The proposed ESPRIT-based method can be extended for the case of unitary-ESPRIT formulation [6]. In doing so, it is required to modify (6) as

$$\mathbf{Z}_T \mathbf{W}_D \mathbf{K}_1 \mathbf{E}_s \boldsymbol{\Upsilon} = \mathbf{Z}_T \mathbf{W}_D \mathbf{K}_2 \mathbf{E}_s \quad (18)$$

where $\mathbf{K}_1 \triangleq 2 \cdot \text{Re}\{\mathbf{Q}_{M-1}^H \mathbf{J}_2 \mathbf{Q}_M\} \in \mathbb{R}^{(M-1) \times M}$, $\mathbf{K}_2 \triangleq 2 \cdot \text{Im}\{\mathbf{Q}_{M-1}^H \mathbf{J}_2 \mathbf{Q}_M\} \in \mathbb{R}^{(M-1) \times M}$, and \mathbf{E}_s is the matrix of K principal left singular vectors of $\varphi(\tilde{\mathbf{X}}) = \mathbf{Q}_M^H \tilde{\mathbf{X}} \mathbf{Q}_{2N} \in \mathbb{R}^{M \times 2N}$ with \mathbf{Q}_M and \mathbf{Q}_{2N} being left Π -real matrices [6]. In addition, $\tilde{\mathbf{X}} \triangleq [\mathbf{X} \quad \Pi_M \mathbf{X}^* \Pi_N] \in \mathbb{C}^{M \times 2N}$, where $\Pi_M \in \mathbb{R}^{M \times M}$ is the exchange matrix with ones on the anti-diagonal and zeros elsewhere [6]. Following the same steps presented for ES ESPRIT, the corresponding real-valued solution of (18) can be determined as

$$\hat{\boldsymbol{\nu}}_{GLS} = \left(\hat{\mathbf{F}}^H \hat{\tilde{\mathbf{W}}}_1 \hat{\mathbf{F}} \right)^{-1} \hat{\mathbf{F}}^H \hat{\tilde{\mathbf{W}}}_1 \hat{\mathbf{f}} \quad (19)$$

where $\hat{\boldsymbol{\nu}}_{GLS} \triangleq \text{vec}\{\hat{\boldsymbol{\Upsilon}}_{GLS}\}$, $\hat{\mathbf{F}} \triangleq (\mathbf{I}_K \otimes \mathbf{Z}_T \mathbf{W}_D \mathbf{K}_1 \hat{\mathbf{E}}_s)$, $\hat{\mathbf{f}} \triangleq \text{vec}\{\mathbf{Z}_T \mathbf{W}_D \mathbf{K}_2 \hat{\mathbf{E}}_s\}$, and $\hat{\tilde{\mathbf{W}}}_1 \triangleq \left[\hat{\mathbf{G}} (\hat{\boldsymbol{\Sigma}}_s^{-2} \otimes \mathbf{I}_M) \hat{\mathbf{G}}^H \right]^{-1}$ with $\hat{\mathbf{G}} \triangleq (\hat{\boldsymbol{\Upsilon}}^T \otimes \mathbf{Z}_T \mathbf{W}_D \mathbf{K}_1) - (\mathbf{I}_K \otimes \mathbf{Z}_T \mathbf{W}_D \mathbf{K}_2)$. Additionally, $\hat{\boldsymbol{\Sigma}}_s \in \mathbb{R}^{K \times K}$ is a diagonal matrix containing the nonzero singular values of $\varphi(\tilde{\mathbf{X}})$, i.e., $\varphi(\tilde{\mathbf{X}})$ can be decomposed via TSVD as $\varphi(\tilde{\mathbf{X}}) = \hat{\mathbf{E}}_s \hat{\boldsymbol{\Sigma}}_s \hat{\mathbf{D}}_s^H$. Similar to the steps presented for ESE, enhanced unitary ESPRIT (EUE) begins with initializing $\hat{\boldsymbol{\Upsilon}}$ via $\hat{\boldsymbol{\Upsilon}}_{LS} = (\mathbf{Z}_T \mathbf{W}_D \mathbf{K}_1 \hat{\mathbf{E}}_s)^\dagger \mathbf{Z}_T \mathbf{W}_D \mathbf{K}_2 \hat{\mathbf{E}}_s$ followed by updating $\hat{\boldsymbol{\nu}}_{GLS}$ and $\hat{\tilde{\mathbf{W}}}_1$ via alternation (5 iterations as suggested for ESE as well). Notice that the spatial frequency estimates $\hat{\mu}_i$, $i = 1, \dots, K$ are extracted from the eigenvalues $\hat{\omega}_i$ of the final estimate of $\hat{\boldsymbol{\Upsilon}}_{GLS} = \text{unvec}\{\hat{\boldsymbol{\nu}}_{GLS}\}$ (i.e., after 5 iterations) by $\hat{\mu}_i = 2 \cdot \arctan\{\hat{\omega}_i\}$.

4. DOA SELECTION STRATEGY

In order to obtain better DOA estimates, we propose to employ ES ESPRIT (EU ESPRIT) twice with $|\mathcal{I}| = M - 1$ and $|\mathcal{I}| = M - 2$, respectively, to produce $2K$ DOA candidates. These choices of cardinality generate the most precise DOA candidates. Then, we employ two selection strategies to pick up K DOA estimates from $2K$ previously estimated DOAs. Following [17], [18], [19], the first selection strategy is using the deterministic ML (DML) cost function [20] to determine the final DOAs. In doing so, $2K$ DOAs are divided into $P_{ESE(EUE)} = \frac{2K!}{K!K!}$ different subsets, each subset Θ_i , $i = 1, \dots, P_{ESE(EUE)}$ containing K different DOAs, which generate $\mathbf{A}(\Theta_i)$, $i = 1, \dots, P_{ESE(EUE)}$. The final DOAs are the members of the subset that minimizes the DML cost function, i.e.,

$$\begin{aligned} \hat{\Theta}_{DML} &= \arg \min_{\Theta_i} \text{tr}((\mathbf{I}_M - \mathbf{A}(\Theta_i))(\mathbf{A}(\Theta_i)^H \mathbf{A}(\Theta_i))^{-1} \\ &\quad \times \mathbf{A}(\Theta_i)^H) \hat{\mathbf{R}}), \quad \forall i = 1, \dots, P_{ESE(EUE)} \end{aligned} \quad (20)$$

where $\hat{\mathbf{R}} = \frac{1}{N} \mathbf{X} \mathbf{X}^H$ is the sample covariance matrix (SCM), and $\text{tr}(\cdot)$ denotes the trace operator. As it can be concluded from (20), the use of DML as the selection strategy imposes computing $P_{ESE(EUE)}$ inverse matrices of dimension $K \times K$, which is of a considerably high computational burden.

However, notice that $P_{ESE(EUE)} = \frac{2K!}{K!K!}$ and $P_{EPUMA} = \frac{(2K+1)!}{(K+1)!K!}$ [19], which makes $P_{ESE(EUE)}$ approximately half of P_{EPUMA} , resulting in fifty percent computational saving via using ESE(EUE) in tandem with DML compared to enhanced principal-singular-vector utilization for modal analysis (EPUMA) [19].

For the second selection strategy, we propose to use a method that we call *sequential likelihood ratio* (SLR). Using the method presented in [21], the final K DOAs are selected sequentially one by one from the estimated $2K$ DOA candidates. Based on SLR, for determining the l th DOA ($l = 1, \dots, K$), we select θ_i from the set of $2K - (l - 1)$ remaining DOAs that maximizes the stochastic maximum likelihood (SML), i.e.,

$$\hat{\theta}_l = \arg \max_{\theta_i} \frac{\mathbf{a}(\theta_i)^H \mathbf{P}_{l-1}^\perp \hat{\mathbf{R}} \mathbf{P}_{l-1}^\perp \mathbf{a}(\theta_i)}{\mathbf{a}(\theta_i)^H \mathbf{P}_{l-1}^\perp \mathbf{a}(\theta_i)}$$

for $i = 1, \dots, 2K - (l - 1)$, $l = 1, \dots, K$ (21)

where

$$\mathbf{P}_{l-1}^\perp \triangleq \begin{cases} \mathbf{I}_M - \mathbf{A}_{l-1}(\mathbf{A}_{l-1}^H \mathbf{A}_{l-1})^{-1} \mathbf{A}_{l-1}^H & l > 1 \\ \mathbf{I}_M & l = 1 \end{cases} \quad (22)$$

with $\mathbf{A}_{l-1} \triangleq [\mathbf{a}(\hat{\theta}_1), \dots, \mathbf{a}(\hat{\theta}_{l-1})] \in \mathbb{C}^{M \times (l-1)}$. Selecting the l th DOA via SLR imposes computing $2K - (l - 1)$ inverse matrices of dimension $(l - 1) \times (l - 1)$. In the case of $l = K$, it is required to compute $K + 1$ inverse matrices of dimension $(K - 1) \times (K - 1)$ for selecting the K th DOA. As a result, using SLR requires much smaller number of arithmetic computations rather than using DML, especially when K is relatively large.

5. SIMULATION RESULT

The performance of the standard ESPRIT using GLS (SE GLS), unitary ESPRIT using GLS (UE GLS) [14], EPUMA [19], and the proposed ES ESPRIT, and EU ESPRIT is evaluated. The deterministic Cramer-Rao bound (CRB) [20] is included as a benchmark. Two examples are considered. In both examples, we assume that a ULA with $M = 16$ isotropic sensors separated by half-wavelength spacing receives the signals. The number of snapshots is $N = 4$. In addition, the sensor noise is drawn from zero-mean circularly symmetric white Gaussian distribution and the curves are averaged over 10000 Monte Carlo trials. Fig. 1 displays the root mean square error (RMSE) versus the SNR for $K = 2$ uncorrelated signals with $\Theta = [5^\circ, 7^\circ]$. As it can be observed from Fig. 1, the best performance is shown by EU ESPRIT and UE GLS.

In Fig. 2, the total RMSE is depicted versus the SNR considering the case of receiving $K = 2$ highly correlated signals with correlation coefficient of $\rho = 0.95$ from $\Theta = [32^\circ, 33^\circ]$. It can be seen from Fig. 2 that the best performance is obtained by EU ESPRIT and EPUMA (using

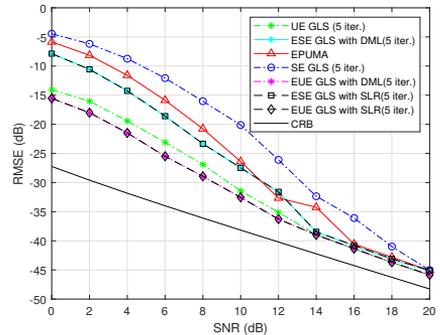


Fig. 1. RMSEs vs. SNR for $K = 2$ uncorrelated signals with $\Theta = [5^\circ, 7^\circ]$, $M = 16$ and $N = 4$.

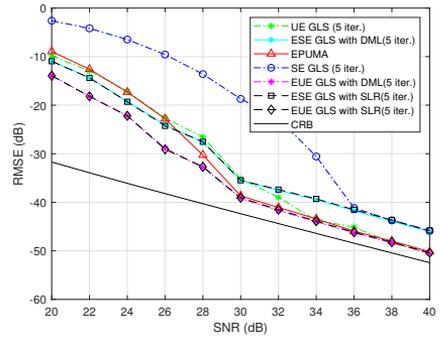


Fig. 2. RMSEs vs. SNR for $K = 2$ correlated signals ($\rho = 0.95$) with $\Theta = [32^\circ, 33^\circ]$, $M = 16$ and $N = 4$.

forward-backward averaging). According to Figs. 1 and 2, it can be concluded that EU ESPRIT has the best performance among the considered methods for both uncorrelated and correlated sources. Moreover, because the use of both DML and SLR as the DOA selection strategy results in the same performance based on the figures, it is more beneficial to employ SLR for selecting the final DOAs due to its much smaller computational burden.

6. CONCLUSION

A novel ESPRIT-type method called ES (EU for the case of unitary formulation) ESPRIT has been developed for estimating DOAs in the presence of coherent sources, small sample size and low SNR. It gives reliable performance when the available sample size is small and/or some of the DOAs are close to each other. Computer simulations have shown that the proposed method outperforms other existing method.

7. REFERENCES

- [1] T. E. Tuncer and B. Friedlander, *Classical and modern direction-of-arrival estimation*, Academic Press, 2009.
- [2] P. Stoica, R. L. Moses, et al., "Spectral analysis of signals," 2005.
- [3] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [4] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.
- [5] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.
- [6] M. Haardt and J. A. Nossek, "Unitary esprit: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1232–1242, 1995.
- [7] B. Ottersten, M. Viberg, and T. Kailath, "Performance analysis of the total least squares esprit algorithm," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1122–1135, 1991.
- [8] T. Amemiya, "Generalized least squares theory," *Advanced Econometrics*, 1985.
- [9] T. Kariya and H. Kurata, *Generalized Least Squares*, John Wiley & Sons, 2004.
- [10] P. Stoica and A. Nehorai, "Performance comparison of subspace rotation and music methods for direction estimation," *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 446–453, 1991.
- [11] C. Qian, L. Huang, H. C. So, N. D. Sidiropoulos, and J. Xie, "Unitary puma algorithm for estimating the frequency of a complex sinusoid," *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5358–5368, 2015.
- [12] A. Eriksson and P. Stoica, "Optimally weighted esprit for direction estimation," *Signal Processing*, vol. 38, no. 2, pp. 223–229, 1994.
- [13] B. D. Rao and K. V. S. Hari, "Performance analysis of esprit and tam in determining the direction of arrival of plane waves in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 1990–1995, 1989.
- [14] J. Steinwandt, F. Roemer, and M. Haardt, "Generalized least squares for esprit-type direction of arrival estimation," *IEEE Signal Processing Letters*, vol. 24, no. 11, pp. 1681–1685, 2017.
- [15] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for doa estimation algorithms: unification, simplification, and observations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1170–1184, 1993.
- [16] J. Steinwandt, F. Roemer, M. Haardt, and G. Del Galdo, "R-dimensional esprit-type algorithms for strictly second-order non-circular sources and their performance analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, pp. 4824–4838, 2014.
- [17] M. Shaghghi and S. A. Vorobyov, "Subspace leakage analysis and improved doa estimation with small sample size," *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3251–3265, 2015.
- [18] A. B. Gershman and P. Stoica, "New mode-based techniques for direction finding with an improved threshold performance," *Signal Processing*, vol. 76, no. 3, pp. 221–235, 1999.
- [19] C. Qian, L. Huang, N. D. Sidiropoulos, and H. C. So, "Enhanced puma for direction-of-arrival estimation and its performance analysis," *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4127–4137, 2016.
- [20] P. Stoica and A. Nehorai, "Music, maximum likelihood, and cramer-rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.
- [21] F. Izedi, M. Karimi, and M. Derakhshan, "Joint doa estimation and source number detection for arrays with arbitrary geometry," *Signal Processing*, vol. 140, pp. 149–160, 2017.

Publication 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.

© 2022

Reprinted with permission.

A NOVEL ANGULAR ESTIMATION METHOD IN THE PRESENCE OF NONUNIFORM NOISE

Majdoddin Esfandiari and Sergiy A. Vorobyov

Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

ABSTRACT

A novel algorithm for direction-of-arrival (DOA) estimation in nonuniform sensor noise is developed. The diagonal nonuniform sensor noise covariance matrix is estimated by an iterative procedure, which only requires a few iterations. Using the generalized eigendecomposition of two matrices and the least squares, the noise subspace is refined and the noise covariance matrix is estimated iteratively. Since there is no need for knowledge of true DOAs when estimating the noise covariance matrix, our method is superior to most existing approaches. For the proposed noise covariance estimator, we also derive the asymptotic variance of one iteration. Numerical simulations are carried out to demonstrate the advantages of the proposed algorithm over existing state-of-the-art methods.

Index Terms— DOA estimation, subspace methods, nonuniform noise.

1. INTRODUCTION

The problem of finding direction-of-arrivals (DOAs) of desired signals arises in diverse practical applications such as wireless communication, automotive radar and sonar [1–6], to name a few. Although multiple rigorous approaches have been developed to tackle the DOA estimation problem in the presence of white uniform sensor noise [7–13], the often more practical assumptions of white nonuniform sensor noise and unknown noise field have drawn attention rather recently [14, 15]. It is worth mentioning that most of the methods designed for the former case are not applicable readily to the later cases. Therefore, devising proper methods which take into consideration the presence of nonuniform sensor noise is vital for many applications.

Several methods designed for the case of white nonuniform sensor noise have been presented recently in the literature. Particularly, the deterministic maximum likelihood (ML) estimator and corresponding Cramer-Rao bound (CRB) for both deterministic and stochastic signals have been derived in [14], while the stochastic ML estimator has been developed in [16]. A simple method has been devised in [17]

that requires less computational cost than ML estimator and improves the DOA estimation accuracy. Moreover, via exploiting the ML and least squares (LS) criteria, two iterative methods referred to as iterative ML subspace estimation (IMLSE) and iterative least squares subspace estimation (ILSSE) have been developed in [18]. The aim of both the IMLSE and ILSSE is to estimate the signal subspace and noise covariance matrices first, followed by finding DOAs by identifying peaks of the multiple signal classification (MUSIC) pseudo-spectrum. In addition, it has been shown in [19] that the signal and noise subspaces can be separated by applying the eigendecomposition (ED) of the reduced covariance matrix (RCM) when sources are uncorrelated. In [19], the authors have developed also a rank minimization based approach for coping with the case of correlated sources. An efficient method referred to as non-iterative subspace-based (NISB) method has been developed recently in [20]. It achieves both high performance and low computational complexity. The essence of the NISB method is to find a proper estimate of the noise covariance matrix by employing the ED of the RCM [19], followed by identifying the noise subspace via applying the generalized eigendecomposition (GED) of the matrix pair of the sample covariance matrix (SCM) and the estimated noise covariance matrix.

In this paper, we develop a new method to address the problem of DOA estimation in the presence of white nonuniform sensor noise. The core of our method is to estimate the noise subspace and noise covariance matrix iteratively. Specifically, the noise subspace and noise covariance matrix are refined by exploiting the GED of two matrices and using the LS criteria, respectively, in iterative manner. A study of the asymptotic variance of such estimator is also conducted. Since we do not need to know the actual DOAs to estimate the noise covariance matrix, our approach can be considered superior to the majority of existing methods. The proposed noise covariance matrix estimator works efficiently for sensor arrays with arbitrary geometries and also in the presence of correlated sources. Our method is also appealing computationally because it only requires a few iterations to obtain a proper noise covariance matrix estimate. Following the computation of the noise subspace, the MUSIC framework is employed to identify DOAs. Simulation results are included to demonstrate the effectiveness of the proposed method.

This work was supported in part by Academy of Finland under Research Grant 319822.

2. SIGNAL MODEL

Consider a uniform linear array (ULA) with M omnidirectional sensors receiving L ($L < M$) independent narrowband signals emitted by L sources. The sources are located in the far-field at distinguished directions, denoted by θ_l , $l = 1, \dots, L$. Then, the signal received by the sensor array at the time instant t is given as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{s}(t) \triangleq [s_1(t) \dots s_L(t)]^T \in \mathbb{C}^L$ is the vector of source signals, $\mathbf{n}(t) \in \mathbb{C}^M$ denotes the sensor noise vector, $\boldsymbol{\theta} \triangleq [\theta_1 \dots \theta_L]^T$ contains the source DOAs, and $\mathbf{A}(\boldsymbol{\theta}) \triangleq [\mathbf{a}(\theta_1) \dots \mathbf{a}(\theta_L)] \in \mathbb{C}^{M \times L}$ with the steering vector $\mathbf{a}(\theta_l) = [1 \ e^{-j2\pi \sin(\theta_l)d/\lambda} \ \dots \ e^{-j2\pi(M-1)\sin(\theta_l)d/\lambda}]^T \in \mathbb{C}^M$ corresponding to l th DOA. Also here λ stands for the carrier wavelength and $d = \lambda/2$. For notation simplicity, \mathbf{A} is used instead of $\mathbf{A}(\boldsymbol{\theta})$ hereafter.

The array covariance matrix can be written as

$$\mathbf{R} \triangleq E\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q} \quad (2)$$

where $\mathbf{P} \in \mathbb{C}^{L \times L}$ and $\mathbf{Q} \in \mathbb{R}^{M \times M}$, respectively, denote the signal and noise covariance matrices defined as

$$\mathbf{P} \triangleq E\{\mathbf{s}(t)\mathbf{s}^H(t)\}, \quad \mathbf{Q} \triangleq E\{\mathbf{n}(t)\mathbf{n}^H(t)\}. \quad (3)$$

Considering the case of spatially and temporally uncorrelated nonuniform sensor noise that is zero-mean Gaussian, the noise covariance matrix can be written as

$$\mathbf{Q} = \text{diag}\{\sigma_1^2, \dots, \sigma_M^2\} \quad (4)$$

where $\text{diag}\{\cdot\}$ stands for a diagonal matrix generated by plugging the entries of the bracketed argument into its main diagonal. In (4), σ_m^2 , $m = 1, \dots, M$ are the noise variances, which are considered to be nonidentical, i.e., $\sigma_i^2 \neq \sigma_j^2$ for $i \neq j$. When $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_M^2 = \sigma^2$, the noise covariance matrix is just a scaled identity matrix $\mathbf{Q} = \sigma^2 \mathbf{I}_M$, i.e., the sensor noise is uniform. Whereas the latter case has been probed to a great extent in the literature, the former has been received more attention only in the recent years.

Because \mathbf{R} is unknown in practice, the SCM is typically used, and it is given by

$$\hat{\mathbf{R}} \triangleq \frac{1}{N} \sum_{t=1}^N \mathbf{x}(t)\mathbf{x}^H(t) = \frac{1}{N} \mathbf{X}\mathbf{X}^H. \quad (5)$$

Here, the matrix signal notation is also used, that is,

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (6)$$

with $\mathbf{X} \triangleq [\mathbf{x}(1) \dots \mathbf{x}(N)]$, $\mathbf{S} \triangleq [\mathbf{s}(1) \dots \mathbf{s}(N)]$, $\mathbf{N} \triangleq [\mathbf{n}(1) \dots \mathbf{n}(N)]$, and N being the number of snapshots.

3. NOISE COVARIANCE MATRIX ESTIMATION

It is desirable for an algorithm estimating \mathbf{Q} that it would not require any knowledge of the true DOAs while providing an acceptable accuracy with an affordable computational cost. Furthermore, such algorithm should provide reliable results in extreme scenarios. The presence of closely located sources and small sample size are two examples of extreme scenarios. To develop such an algorithm, both sides of (2) are multiplied by $\mathbf{U} \in \mathbb{C}^{M \times (M-L)}$ which satisfies the following condition

$$\mathbf{A}^H \mathbf{U} = \mathbf{0}_{L \times (M-L)} \quad (7)$$

where the constraint $\mathbf{U}^H \mathbf{U} = \mathbf{I}_{M-L}$ is also imposed to avoid ambiguities in obtaining \mathbf{U} . It can be seen from (7) that the columns of \mathbf{U} span the noise subspace. For the case of nonuniform noise, finding \mathbf{U} is not as simple as for the uniform noise case where \mathbf{U} is obtained by just calculating the eigenvectors of $\hat{\mathbf{R}}$.

Multiplying (2) by \mathbf{U} and using (7), we find that [20]

$$\hat{\mathbf{R}}\mathbf{U} = \mathbf{Q}\mathbf{U} \quad (8)$$

where \mathbf{R} is replaced by $\hat{\mathbf{R}}$. Knowing \mathbf{Q} , it can be observed from (8) that the columns of the best estimate of \mathbf{U} , denoted as $\hat{\mathbf{U}}$, are calculated as the $M-L$ eigenvectors corresponding to the $M-L$ smallest eigenvalues obtained after applying the GED to the pair of matrices $\{\hat{\mathbf{R}}, \mathbf{Q}\}$.

Moreover, according to (8), \mathbf{Q} can be estimated by the LS approach. Towards this end, we formulate the following LS minimization problem with respect to \mathbf{Q}

$$\hat{\mathbf{Q}} = \arg \min_{\mathbf{Q}} \|(\hat{\mathbf{R}} - \mathbf{Q})\hat{\mathbf{U}}\|_{\mathbb{F}}^2 \quad (9)$$

where \mathbf{U} is replaced by $\hat{\mathbf{U}}$ and $\|\cdot\|_{\mathbb{F}}$ denotes the the Frobenius norm of a matrix. Problem (9) needs to be solved subject to the constraint of \mathbf{Q} being a diagonal matrix. The objective function of (9) can be rewritten as

$$\begin{aligned} f(\mathbf{Q}) &\triangleq \|(\hat{\mathbf{R}} - \mathbf{Q})\hat{\mathbf{U}}\|_{\mathbb{F}}^2 \\ &= \text{trace} \left\{ \left((\hat{\mathbf{R}} - \mathbf{Q})\hat{\mathbf{U}} \right) \left((\hat{\mathbf{R}} - \mathbf{Q})\hat{\mathbf{U}} \right)^H \right\} \\ &= \text{trace} \left\{ \hat{\mathbf{U}}\hat{\mathbf{U}}^H \hat{\mathbf{R}}^2 \right\} - \text{trace} \left\{ \hat{\mathbf{R}}\hat{\mathbf{U}}\hat{\mathbf{U}}^H \mathbf{Q} \right\} \\ &\quad - \text{trace} \left\{ \hat{\mathbf{U}}\hat{\mathbf{U}}^H \hat{\mathbf{R}}\mathbf{Q} \right\} + \text{trace} \left\{ \hat{\mathbf{U}}\hat{\mathbf{U}}^H \mathbf{Q}^2 \right\} \end{aligned} \quad (10)$$

where $\text{trace}\{\cdot\}$ stands for the trace of a square matrix, and the properties $\|\mathbf{X}\|_{\mathbb{F}}^2 = \text{trace}\{\mathbf{X}\mathbf{X}^H\}$, $\text{trace}\{\mathbf{X}\mathbf{Y}\} = \text{trace}\{\mathbf{Y}\mathbf{X}\}$, $\hat{\mathbf{R}} = \hat{\mathbf{R}}^H$, and $\mathbf{Q} = \mathbf{Q}^H$ are used. The partial derivative of (10) with respect to \mathbf{Q} is obtained as

$$\frac{\partial f(\mathbf{Q})}{\partial \mathbf{Q}} = 2\mathcal{D} \left\{ \hat{\mathbf{U}}\hat{\mathbf{U}}^H \right\} \mathbf{Q} - \mathcal{D} \left\{ \hat{\mathbf{R}}\hat{\mathbf{U}}\hat{\mathbf{U}}^H + \hat{\mathbf{U}}\hat{\mathbf{U}}^H \hat{\mathbf{R}} \right\} \quad (11)$$

where the operator $\mathcal{D}\{\cdot\}$ generates a diagonal matrix by preserving the main diagonal of the bracketed matrix and setting all other entries to zero (see details in [21]). Equating (11) to zero, the optimal estimate of \mathbf{Q} is obtained as

$$\hat{\mathbf{Q}} = \frac{1}{2} \mathcal{D} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}} \hat{\mathbf{U}}^H + \hat{\mathbf{U}} \hat{\mathbf{U}}^H \hat{\mathbf{R}} \right\} \mathcal{D} \left\{ \hat{\mathbf{U}} \hat{\mathbf{U}}^H \right\}^{-1}. \quad (12)$$

Because of the dependencies of calculating $\hat{\mathbf{Q}}$ and $\hat{\mathbf{U}}$ in (8) and (12), it is natural to use an iterative scheme for estimating $\hat{\mathbf{Q}}$ and $\hat{\mathbf{U}}$. It starts with properly initializing $\hat{\mathbf{Q}}$, denoted as $\hat{\mathbf{Q}}^{(0)}$. Then $\hat{\mathbf{U}}^{(0)}$ is estimated as the $M - L$ generalized eigenvectors of the pair $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}^{(0)}\}$ corresponding to the $M - L$ smallest eigenvalues. Next, $\hat{\mathbf{Q}}^{(1)}$ is obtained via (12) after replacing $\hat{\mathbf{U}}$ with $\hat{\mathbf{U}}^{(0)}$. The alternations carry on until a predefined stopping criterion is satisfied. It is worth noting that any diagonal matrix with positive diagonal entries can be used as $\hat{\mathbf{Q}}^{(0)}$, however, we suggest to employ $\hat{\mathbf{Q}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$.

Algorithm 1: Noise Covariance Matrix Estimation

- 1: Compute $\hat{\mathbf{R}} = 1/N \sum_{t=1}^N \mathbf{x}(t) \mathbf{x}^H(t)$.
 - 2: Set $i = 0$, $\hat{\mathbf{Q}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$, and the maximum number of iterations $i_{max} = 5$.
 - while** $i \leq i_{max}$
 - 3: Carry out the GED of the pair $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}^{(i)}\}$ to obtain $\hat{\mathbf{U}}^{(i)}$ as the $M - L$ eigenvectors corresponding to the $M - L$ smallest eigenvalues.
 - 4: Calculate $\hat{\mathbf{Q}}^{(i+1)}$ using (12).
 - 5: set $i = i + 1$.
 - end**
-

The steps of the proposed algorithm for the noise covariance matrix estimation are outlined in Algorithm 1. Carrying out the GED of the pair $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}\}$,¹ the noise subspace is obtained as the $M - L$ eigenvectors associated with the $M - L$ smallest eigenvalues. Although any subspace-based method can be adopted for DOA estimation, we use the spectral MUSIC method via finding the locations of L peaks in the following pseudo-spectrum

$$\mathbf{S}_{MU}(\theta) = \frac{1}{\mathbf{a}^H(\theta) \hat{\mathbf{U}} \hat{\mathbf{U}}^H \mathbf{a}(\theta)}. \quad (13)$$

The asymptotic mean square error (MSE) of the proposed noise covariance estimation method in (12) is given in the following proposition.

Proposition 1: The asymptotic variance of estimating each diagonal element of $\hat{\mathbf{Q}}$ in (12), given a particular $\hat{\mathbf{U}}$, is

$$\begin{aligned} \mathbb{E} \left\{ (\Delta \sigma_m^2)^2 \right\} &= \frac{1}{2N\tau_m^2} \left(\Re \left\{ [\mathbf{R}]_{mm} \mathbf{v}_m^H \mathbf{R} \mathbf{v}_m \right\} \right. \\ &\quad \left. + \Re \left\{ \mathbf{v}_m^H (\mathbf{r}_m^T \otimes \mathbf{R}) \mathbf{K} (\mathbf{d}_m \otimes \mathbf{I}_M) \mathbf{v}_m^* \right\} \right) \end{aligned} \quad (14)$$

¹If there exist correlated sources, $\hat{\mathbf{R}}_{\text{FB}} = \frac{1}{2} (\hat{\mathbf{R}} + \mathbf{J}_M \hat{\mathbf{R}}^* \mathbf{J}_M)$ is preferred over $\hat{\mathbf{R}}$ where \mathbf{J}_M is the exchange matrix.

where $\tau_m \triangleq [\hat{\mathbf{U}} \hat{\mathbf{U}}^H]_{mm}$ is the entry in the intersection of the i th row and j th column of the matrix $\hat{\mathbf{U}} \hat{\mathbf{U}}^H$, $\mathbf{v}_m \triangleq [\hat{\mathbf{U}} \hat{\mathbf{U}}^H]_{:,m}$ is the m th column of the aforementioned matrix, $\Re\{\cdot\}$ returns the real part of the bracketed argument, \mathbf{r}_m denotes the m th column of \mathbf{R} , \otimes denotes the Kronecker product, \mathbf{K} is the commutation matrix and $\mathbf{d}_m \in \mathbb{R}^M$ is a vector with 1 on the m th position and 0 elsewhere.

Proof: Using (12), the m th diagonal entry of $\hat{\mathbf{Q}}$, denoted by $\hat{\sigma}_m^2$, can be written as

$$\hat{\sigma}_m^2 = \frac{(\mathbf{v}_m^H \hat{\mathbf{r}}_m + (\mathbf{v}_m^H \hat{\mathbf{r}}_m)^H)}{2\tau_m} = \frac{\Re\{\mathbf{v}_m^H \hat{\mathbf{r}}_m\}}{\tau_m} \quad (15)$$

where $\hat{\mathbf{r}}_m$ denotes the m th column of $\hat{\mathbf{R}}$. Expressing $\hat{\mathbf{r}}_m$ as $\hat{\mathbf{r}}_m = \mathbf{r}_m + \Delta \mathbf{r}_m$, where $\Delta \mathbf{r}_m$ is the estimation error of the m th column of the SCM, it can be found that

$$\Delta \sigma_m^2 = \frac{\Re\{\mathbf{v}_m^H \Delta \mathbf{r}_m\}}{\tau_m} \quad (16)$$

where $\Delta \sigma_m^2$ is the difference between the actual σ_m^2 and the estimate $\hat{\sigma}_m^2$, i.e., $\Delta \sigma_m^2 = \hat{\sigma}_m^2 - \sigma_m^2$. Consequently, the variance of $\Delta \sigma_m^2$ can be expressed as

$$\begin{aligned} \mathbb{E} \left\{ (\Delta \sigma_m^2)^2 \right\} &= \frac{1}{4\tau_m^2} \mathbb{E} \left\{ (\mathbf{v}_m^H \Delta \mathbf{r}_m + \mathbf{v}_m^T \Delta \mathbf{r}_m^*) \right. \\ &\quad \left. \times (\Delta \mathbf{r}_m^H \mathbf{v}_m + \Delta \mathbf{r}_m^T \mathbf{v}_m^*) \right\} \\ &= \frac{1}{4\tau_m^2} \left(\mathbf{v}_m^H \mathbb{E} \left\{ \Delta \mathbf{r}_m \Delta \mathbf{r}_m^H \right\} \mathbf{v}_m + \mathbf{v}_m^H \mathbb{E} \left\{ \Delta \mathbf{r}_m \Delta \mathbf{r}_m^T \right\} \mathbf{v}_m^* \right. \\ &\quad \left. + \mathbf{v}_m^T \mathbb{E} \left\{ \Delta \mathbf{r}_m^* \Delta \mathbf{r}_m^H \right\} \mathbf{v}_m + \mathbf{v}_m^T \mathbb{E} \left\{ \Delta \mathbf{r}_m^* \Delta \mathbf{r}_m^T \right\} \mathbf{v}_m^* \right). \end{aligned} \quad (17)$$

According to [22], the asymptotic covariance and pseudo-covariance matrices of the vector $\Delta \mathbf{r} \triangleq \text{vec}\{(\hat{\mathbf{R}} - \mathbf{R})\} \in \mathbb{C}^{M^2}$ are

$$\mathbb{E} \left\{ \Delta \mathbf{r} \Delta \mathbf{r}^H \right\} = \frac{1}{N} (\mathbf{R}^T \otimes \mathbf{R}) \quad (18)$$

$$\mathbb{E} \left\{ \Delta \mathbf{r} \Delta \mathbf{r}^T \right\} = \frac{1}{N} (\mathbf{R}^T \otimes \mathbf{R}) \mathbf{K}. \quad (19)$$

Using (18) and (19), it is straightforward to show that [23]

$$\mathbb{E} \left\{ \Delta \mathbf{r}_m \Delta \mathbf{r}_m^H \right\} = \left(\frac{[\mathbf{R}]_{mm}}{N} \right) \mathbf{R} \quad (20)$$

$$\mathbb{E} \left\{ \Delta \mathbf{r}_m \Delta \mathbf{r}_m^T \right\} = \frac{1}{N} (\mathbf{r}_m^T \otimes \mathbf{R}) \mathbf{K} (\mathbf{d}_m \otimes \mathbf{I}_M). \quad (21)$$

Plugging (20) and (21) into (17) yields (14), which completes the proof. \blacksquare

Remark 1: It is worth noting that (12) represents the power domain (PD) method [17] in an alternative way. The PD method estimates σ_m^2 's as

$$\sigma_m^2 = \frac{(\mathbf{d}_m^T \mathbf{P}_A^\perp \hat{\mathbf{r}}_m + \hat{\mathbf{r}}_m^H \mathbf{P}_A^\perp \mathbf{d}_m)}{2\mathbf{d}_m^T \mathbf{P}_A^\perp \mathbf{d}_m}, \quad m = 1, \dots, M \quad (22)$$

where $\mathbf{P}_A^\perp \in \mathbb{C}^{M \times M}$ is the orthogonal projection matrix of the signal subspace, i.e., $\mathbf{P}_A^\perp = \mathbf{I}_M - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$. The primary difference between (12) and (22) is the use of $\hat{\mathbf{U}}\hat{\mathbf{U}}^H$ as an estimate of \mathbf{P}_A^\perp rather than conducting a multidimensional search like in [17] for an ML estimate. In general, such multidimensional search is known to be very computationally demanding. Due to its reduced computational requirements, the proposed method has significant advantage over the PD technique.

Remark 2: The sufficient number of iterations for achieving a precise estimate of the noise covariance matrix for $\hat{\mathbf{Q}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$ is 3–5 as will be shown in the next section.

4. SIMULATION RESULTS

We evaluate the performance of the proposed method and compare it to that of the other state-of-the-art algorithms. The methods used for comparison are the “NISB+MUSIC” method of [20], “IMLSE+MUSIC” method of [18], and “RTM+MUSIC” method of [19]. The nonuniform stochastic CRB [14] is used as the benchmark. A ULA with $M = 16$ sensors separated by half wavelength collecting $N = 8$ snapshots is considered, and 2000 Monte Carlo runs are conducted to calculate the root mean square error (RMSE) defined as

$$\text{RMSE} = 10 \log_{10} \sqrt{\frac{1}{2000L} \sum_{l=1}^L \sum_{i=1}^{2000} (\hat{\theta}_{l,i} - \theta_i)^2}.$$

The SNR is computed as $\text{SNR} = \frac{\sigma_s^2}{M} \sum_{m=1}^M \frac{1}{\sigma_m^2}$ where σ_s^2 represents the identical powers of different sources. The sensor noise covariance matrix is set to $\mathbf{Q} = \text{diag}\{[6, 2, 0.5, 2.5, 3, 10, 5.5, 30, 11, 1.2, 3.5, 18, 2, 8.5, 36, 6.5]\}$. As a result, the worst noise power ratio (WNPR) in these examples is

$$\text{WNPR} = \frac{\sigma_{max}^2}{\sigma_{min}^2} = \frac{36}{0.5} = 72.$$

Fig. 1 shows the RMSE performance of the methods tested versus SNR for the setup of two uncorrelated sources with $\theta = [-29^\circ, 18^\circ]$. We observe that the proposed method achieves a higher threshold performance than other methods tested. In addition, Fig. 2 shows the strengths of the methods tested against the presence of closely located sources for the case of uncorrelated sources. The setup used for producing Fig. 2 is $\theta = [18^\circ, (18 + \Delta\theta)^\circ]$ and SNR = 10 dB with $\Delta\theta$ varying from 4° to 10° . Compared to the other methods tested, the performance of the proposed method is better as illustrated in Fig. 2.

5. CONCLUSION

A novel algorithm is presented to estimate DOAs when the sensor noise is nonuniform. Our algorithm iteratively estimates the nonuniform noise covariance matrix. Each iteration involves estimation of the noise subspace using GED

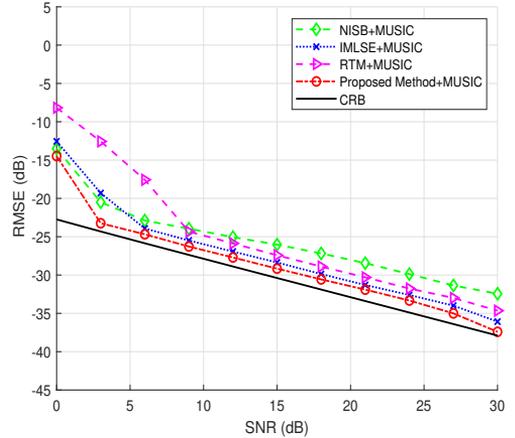


Fig. 1. RMSE vs. SNR for $L = 2$ uncorrelated sources with $\theta = [-29^\circ, 18^\circ]$, $M = 16$, and $N = 8$.

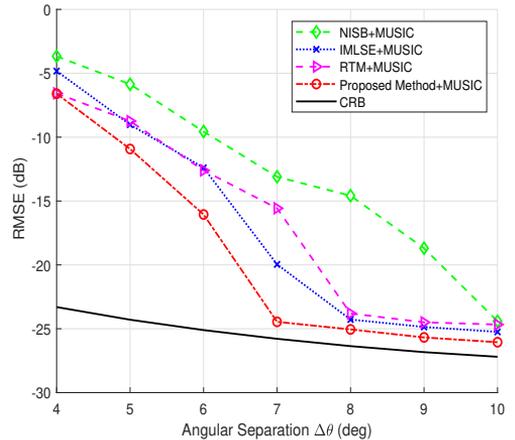


Fig. 2. RMSE vs. the angular separation for $L = 2$ uncorrelated sources with $\theta = [18^\circ, (18 + \Delta\theta)^\circ]$, SNR = 10 dB, $M = 16$, and $N = 8$.

first, followed by updating the noise covariance matrix using LS. In addition to being applicable to a wide variety of array geometries, the proposed noise covariance matrix estimator is also fast and easy to implement since it only requires a few iterations to provide accurate estimates. Simulation examples show that the proposed algorithm is superior to the existing approaches.

6. REFERENCES

- [1] 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.
- [2] K. Upadhyaya and S. A. Vorobyov, "An array processing approach to pilot decontamination for massive mimo," in *Proc. IEEE 6th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process., Cancun, Mexico*, Dec. 2015, pp. 453–456.
- [3] K. Upadhyaya, S. A. Vorobyov, and R. W. Heath, "Low-overhead receiver-side channel tracking for mmwave mimo," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Alberta, Canada*, Apr. 2018, pp. 3859–3863.
- [4] A. Hassanien and S. A. Vorobyov, "Transmit energy focusing for doa estimation in mimo radar with colocated antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2669–2682, Jun. 2011.
- [5] A. Khabbazibasmenj, A. Hassanien, S. A. Vorobyov, and M. W. Morency, "Efficient transmit beamspace design for search-free based doa estimation in mimo radar," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1490–1500, Mar. 2014.
- [6] J. Zhang, D. Rakhimov, and M. Haardt, "Gridless channel estimation for hybrid mmwave mimo systems via tensor-esprit algorithms in dft beamspace," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 816–831, Apr. 2021.
- [7] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [8] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 1983, vol. 8, pp. 336–339.
- [9] 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.
- [10] M. Shaghghi 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] S.A. Vorobyov, A.B. Gershman, and Wong K.M., "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.
- [16] 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.
- [17] D. Madurasinghe, "A new doa estimator in nonuniform noise," *IEEE Signal Process. Lett.*, vol. 12, no. 4, pp. 337–339, Apr. 2005.
- [18] 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.
- [19] 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 Sensors J.*, vol. 16, no. 24, pp. 8982–8989, Dec. 2016.
- [20] 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.
- [21] M. Esfandiari and S. A. Vorobyov, "DOA estimation in nonuniform sensor noise," *arXiv e-prints*, pp. 1–35, arXiv:2109.15043, Sept. 2021.
- [22] M. Bilodeau and D. Brenner, *Theory of multivariate statistics*, Springer Science & Business Media, 2008.
- [23] 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.

Publication 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.

© 2022

Reprinted with permission.

SPARSITY ENFORCING WITH TOEPLITZ MATRIX RECONSTRUCTION METHOD FOR MMWAVE UL CHANNEL ESTIMATION WITH ONE-BIT ADCS

*Majdoddin Esfandiari**, *Sergiy A. Vorobyov**, and *Robert W. Heath Jr.[†]*

*Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

[†]Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, USA

ABSTRACT

One-bit analog-to-digital converters enable digital beamforming in millimeter wave (mmWave) multi-input multi-output communication systems with low power consumption. Conventional signal processing tasks like channel estimation, though, are challenging due to the extreme quantization, making it challenging to implement uplink (UL) multiuser receivers. We reformulate the UL channel estimation problem as a multiplication of two specific matrices, and then we leverage Toeplitz matrix reconstruction in conjunction with the angular domain sparsity of the UL channel to recover UL channel via solving a properly designed optimization problem. Our new approach is called the sparsity enforcing with Toeplitz matrix reconstruction (SE-TMR) method. Numerical simulations are carried out to showcase the advantages of SE-TMR over existing competitive methods in terms of normalized mean squared error in clustered narrowband channels.

Index Terms— Multi-user MIMO, uplink channel estimation, one-bit analog-to-digital converter, angular domain, Toeplitz matrix reconstruction.

1. INTRODUCTION

Millimeter-wave (mmWave) multiple-input multiple-output (MIMO) communications is an approach that extends the high data rate benefits associated with the use of mmWave to take advantage of multiuser multiplexing provided by massive MIMO [1–5]. Despite the advantages offered by mmWave MIMO system, the need to deploy high-resolution analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) for the large number of antenna elements in the base station (BS) makes system impractical in terms of the power consumption for the large arrays found in massive MIMO and the higher bandwidths in mmWave and sub-THz communications. The use of low-resolution (1-3 bits) ADCs/DACs is one approach to reduce overall transceiver power consumption [7, 9–12]. The most benefits in terms of power consumption and reduced analog hardware complexity are found with one-bit data converters, where the

conventional treatment of quantization error as additive noise may be a poor assumption.

Channel estimation is required for most one-bit receiver designs and has been widely considered in prior works [8, 11, 13–17]. Of relevance to our paper, in [13] and [14], the maximum-likelihood (ML) approach has been used to handle uplink (UL) channel estimation problem, but the prohibitive computational cost required for implementing the proposed algorithms has been the main impediment in practical scenarios. In [15], an algorithm referred to as Bussgang linear minimum mean squared error (BLMMSE) channel estimation has been developed for both flat and frequency-selective channel models. The essence of BLMMSE is to approximate the non-linear one-bit quantizer as a linear function via the Bussgang decomposition [18]. In [11], the authors have proposed a generalized approximate message passing (GAMP) based algorithm, where a compressive sensing (CS) approach has been employed to estimate the angular domain parameters of UL channel. In [16], an amplitude retrieval (AR) based algorithm has been derived that is based on completing the lost amplitudes of one-bit measurements along with ML direction-of-arrival (DOA) estimation method. Recently, an algorithm referred to as gridless GAMP (GL-GAMP) has been developed in [17] for UL channel estimation. GL-GAMP utilizes modified expectation-maximization GAMP (EM-GAMP) method in conjunction with the well-known RELAX [19] method to reconstruct the channel. Motivated by the promising use of the structure-based methods such as the AR and GL-GAMP methods, we develop a channel estimation approach by leveraging the underlying Toeplitz structure of the UL channel along with angular domain sparsity, which has not been fully exploited in prior works.

In this paper, we develop a new method referred to as sparsity enforcing with Toeplitz matrix reconstruction (SE-TMR) method to address the problem of UL channel estimation for mmWave MIMO communications when one-bit ADCs are deployed at the BS. The core idea of SE-TMR is to first reformulate UL channel estimation problem in terms of multiplication of two specific matrices, followed by using the combination of Toeplitz matrix reconstruction and UL channel sparsity in the angular domain. SE-TMR reconstructs efficiently UL channel up to a scale factor. Numerical simula-

This work was supported in part by Academy of Finland under Research Grant 319822.

tions validate its performance improvement and compare it to the existing competitive methods.

Notations: Upper-case and lower-case bold-face letters denote matrices and vectors, respectively, while scalars are denoted by lower-case letters. The transpose, and Hermitian transpose are denoted by $\{\cdot\}^T$ and $\{\cdot\}^H$, respectively, while $\|\cdot\|$ and $\|\cdot\|_1$ denote the l_2 and l_1 norms of a vector, and $\|\cdot\|_F$ stands for the Frobenius norm of a matrix. The notation $\mathcal{T}(\cdot)$ stands for an operation of building a square Hermitian Toeplitz matrix with its first column being the bracketed vector, and $\text{trace}\{\cdot\}$ stands for the trace of a square matrix. The $n \times n$ identity matrix is denoted by \mathbf{I}_n . The $n \times 1$ vector with all its entries equal to one is denoted by $\mathbf{1}_n$. The i th entry of the vector $\boldsymbol{\pi}$ is denoted by $[\boldsymbol{\pi}]_i$, while the entry in the intersection of the i th row and j th column of the matrix $\mathbf{\Pi}$ is denoted as $[\mathbf{\Pi}]_{i,j}$. The operator $\text{vec}\{\cdot\}$ stacks the columns of a matrix into a long vector. The operator $\text{diag}\{\boldsymbol{\pi}\}$ generates a diagonal matrix by plugging the entries of the vector $\boldsymbol{\pi}$ into its main diagonal. Finally, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ return the real and imaginary parts of the bracketed argument, respectively.

2. UL CHANNEL ESTIMATION

2.1. System Model

Consider an UL multiuser mmWave MIMO system equipped with a uniform linear array (ULA) of M antenna elements at the base station (BS) and K single antenna users. The BS is equipped with one-bit ADCs, whereas all users deploy high-resolution DACs. Then, the UL channel between user k and the BS can be formulated as

$$\mathbf{h}_k = \sum_{l=1}^{L_k} \sum_{m=1}^{M_{\text{path}}^{k,l}} \gamma_{k,l,m} \mathbf{a}(\boldsymbol{\theta}_{k,l,m})$$

$$= [\mathbf{A}(\boldsymbol{\theta}_{k,1}), \dots, \mathbf{A}(\boldsymbol{\theta}_{k,L_k})] \begin{bmatrix} \gamma_{k,1} \\ \vdots \\ \gamma_{k,L_k} \end{bmatrix} = \mathbf{A}(\boldsymbol{\theta}_k) \boldsymbol{\gamma}_k \quad (1)$$

where L_k denotes the number of multipath clusters between the BS and the user k , the l th cluster encompasses $M_{\text{path}}^{k,l}$ paths concentrated around an angular area defined by the corresponding angle spread [20], $\gamma_{k,l,m}$ and $\boldsymbol{\theta}_{k,l,m}$ are respectively the gain and DOA of the m th path in the l th cluster, the steering vector is set as $\mathbf{a}(\boldsymbol{\theta}_{k,l,m}) = [1, e^{-j\pi \sin(\boldsymbol{\theta}_{k,l,m})}, \dots, e^{-j(M-1)\pi \sin(\boldsymbol{\theta}_{k,l,m})}]^T \in \mathbb{C}^{M \times 1}$, $\boldsymbol{\theta}_{k,l} \triangleq [\boldsymbol{\theta}_{k,l,1}, \dots, \boldsymbol{\theta}_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{R}^{M_{\text{path}}^{k,l} \times 1}$ for $l = 1, \dots, L_k$, $\mathbf{A}(\boldsymbol{\theta}_{k,l}) \triangleq [\mathbf{a}(\boldsymbol{\theta}_{k,l,1}), \dots, \mathbf{a}(\boldsymbol{\theta}_{k,l,M_{\text{path}}^{k,l}})] \in \mathbb{C}^{M \times M_{\text{path}}^{k,l}}$, $\boldsymbol{\gamma}_{k,l} \triangleq [\gamma_{k,l,1}, \dots, \gamma_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{C}^{M_{\text{path}}^{k,l} \times 1}$, $\boldsymbol{\theta}_k \triangleq [\boldsymbol{\theta}_{k,1}^T, \dots, \boldsymbol{\theta}_{k,L_k}^T]^T$, $\mathbf{A}(\boldsymbol{\theta}_k) \triangleq [\mathbf{A}(\boldsymbol{\theta}_{k,1}), \dots, \mathbf{A}(\boldsymbol{\theta}_{k,L_k})]$, and $\boldsymbol{\gamma}_k \triangleq [\boldsymbol{\gamma}_{k,1}^T, \dots, \boldsymbol{\gamma}_{k,L_k}^T]^T$. The overall channel between the BS and K users is

$$\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K] = [\mathbf{A}(\boldsymbol{\theta}_1) \boldsymbol{\gamma}_1, \dots, \mathbf{A}(\boldsymbol{\theta}_K) \boldsymbol{\gamma}_K]. \quad (2)$$

In the training stage, a pilot sequence of length N_s ($N_s \geq K$) is transmitted by K users. The received signal at the BS is

$$\mathbf{Y} = \mathcal{Q}(\mathbf{H}\mathbf{S} + \mathbf{N}) \quad (3)$$

where $\mathcal{Q}(\cdot) \triangleq \text{sign}(\Re\{\cdot\}) + j\text{sign}(\Im\{\cdot\})$ is the element-wise one-bit quantizer which maps an argument to one of the members of the set $\mathcal{S} = \{1+j, 1-j, -1+j, -1-j\}$, $\mathbf{S} \in \mathbb{C}^{K \times N_s}$ represents the orthogonal pilot matrix, and $\mathbf{N} \in \mathbb{C}^{M \times N_s}$ is the additive circularly symmetric complex-valued Gaussian noise with zero mean and variance σ^2 . The aim is to recover (scaled) $\mathbf{H} \in \mathbb{C}^{M \times K}$ by processing the received signal $\mathbf{Y} \in \mathbb{C}^{M \times N_s}$.

2.2. Proposed UL Channel Estimation

Our UL channel estimation method will be developed by employing the notion of Toeplitz matrix reconstruction combined with the angular domain sparsity of UL channel. Although the actual channel between the BS and the k th user presented in (1) is composed of many paths, it is still sparse in an angular-based dictionary (e.g., the normalized discrete Fourier transform (DFT) matrix). Therefore, \mathbf{h}_k can be approximated as an unknown linear combination of a few atoms of the angular-based dictionary which correspond to those paths that contribute the most. These paths are entitled as the “*basis paths*” henceforth. From this point of view, we represent the k th column of \mathbf{H} (i.e., \mathbf{h}_k) as a linear combination of L_k basis paths with L_k path gains and DOAs. As a result, (2) can be recast as

$$\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K] = [\mathbf{A}(\bar{\boldsymbol{\theta}}_1) \bar{\boldsymbol{\gamma}}_1, \dots, \mathbf{A}(\bar{\boldsymbol{\theta}}_K) \bar{\boldsymbol{\gamma}}_K] \quad (4)$$

where $\bar{\boldsymbol{\theta}}_k \triangleq [\bar{\boldsymbol{\theta}}_{k,1}, \dots, \bar{\boldsymbol{\theta}}_{k,L_k}]^T \in \mathbb{R}^{L_k \times 1}$ and $\bar{\boldsymbol{\gamma}}_k \triangleq [\bar{\gamma}_{k,1}, \dots, \bar{\gamma}_{k,L_k}]^T \in \mathbb{C}^{L_k \times 1}$ are respectively the DOAs and path gains which correspond to L_k basis paths between the BS and user k .

To develop our method, we first reformulate (4) as

$$\mathbf{H} = \mathbf{A} \boldsymbol{\Gamma} \bar{\mathbf{G}} = \bar{\mathbf{H}} \bar{\mathbf{G}} \quad (5)$$

where

$$\mathbf{A} \triangleq [\mathbf{A}(\bar{\boldsymbol{\theta}}_1), \dots, \mathbf{A}(\bar{\boldsymbol{\theta}}_K)] \in \mathbb{C}^{M \times L} \quad (6)$$

$$\boldsymbol{\Gamma} \triangleq \begin{bmatrix} \text{diag}(\bar{\boldsymbol{\gamma}}_1) & & \\ & \ddots & \\ & & \text{diag}(\bar{\boldsymbol{\gamma}}_K) \end{bmatrix} \in \mathbb{C}^{L \times L} \quad (7)$$

$$\bar{\mathbf{G}} \triangleq \begin{bmatrix} \mathbf{1}_{L_1} & & \\ & \mathbf{1}_{L_2} & \\ & & \ddots \\ & & & \mathbf{1}_{L_K} \end{bmatrix} \in \mathbb{R}^{L \times K} \quad (8)$$

$$\bar{\mathbf{H}} \triangleq \mathbf{A} \boldsymbol{\Gamma} \in \mathbb{C}^{M \times L} \quad (9)$$

and $L \triangleq \sum_{k=1}^K L_k$. The significant point stated in (5) is that estimating \mathbf{H} is equivalent to estimating $\bar{\mathbf{H}}$. Therefore, we develop SE-TMR by formulating an optimization problem with respect to $\bar{\mathbf{H}}$ because of good properties that $\bar{\mathbf{H}}$ possesses as it will be elaborated in the sequel.

Due to the special structures of \mathbf{A} and $\mathbf{\Gamma}$, i.e., \mathbf{A} and $\mathbf{\Gamma}$ being respectively Vandermonde and diagonal matrices, it is straightforward to show that

$$\bar{\mathbf{H}}\bar{\mathbf{H}}^H = \mathcal{T}(\mathbf{u}) \quad (10)$$

where $\mathbf{u} \in \mathbb{C}^M$ and $[\mathbf{u}]_1$ belongs to the real numbers field.

Moreover, the channel estimation problem can be converted into a sparse recovery problem by defining an angular-based dictionary [21, 22] as explained above. In the case of ULA, the normalized DFT matrix can be a proper dictionary [17] because of zero intra-column correlation, i.e., the DFT matrix columns are orthonormal to each other. Hence, multiplying $\bar{\mathbf{H}}$ by the DFT matrix, we have

$$\mathbf{X}(\bar{\theta}) = \mathbf{F}\bar{\mathbf{H}} = \mathbf{F}\bar{\mathbf{H}}\bar{\mathbf{G}} \quad (11)$$

where $\mathbf{F} \in \mathbb{C}^{M \times M}$ denotes the normalized DFT matrix. Taking into consideration (5), it can be concluded that each column of $\mathbf{X}(\bar{\theta})$ defined in (11) is sparse. It is well known that the optimal way to enforce sparsity is to use l_0 pseudo-norm, however, the corresponding optimization problem is known to be NP hard. Therefore, we exploit l_1 norm to enforce sparsity, which is a widely used alternative.

Considering (10) and (11), the following optimization problem is introduced to recover $\bar{\mathbf{H}}$ efficiently given the one-bit measurement matrix \mathbf{Y}

$$\min_{\bar{\mathbf{H}}, \mathbf{u}, \mathbf{E}^R, \mathbf{E}^I} \|\text{vec}\{\mathbf{F}\bar{\mathbf{H}}\bar{\mathbf{G}}\}\|_1 + \lambda \left(\sum_{i=1}^M \sum_{j=1}^{N_s} ([\mathbf{E}^R]_{i,j} + [\mathbf{E}^I]_{i,j}) \right) \quad (12)$$

$$\text{s.t. } \bar{\mathbf{H}}\bar{\mathbf{H}}^H = \mathcal{T}(\mathbf{u}) \quad (13)$$

$$\text{trace}\{\mathcal{T}(\mathbf{u})\} = C \quad (14)$$

$$\Re\{[\bar{\mathbf{H}}\bar{\mathbf{G}}\mathbf{S}]_{i,j}\}\Re\{[\mathbf{Y}]_{i,j}\} \geq -[\mathbf{E}^R]_{i,j}, \quad (15)$$

$$\Im\{[\bar{\mathbf{H}}\bar{\mathbf{G}}\mathbf{S}]_{i,j}\}\Im\{[\mathbf{Y}]_{i,j}\} \geq -[\mathbf{E}^I]_{i,j}, \quad (16)$$

$$[\mathbf{E}^R]_{i,j} \geq 0, \quad i = 1, \dots, M, \quad j = 1, \dots, N_s \quad (17)$$

$$[\mathbf{E}^I]_{i,j} \geq 0, \quad i = 1, \dots, M, \quad j = 1, \dots, N_s \quad (18)$$

where $\lambda > 0$ 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 introduced to handle probable sign flips due to the impact of noise [23]. In this problem, (14) prevents the scaling ambiguity with $C > 0$, and (15)–(16) are imposed to maintain the consistency with the observation matrix \mathbf{Y} . The optimization problem (12)–(18) is non-convex because of the constraint

(13) which is difficult to address in an efficient manner. For alleviating difficulties caused by imposing (13), semi-definite programming (SDP) relaxation can be exploited to turn the non-convex constraint (13) into a convex one. Therefore, the optimization problem (12)–(18) can be modified by means of SDP relaxation as

$$\min_{\bar{\mathbf{H}}, \mathbf{u}, \mathbf{E}^R, \mathbf{E}^I} \|\text{vec}\{\mathbf{F}\bar{\mathbf{H}}\bar{\mathbf{G}}\}\|_1 + \lambda \left(\sum_{i=1}^M \sum_{j=1}^{N_s} ([\mathbf{E}^R]_{i,j} + [\mathbf{E}^I]_{i,j}) \right) \quad (19)$$

$$\text{s.t. } \begin{bmatrix} \mathbf{I}_L & \bar{\mathbf{H}}^H \\ \bar{\mathbf{H}} & \mathcal{T}(\mathbf{u}) \end{bmatrix} \succeq 0 \quad (20)$$

$$[\mathbf{u}]_1 = \frac{C}{M} \quad (21)$$

$$\Re\{[\bar{\mathbf{H}}\bar{\mathbf{G}}\mathbf{S}]_{i,j}\}\Re\{[\mathbf{Y}]_{i,j}\} \geq -[\mathbf{E}^R]_{i,j}, \quad (22)$$

$$\Im\{[\bar{\mathbf{H}}\bar{\mathbf{G}}\mathbf{S}]_{i,j}\}\Im\{[\mathbf{Y}]_{i,j}\} \geq -[\mathbf{E}^I]_{i,j}, \quad (23)$$

$$[\mathbf{E}^R]_{i,j} \geq 0, \quad i = 1, \dots, M, \quad j = 1, \dots, N_s \quad (24)$$

$$[\mathbf{E}^I]_{i,j} \geq 0, \quad i = 1, \dots, M, \quad j = 1, \dots, N_s \quad (25)$$

where the convex constraint (20) is imposed instead of (13) via SDP relaxation, and (14) is replaced by (21) as they are interchangeable. The optimization problem (19)–(25) is convex and therefore can be solved by off-the-shelf solvers like CVX [24].

After estimating $\bar{\mathbf{H}}$, we recover \mathbf{H} using (5). Due to the angular-based structure of the columns of $\bar{\mathbf{H}}$, further refinement can be attained through recovering the L_k^{HR} path gains and L_k^{HR} DOAs associated with the k th column of $\bar{\mathbf{H}}$ for $k = 1, \dots, K$. In doing so, conventional one-dimensional harmonic retrieval (HR) methods such as RELAX [19] can be applied to each column of $\bar{\mathbf{H}}$ in order to estimate the L_k^{HR} path gains and L_k^{HR} DOAs, and then reconstruct the refined \mathbf{H} [17], denoted by $\hat{\mathbf{H}}$. Note that L_k^{HR} can be opted greater than L_k due to the presence of actual larger number of multipaths than the basis paths as given in (1). As a matter of fact, it will be shown in the next section by means of simulation that choosing L_k^{HR} greater than L_k leads to performance improvement of the SE-TMR method. The steps of the proposed UL channel estimator are outlined in Algorithm 1.

3. SIMULATION RESULTS

We evaluate the performance of the proposed SE-TMR method and compare it to that of other competitive algorithms. The methods used for comparison are near ML (nML) [13], AR [16], and zero-forcing (ZF) method of [13]. A Zadoff-Chu (ZC) sequence with length N_s is used to construct the pilot sequence, such that each row of \mathbf{S} is a circularly shifted replica of the ZC sequence and is therefore

Algorithm 1: SE-TMR Algorithm

Input: \mathbf{Y} , $\bar{\mathbf{G}}$, λ , L_k 's, L_k^{HR} 's

1: Obtain $\hat{\mathbf{H}}$, \mathbf{u} , \mathbf{E}^R , and \mathbf{E}^I by solving the optimization problem (19)-(25).

2: Calculate $\hat{\mathbf{H}} = \bar{\mathbf{H}}\bar{\mathbf{G}}$.

Refinement with RELAX:

3: Apply RELAX to each column of $\hat{\mathbf{H}}$ to estimate each user's L_k^{HR} DOAs and path gains.

4: Employ the estimated DOAs and path gains of each user to recover the channel matrix $\hat{\hat{\mathbf{H}}}$.

Output: $\hat{\hat{\mathbf{H}}}$

orthogonal to the other rows.¹ The signal-to-noise ratio (SNR) and normalized mean square error (NMSE) are respectively defined as $\text{SNR} \triangleq 10 \log_{10} \left(\frac{\|\mathbf{S}\|_F^2}{N_s \sigma^2} \right)$ and $\text{NMSE} \triangleq$

$$\frac{1}{KN} \sum_{k=1}^K \sum_{n=1}^N \left\| \frac{\hat{\mathbf{h}}_k^{(n)}}{\|\hat{\mathbf{h}}_k^{(n)}\|} - \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|} \right\|^2, \text{ where } \hat{\mathbf{h}}_k^{(n)} \text{ is the } k\text{th}$$

column of $\hat{\mathbf{H}}$ estimated in the n th Monte Carlo run, and N is the total number of Monte Carlo runs which is $N = 100$ in this paper. We consider $\lambda = 10$, and $K = 8$ in all examples.² Moreover, the same number of channel clusters and within cluster multipaths for each user are considered in the upcoming examples,³ i.e., $L_1 = \dots = L_K$ and $M_{\text{path}}^{1,1} = \dots = M_{\text{path}}^{1,L_1} = \dots = M_{\text{path}}^{K,1} = \dots = M_{\text{path}}^{1,L_K} = 100$. All UL DOAs are generated randomly once and remain the same throughout the Monte Carlo runs. Moreover, the channel path gains are distributed as $\mathcal{CN}(0, 1)$ for all users. Fig. 1 shows the NMSE performance of the methods tested versus SNR for the setup of $M = 16$, $N_s = 128$, and $L_k = 1$. Moreover, the angle spread within a cluster is 8 degrees. Note that L_k^{HR} is set as 4 for the SE-TMR method in this example. It shows that the performance of the proposed SE-TMR method is substantially better than that of the other methods tested. Note that RELAX method is also applied to nML and ZF methods (see steps 3 and 4 in Algorithm 1) to reconstruct the structured estimate of the UL channel. Hence, the original methods without RELAX-based refinement are called “unstructured” in the figures.⁴ The setup considered for Fig. 2 is $M = 16$, $N_s = 128$, and $L_k = 2$. The angle spread within one cluster is 8 degrees, while it is 10 degrees for the other one. In Fig. 2, the proposed SE-TMR method outperforms the other

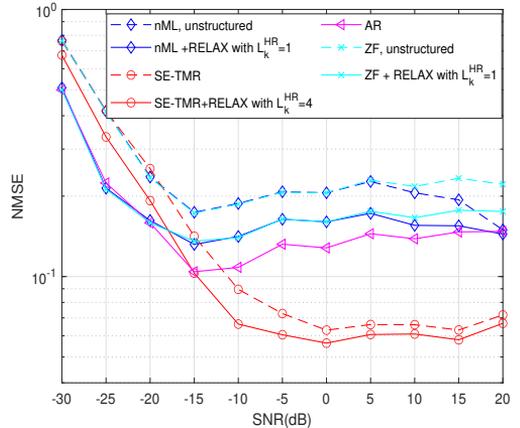


Fig. 1. NMSE vs. SNR for $M = 16$, $N_s = 128$, and $L_k = 1$.

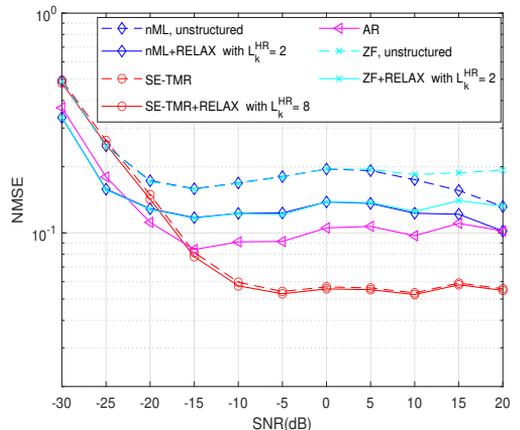


Fig. 2. NMSE vs. SNR for $M = 16$, $N_s = 128$, and $L_k = 2$.

methods tested.

4. CONCLUSION

A novel algorithm called SE-TMR is presented to estimate UL channel in mmWave multi-user MIMO systems with one-bit ADCs at the BS. The essence of the SE-TMR method is to reconstruct UL channel by solving a properly designed optimization problem, which leverages the Toeplitz matrix reconstruction and sparsity of UL channel in the DFT domain. Based on the numerical simulations provided, the SE-TMR method outperforms existing competitive methods in different scenarios with diverse number of dominated paths between the BS and users.

¹In our algorithm, the choice of pilot sequence is not of significance, i.e., our algorithm also works with other pilot sequences. ZC sequence is selected here solely due to its popularity.

²It is observed that setting λ much greater than one leads to better performance, i.e., $\lambda \gg 1$.

³Different users may have different numbers of channel clusters and multipaths within clusters. For simplicity, we use the same number of channel clusters and within cluster multipaths for all users here.

⁴RELAX method is not applied to AR since the AR algorithm uses an ML DOA estimator as one step of the channel estimation algorithm, therefore, it belongs to the category of structured methods.

5. REFERENCES

- [1] 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.
- [2] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Millimeter-wave massive MIMO: The next wireless revolution?," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [3] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capollio, "Five disruptive technology directions for 5G," *IEEE Comm. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.
- [4] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roth, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [5] T. S. Rappaport *et al.*, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access*, vol. 1, pp. 335–349, May. 2013.
- [6] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.
- [7] A. Mezghani, F. Antreich, and J. A. Nossek, "Multiple parameter estimation with quantized channel output," in *Proc. Int. ITG Workshop Smart Antennas*, Bremen, Germany, Feb. 2010, pp. 143–150.
- [8] C. Mollen, J. Choi, E. G. Larsson, and R. W. Heath Jr., "Uplink Performance of Wideband Massive MIMO With One-Bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Jan. 2017.
- [9] C. Wen, C. Wang, S. Jin, 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.
- [10] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.
- [11] 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.
- [12] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, Jul. 2018.
- [13] 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.
- [14] F. Wang, J. Fang, H. Li, Z. Chen, and S. Li, "One-bit quantization design and channel estimation for massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10921–10934, Nov. 2018.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. 216, 1952.
- [19] 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.
- [20] "5G channel model for bands up to 100 GHz (v2.3)," Tech. Rep., 2016. [Online]. Available: <http://www.5gworkshops.com/5gcm.html>
- [21] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.
- [22] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial-and-frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3393–3406, Jul. 2018.
- [23] Y. Gao, D. Hu, Y. Chen, and Y. Ma, "Gridless 1-b DOA estimation exploiting SVM approach," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2210–2213, Oct. 2017.
- [24] M. Grant, S. Boyd, and Y. Ye, CVX: Matlab software for disciplined convex programming. 2008 [Online]. Available: <http://stanford.edu/~boyd/cvx>

Publication 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.

© 2023

Reprinted with permission.

ADMM-based Solution for mmWave UL Channel Estimation with One-bit ADCs via Sparsity Enforcing and Toeplitz Matrix Reconstruction

Majdoddin Esfandiari*, Sergiy A. Vorobyov*, and Robert W. Heath Jr.†

*Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

†Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, USA

ABSTRACT

Low-power millimeter wave (mmWave) multi-input multi-output communication systems can be enabled with the use of one-bit analog-to-digital converters. Owing to the extreme quantization, conventional signal processing tasks such as channel estimation are challenging, making uplink (UL) multi-user receivers difficult to implement. To address this issue, we first reformulate the UL channel estimation problem, and then combine the idea of ℓ_1 regularized logistic regression classification and Toeplitz matrix reconstruction in a properly designed optimization problem. Our new method is referred to as ℓ_1 regularized logistic regression with Toeplitz matrix reconstruction (L1-RLR-TMR). In addition, we develop a computationally efficient alternating direction method of multipliers (ADMM)-based implementation for the L1-RLR-TMR method. Numerical results demonstrate the performance of the L1-RLR-TMR method in comparison with other existing methods.

Index Terms— Multi-user MIMO, uplink channel estimation, one-bit ADC, angular domain, Toeplitz matrix reconstruction, ℓ_1 regularized logistic regression.

1. INTRODUCTION

Millimeter-wave (mmWave) multiple-input multiple-output (MIMO) communication is an approach to enhance the high data rate benefits provided by mmWave while enjoying the advantages of multi-user multiplexing offered by massive MIMO [1–4]. One approach to reduce transceiver power consumption is to use low-resolution ADCs/DACs (1–3 bits) [6, 8–10]. With one-bit data converters, the power consumption can be reduced to a minimum. Therefore, the methods developed for the case of high resolution converters for tasks like channel estimation or symbol detection have poor performance in the case of one-bit converters utilization which emphasize the need for designing methods consistent with one-bit converters setups.

Most one-bit receiver designs require channel estimation [7, 9, 11–13]. A near maximum-likelihood (nML) was proposed in [11], but its performance was not good at high signal-to-noise ratio (SNR). In [12], an algorithm known as Bussgang linear minimum mean squared error (BLMMSE)

has been proposed that can handle the estimation of both flat and frequency-selective channels. BLMMSE approximates the nonlinear one-bit quantizer as a linear function by employing the Bussgang decomposition [16]. A generalized approximate message passing (GAMP) based algorithm has been presented in [9], in which the notion of compressive sensing (CS) is used to estimate the angular domain parameters of UL channels. Using ML direction-of-arrival estimation method together with restoring the lost amplitudes of one-bit measurements, [13] describes a method called the amplitude retrieval (AR). In [14], the problem of UL channel estimation is formulated as a binary classification task, in which the conventional support vector machine (SVM) and its modified version are adopted to handle the estimation of the uncorrelated and spatially correlated channels, respectively. Recently, a method known as the sparsity enforcing with Toeplitz matrix reconstruction (SE-TMR) has been presented in [15] where the combination of the angular domain sparsity and Toeplitz matrix reconstruction is leveraged to derive the SE-TMR estimator. The purpose of our work is to develop an UL channel estimation that not only has a good performance compared to other existing methods, but also requires a low computational complexity to implement.

Inspired by the use of the classification-based methods such as SVM, for instance in [14], as well as the promising underlying Toeplitz structure exploited in the SE-TMR method [15], we develop a computationally efficient UL channel estimation for one-bit ADCs. This new method for UL channel estimation for narrowband mmWave MIMO communications when the BS deploys one-bit ADCs is called ℓ_1 regularized logistic regression with Toeplitz matrix reconstruction (L1-RLR-TMR). A computationally efficient alternating direction method of multipliers (ADMM)-based implementation of the L1-RLR-TMR method is also developed. Numerical simulations are included to showcase the efficiency of the L1-RLR-TMR method compared to existing competitive methods.

Notations: Upper-case and lower-case bold-face letters denote matrices and vectors, respectively, while scalars are denoted by lower-case letters. The transpose, and Hermitian transpose are denoted by $\{\cdot\}^T$ and $\{\cdot\}^H$, respectively, while $\|\cdot\|_2$ and $\|\cdot\|_1$ are the ℓ_2 and ℓ_1 norms of a vector, and

$\|\cdot\|_F$ stands for the Frobenius norm of a matrix. The notation $\mathcal{T}(\cdot)$ stands for an operation of building a square Hermitian Toeplitz matrix with its first column being the bracketed vector. The $n \times n$ identity matrix is denoted by \mathbf{I}_n . The $n \times 1$ vector with all its entries equal to one is denoted by $\mathbf{1}_n$. The i th entry of the vector $\boldsymbol{\pi}$ is denoted by $[\boldsymbol{\pi}]_i$, while the entry in the intersection of the i th row and j th column of the matrix $\boldsymbol{\Pi}$ is denoted as $[\boldsymbol{\Pi}]_{i,j}$. The operator $\text{vec}\{\cdot\}$ stacks the columns of a matrix into a long vector, while $\text{unvec}\{\cdot\}$ forms a matrix by splitting the argument and putting them in the columns of that matrix. The operator $\text{diag}\{\boldsymbol{\pi}\}$ generates a diagonal matrix by plugging the entries of the vector $\boldsymbol{\pi}$ into its main diagonal, while $\text{blkdiag}\{\boldsymbol{\Pi}_1, \dots, \boldsymbol{\Pi}_n\}$ generates a block-diagonal matrix using the bracketed matrices. The notation $\boldsymbol{\Pi} \succeq 0$ means that $\boldsymbol{\Pi}$ is Hermitian positive semidefinite. The least non-negative remainder in the division of a by b is denoted by $\text{rem}(a, b)$. Finally, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ return the real and imaginary parts of the bracketed argument, respectively.

2. UL CHANNEL ESTIMATION

2.1. System Model

Consider an UL multi-user mmWave MIMO system where the BS deploys a uniform linear array (ULA) with M antenna elements that serves K single antenna users.¹ Each antenna of the BS array is connected to two one-bit ADCs for converting the real and imaginary parts of the received signals, while all users are equipped with high-resolution DACs. The UL channel between user k and the BS is given by

$$\mathbf{h}_k = \sum_{l=1}^{L_k} \sum_{m=1}^{M_{\text{path}}^{k,l}} \gamma_{k,l,m} \mathbf{a}(\theta_{k,l,m})$$

$$= [\mathbf{A}(\boldsymbol{\theta}_{k,1}), \dots, \mathbf{A}(\boldsymbol{\theta}_{k,L_k})] \begin{bmatrix} \gamma_{k,1} \\ \vdots \\ \gamma_{k,L_k} \end{bmatrix} = \mathbf{A}(\boldsymbol{\theta}_k) \boldsymbol{\gamma}_k \quad (1)$$

where L_k indicates the number of multipath clusters between the BS and the user k , the l th cluster consists of $M_{\text{path}}^{k,l}$ paths densely grouped around a mean direction with the corresponding angle spread [19], $\gamma_{k,l,m}$ and $\theta_{k,l,m}$ are respectively the gain and DOA of the m th path in the l th cluster, the steering vector is given by $\mathbf{a}(\theta_{k,l,m}) = [1, e^{-j\pi \sin(\theta_{k,l,m})}, \dots, e^{-j(M-1)\pi \sin(\theta_{k,l,m})}]^T \in \mathbb{C}^{M \times 1}$, $\boldsymbol{\theta}_{k,l} \triangleq [\theta_{k,l,1}, \dots, \theta_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{R}^{M_{\text{path}}^{k,l} \times 1}$ for $l = 1, \dots, L_k$, $\mathbf{A}(\boldsymbol{\theta}_{k,l}) \triangleq [\mathbf{a}(\theta_{k,l,1}), \dots, \mathbf{a}(\theta_{k,l,M_{\text{path}}^{k,l}})] \in \mathbb{C}^{M \times M_{\text{path}}^{k,l}}$, $\boldsymbol{\gamma}_{k,l} \triangleq [\gamma_{k,l,1}, \dots, \gamma_{k,l,M_{\text{path}}^{k,l}}]^T \in \mathbb{C}^{M_{\text{path}}^{k,l} \times 1}$, $\boldsymbol{\theta}_k \triangleq [\boldsymbol{\theta}_{k,1}^T, \dots, \boldsymbol{\theta}_{k,L_k}^T]^T$, $\mathbf{A}(\boldsymbol{\theta}_k) \triangleq [\mathbf{A}(\boldsymbol{\theta}_{k,1}), \dots, \mathbf{A}(\boldsymbol{\theta}_{k,L_k})]$, and $\boldsymbol{\gamma}_k \triangleq [\boldsymbol{\gamma}_{k,1}^T, \dots, \boldsymbol{\gamma}_{k,L_k}^T]^T$. Accordingly, the channel between the BS and K users is

given as

$$\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K] = [\mathbf{A}(\boldsymbol{\theta}_1) \boldsymbol{\gamma}_1, \dots, \mathbf{A}(\boldsymbol{\theta}_K) \boldsymbol{\gamma}_K]. \quad (2)$$

The users transmit a pilot sequence of length N_s ($N_s \geq K$) during the training stage, and the received signal at the BS is

$$\mathbf{Y} = \mathcal{Q}(\mathbf{H}\mathbf{S} + \mathbf{N}) \quad (3)$$

where $\mathcal{Q}(\cdot) \triangleq \text{sign}(\Re\{\cdot\}) + j \text{sign}(\Im\{\cdot\})$ denotes the element-wise one-bit quantizer which only preserves the sign of arguments and its output, i.e., takes values from the set $\mathcal{S} = \{1 + j, 1 - j, -1 + j, -1 - j\}$, $\mathbf{S} \in \mathbb{C}^{K \times N_s}$ represents the orthogonal pilot matrix, and $\mathbf{N} \in \mathbb{C}^{M \times N_s}$ is the additive circularly symmetric complex Gaussian noise with zero mean and variance σ^2 . The task is to restore (scaled) $\mathbf{H} \in \mathbb{C}^{M \times K}$ from the received signal $\mathbf{Y} \in \mathbb{C}^{M \times N_s}$.

2.2. Logistic Regression for Binary Classification

Consider a binary classification task where the goal is to separate a training data set $\mathcal{D} = \{(\mathbf{x}_p, y_p)\}_{p=1}^P$ with \mathbf{x}_p and $y_p \in \{\pm 1\}$ being a feature and a corresponding binary label, respectively. The logistic regression classifier separates the data space into two regions via finding vectors \mathbf{w} and \mathbf{b} that minimize

$$\min_{\mathbf{w}, \mathbf{b}} \sum_{p=1}^P \log \left(1 + e^{-y_p (\mathbf{w}^T \mathbf{x}_p)} \right) \quad (4)$$

where \mathbf{w} and \mathbf{b} are referred to as the weight vector and the bias, respectively [17, 18].

2.3. Proposed UL Channel Estimation

Because of the clustered-based representation of \mathbf{h}_k in (1), the channel between the BS and the k th user can be still considered to be sparse with respect to an angular-based dictionary in spite of being composed of many closely located paths. In other words, \mathbf{h}_k can be approximated by a linear combination of a few atoms of a proper angular-based dictionary (e.g., the normalized discrete Fourier transform (DFT)). We approximate each \mathbf{h}_k as a linear combination of the corresponding L_k basis paths with L_k DOAs and path gains.² Hence, (2) can be reformulated as

$$\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K] = [\mathbf{A}(\bar{\boldsymbol{\theta}}_1) \bar{\boldsymbol{\gamma}}_1, \dots, \mathbf{A}(\bar{\boldsymbol{\theta}}_K) \bar{\boldsymbol{\gamma}}_K] \quad (5)$$

where $\bar{\boldsymbol{\theta}}_k \triangleq [\bar{\theta}_{k,1}, \dots, \bar{\theta}_{k,L_k}]^T \in \mathbb{R}^{L_k \times 1}$ and $\bar{\boldsymbol{\gamma}}_k \triangleq [\bar{\gamma}_{k,1}, \dots, \bar{\gamma}_{k,L_k}]^T \in \mathbb{C}^{L_k \times 1}$ denote respectively the DOAs and path gains associating with L_k basis paths between the BS and user k . Manipulating (5), we obtain

$$\mathbf{H} = \mathbf{A} \mathbf{\Gamma} \bar{\mathbf{G}} = \bar{\mathbf{H}} \bar{\mathbf{G}} \quad (6)$$

¹The assumption of single antenna users is made for simplicity, while a more practical configuration is to consider users with multiple antennas.

²Note that increasing the number of atoms for approximating each \mathbf{h}_k leads to a more accurate modeling indeed, however, it also increases the computational complexity.

where $\mathbf{A} \triangleq [\mathbf{A}(\bar{\theta}_1), \dots, \mathbf{A}(\bar{\theta}_K)] \in \mathbb{C}^{M \times L}$, $\mathbf{\Gamma} \triangleq \text{diag}\{\bar{\gamma}_1^T, \bar{\gamma}_2^T, \dots, \bar{\gamma}_K^T\} \in \mathbb{C}^{L \times L}$, $\mathbf{G} \triangleq \text{blkdiag}\{\mathbf{1}_{L_1}, \mathbf{1}_{L_2}, \dots, \mathbf{1}_{L_K}\} \in \mathbb{R}^{L \times K}$, $\bar{\mathbf{H}} \triangleq \mathbf{A}\mathbf{\Gamma} \in \mathbb{C}^{M \times L}$, and $L \triangleq \sum_{k=1}^K L_k$. Thanks to \mathbf{A} and $\mathbf{\Gamma}$ being respectively Vandermonde and diagonal matrices, we have

$$\bar{\mathbf{H}}\bar{\mathbf{H}}^H = \mathcal{T}(\mathbf{u}) \quad (7)$$

where $\mathbf{u} \in \mathbb{C}^M$ and $[\mathbf{u}]_1$ is a real number according to (7). On the other hand, with the DOA-based structure of $\bar{\mathbf{H}}$ through \mathbf{A} , $\bar{\mathbf{H}}$ can be multiplied by a proper angular-based dictionary matrix like the normalized DFT matrix to make each column of the resultant sparse, i.e.,

$$\mathbf{X}(\bar{\theta}) = \mathbf{F}\bar{\mathbf{H}} = \mathbf{F}\bar{\mathbf{H}}\bar{\mathbf{G}} \quad (8)$$

where $\mathbf{F} \in \mathbb{C}^{M \times M}$ denotes the normalized DFT matrix.

Applying the vectorization operator to (3) together with the use of (6), we have

$$\mathbf{y} \triangleq \text{vec}\{\mathbf{Y}\} = \mathcal{Q}\left(\left((\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\right) \bar{\mathbf{h}} + \mathbf{n}\right) \quad (9)$$

where $\bar{\mathbf{h}} \triangleq \text{vec}\{\bar{\mathbf{H}}\}$ and $\mathbf{n} \triangleq \text{vec}\{\mathbf{N}\}$. For convenience in later derivations, the notation in (9) is converted to the real domain as

$$\mathbf{y}_R \triangleq [\Re\{\mathbf{y}\}^T, \Im\{\mathbf{y}\}^T]^T = \bar{\mathbf{S}}\bar{\mathbf{h}}_R \quad (10)$$

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

$$\begin{aligned} \bar{\mathbf{S}} &\triangleq \begin{bmatrix} \Re\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} & -\Im\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} \\ \Im\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} & \Re\{(\bar{\mathbf{G}}\mathbf{S})^T \otimes \mathbf{I}_M\} \end{bmatrix} \\ &= [\bar{\mathbf{s}}_1, \bar{\mathbf{s}}_2, \dots, \bar{\mathbf{s}}_{2MN_s}]^T. \end{aligned} \quad (11)$$

Note that $\mathbf{y}_R \in \{\pm 1\}^{2MN_s \times 1}$ and $\bar{\mathbf{h}}_R \in \mathbb{R}^{2ML \times 1}$. In addition, $\bar{\mathbf{h}}_m^T \in \mathbb{C}^{1 \times L}$ with $m \in \{1, \dots, M\}$ and $\bar{\mathbf{s}}_t^T \in \mathbb{R}^{1 \times 2ML}$ with $t \in \{1, \dots, 2MN_s\}$ denote the m th and t th rows of $\bar{\mathbf{H}}$ and $\bar{\mathbf{S}}$, respectively. Using (4), (7)-(8), and (10)-(11), the minimization problem associated with the L1-RLR-TMR method can be written as

$$\begin{aligned} \min_{\bar{\mathbf{h}}_R, \mathbf{u}} & \|\bar{\mathbf{F}}\bar{\mathbf{h}}_R\|_1 + \lambda \sum_{t=1}^{2MN_s} \log\left(1 + e^{-\kappa[\mathbf{y}_R]_t(\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)}\right) \\ \text{s.t.} & \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I)^H \\ \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I & \mathcal{T}(\mathbf{u}) \end{bmatrix} \succeq 0 \\ & \|\bar{\mathbf{h}}_m\|_2^2 = c, \quad m = 1, \dots, M \end{aligned} \quad (12)$$

where $\bar{\mathbf{F}} \triangleq \begin{bmatrix} \Re\{\mathbf{G}^T \otimes \mathbf{F}\} & -\Im\{\mathbf{G}^T \otimes \mathbf{F}\} \\ \Im\{\mathbf{G}^T \otimes \mathbf{F}\} & \Re\{\mathbf{G}^T \otimes \mathbf{F}\} \end{bmatrix}$, $\lambda > 0$ is a regularization parameter, and $\kappa > 1$ is added for accelerating the convergence of the ADMM-based implementation of the L1-RLR-TMR method in the next Subsection. Note that the two constrains of (12) are added to impose (7), where c is a tunable parameter. We suggest to set $c = 1$. Moreover, we set $\kappa = 10$. The minimization problem presented in (12) is non-convex because of the equality constraints.

2.4. The ADMM Implementation

To develop the ADMM-based method, an auxiliary variable \mathbf{Z} is first introduced in order to modify (12) as

$$\begin{aligned} \min_{\bar{\mathbf{h}}_R, \mathbf{u}} & \|\bar{\mathbf{F}}\bar{\mathbf{h}}_R\|_1 + \lambda \sum_{t=1}^{2MN_s} \log\left(1 + e^{-\kappa[\mathbf{y}_R]_t(\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)}\right) \\ \text{s.t.} & \mathbf{Z} = \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I)^H \\ \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I & \mathcal{T}(\mathbf{u}) \end{bmatrix} \\ & \mathbf{Z} \succeq 0 \\ & \|\bar{\mathbf{h}}_m\|_2^2 = c, \quad m = 1, \dots, M \end{aligned} \quad (13)$$

Consequently, the scaled augmented Lagrangian of (13) is expressed as

$$\begin{aligned} \mathcal{L}_\rho(\bar{\mathbf{h}}_R, \mathbf{u}, \mathbf{Z}, \mathbf{\Lambda}) &= \|\bar{\mathbf{F}}\bar{\mathbf{h}}_R\|_1 \\ &+ \lambda \sum_{t=1}^{2MN_s} \log\left(1 + e^{-\kappa[\mathbf{y}_R]_t(\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)}\right) \\ &+ \frac{\rho}{2} \left\| \mathbf{Z} - \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I)^H \\ \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I & \mathcal{T}(\mathbf{u}) \end{bmatrix} + \mathbf{\Lambda} \right\|_F^2 \end{aligned} \quad (14)$$

where $\rho > 0$ is a penalty parameter, and $\mathbf{\Lambda}$ is the dual variable. For convenience, \mathbf{Z} and $\mathbf{\Lambda}$ are partitioned as $\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_0 & (\mathbf{Z}_R + j\mathbf{Z}_I)^H \\ \mathbf{Z}_R + j\mathbf{Z}_I & \mathbf{Z}_1 \end{bmatrix}$, and $\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_0 & (\mathbf{\Lambda}_R + j\mathbf{\Lambda}_I)^H \\ \mathbf{\Lambda}_R + j\mathbf{\Lambda}_I & \mathbf{\Lambda}_1 \end{bmatrix}$. Therefore, the updating rules of the ADMM for solving (12) are

$$\begin{aligned} (\bar{\mathbf{h}}_R^{l+1}, \mathbf{u}^{l+1}) &= \arg \min_{\bar{\mathbf{h}}_R, \mathbf{u}} \mathcal{L}_\rho(\bar{\mathbf{h}}_R, \mathbf{u}, \mathbf{Z}^l, \mathbf{\Lambda}^l) \\ \text{s.t.} & \|\bar{\mathbf{h}}_m\|_2^2 = c, \quad m = 1, \dots, M \end{aligned} \quad (15)$$

$$\mathbf{Z}^{l+1} = \left[\mathbf{D}^{l+1} - \mathbf{\Lambda}^l \right]_+ \quad (16)$$

$$\mathbf{\Lambda}^{l+1} = \mathbf{\Lambda}^l + \mathbf{Z}^{l+1} - \mathbf{D}^{l+1} \quad (17)$$

where $\mathbf{D}^{l+1} \triangleq \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R^{l+1} + j\bar{\mathbf{H}}_I^{l+1})^H \\ \bar{\mathbf{H}}_R^{l+1} + j\bar{\mathbf{H}}_I^{l+1} & \mathcal{T}(\mathbf{u}^{l+1}) \end{bmatrix}$, and the notation $(\cdot)^l$ represents the estimates at l th iteration. In addition, $[\cdot]_+$ is a projection function used in (16) to project the argument onto the positive semidefinite cone via carrying out the eigenvalue decomposition and setting all the negative eigenvalues to zero. As it is presented in (15), a minimization problem has to be solved to estimates $\bar{\mathbf{h}}_R^{l+1}$ and \mathbf{u}^{l+1} . (15) is convex with respect to \mathbf{u} with the closed-form solution as

$$\mathbf{u}^{l+1} = \mathbf{W}(\mathcal{T}^*(\mathbf{Z}_1^l + \mathbf{\Lambda}_1^l)) \quad (18)$$

where $\mathbf{W} \triangleq \text{diag}\{\frac{1}{M}, \frac{1}{2(M-1)}, \dots, \frac{1}{2}\}^T$, and $\mathcal{T}^*(\cdot)$ denotes the Toeplitz adjoint operator. However, estimating $\bar{\mathbf{h}}_R^{l+1}$ can be handled using an inner ADMM formulation for solving

(15) with respect to $\bar{\mathbf{h}}_R$. In doing so, an auxiliary variable \mathbf{w} is introduced to modify (15) as

$$\begin{aligned} \bar{\mathbf{h}}_R^{l+1} = \arg \min_{\bar{\mathbf{h}}_R} & \|\mathbf{w}\|_1 + \lambda \sum_{t=1}^{2MN_s} \log \left(1 + e^{-\kappa[\mathbf{y}_R]_t (\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)} \right) \\ & + \frac{\rho}{2} \left\| \mathbf{Z}^l - \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I)^H \\ \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I & \mathcal{T}(\mathbf{u}^l) \end{bmatrix} + \mathbf{\Lambda}^l \right\|_F^2 \\ \text{s.t. } & \|\bar{\mathbf{h}}_m\|_2^2 = c, \quad m = 1, \dots, M \\ & \bar{\mathbf{F}}\bar{\mathbf{h}}_R - \mathbf{w} = \mathbf{0} \end{aligned} \quad (19)$$

Hence, its scaled augmented Lagrangian of (19) is given as

$$\begin{aligned} \mathcal{L}_{\bar{\rho}}(\bar{\mathbf{h}}_R, \mathbf{w}, \mathbf{v}) = & \|\mathbf{w}\|_1 + \lambda \sum_{t=1}^{2MN_s} \log \left(1 + e^{-\kappa[\mathbf{y}_R]_t (\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)} \right) \\ & + \frac{\rho}{2} \left\| \mathbf{Z}^l - \begin{bmatrix} \mathbf{I}_L & (\bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I)^H \\ \bar{\mathbf{H}}_R + j\bar{\mathbf{H}}_I & \mathcal{T}(\mathbf{u}^l) \end{bmatrix} + \mathbf{\Lambda}^l \right\|_F^2 \\ & + \frac{\bar{\rho}}{2} \|\bar{\mathbf{F}}\bar{\mathbf{h}}_R - \mathbf{w} + \mathbf{v}\|_2^2 \end{aligned} \quad (20)$$

where $\bar{\rho} > 0$ is a penalty parameter associated with the inner ADMM, and \mathbf{v} is the dual variable. In each outer ADMM iteration l , the inner ADMM update rules are

$$\begin{aligned} \bar{\mathbf{h}}_R^+ = \arg \min_{\bar{\mathbf{h}}_R} & \mathcal{L}_{\bar{\rho}}(\bar{\mathbf{h}}_R, \mathbf{w}^-, \mathbf{v}^-) \\ \text{s.t. } & \|\bar{\mathbf{h}}_m\|_2^2 = c, \quad m = 1, \dots, M \end{aligned} \quad (21)$$

$$\mathbf{w}^+ = \mathcal{S}_{1/\bar{\rho}}(\bar{\mathbf{F}}\bar{\mathbf{h}}_R^+ + \mathbf{v}^-) \quad (22)$$

$$\mathbf{v}^+ = \mathbf{v}^- + \bar{\mathbf{F}}\bar{\mathbf{h}}_R^+ - \mathbf{w}^+ \quad (23)$$

where $\mathcal{S}_\alpha(\cdot)$ is the soft thresholding operator. Note that the notation relations to l is omitted for simplicity in (21)–(23), and also the notation $(\cdot)^+$ is adopted to show the updates related to the inner ADMM iterations. Finally, we exploit the projected gradient descent (PGD) to handle the minimization problem (21) where the gradient of its objective with respect to $\bar{\mathbf{h}}_R$ is given as

$$\begin{aligned} \nabla \mathcal{L}_{\bar{\rho}}(\bar{\mathbf{h}}_R) = & -\lambda \sum_{t=1}^{2MN_s} \frac{\kappa[\mathbf{y}_R]_t}{1 + e^{\kappa[\mathbf{y}_R]_t (\bar{\mathbf{s}}_t^T \bar{\mathbf{h}}_R)}} \bar{\mathbf{s}}_t + 2\rho(\bar{\mathbf{h}}_R - \mathbf{q}^l) \\ & + \bar{\rho}(\bar{\mathbf{F}}^T \bar{\mathbf{F}}\bar{\mathbf{h}}_R - \bar{\mathbf{F}}^T(\mathbf{w}^- - \mathbf{v}^-)) \end{aligned} \quad (24)$$

where $\mathbf{q}^l \triangleq [\text{vec}\{\mathbf{Z}_R^l + \mathbf{\Lambda}_R^l\}^T, \text{vec}\{\mathbf{Z}_I^l + \mathbf{\Lambda}_I^l\}^T]^T$. It is worth mentioning that we initialize the aforementioned PGD by the maximum ratio estimate [11] as

$$\bar{\mathbf{h}}_R^{(0)} = \frac{\tilde{\mathbf{X}}_R^T \mathbf{1}_{2MN}}{\|\tilde{\mathbf{X}}_R^T \mathbf{1}_{2MN}\|_2} \sqrt{M^c} \quad (25)$$

where $\tilde{\mathbf{X}}_R \triangleq \text{diag}\{\mathbf{y}_R\}\bar{\mathbf{S}}$. The steps of the proposed UL channel estimator are outlined in Algorithm 1. The following

definitions of the primal and dual variables are used for the ADMM implementation [22] in Algorithm 1:

$$\begin{aligned} \mathbf{x}^{l+1} & \triangleq \left[(\bar{\mathbf{h}}_R^{l+1})^T, \text{vec}\{\Re\{\mathcal{T}(\mathbf{u}^{l+1})\}\}^T, \right. \\ & \left. \text{vec}\{\Im\{\mathcal{T}(\mathbf{u}^{l+1})\}\}^T \right]^T, \quad \mathbf{z}^{l+1} \triangleq \left[\text{vec}\{\mathbf{Z}_R^{l+1}\}^T, \right. \\ & \left. \text{vec}\{\mathbf{Z}_I^{l+1}\}^T, \text{vec}\{\Re\{\mathbf{Z}_1^{l+1}\}\}^T, \text{vec}\{\Im\{\mathbf{Z}_1^{l+1}\}\}^T \right]^T \\ \mathbf{y}^{l+1} & \triangleq \rho \left[\text{vec}\{\mathbf{\Lambda}_R^{l+1}\}^T, \text{vec}\{\mathbf{\Lambda}_I^{l+1}\}^T, \text{vec}\{\Re\{\mathbf{\Lambda}_1^{l+1}\}\}^T, \right. \\ & \left. \text{vec}\{\Im\{\mathbf{\Lambda}_1^{l+1}\}\}^T \right]^T, \\ \epsilon_{\text{pri}}^{l+1} & \triangleq \sqrt{2M(L+M)}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|\mathbf{x}^{l+1}\|_2, \|\mathbf{z}^{l+1}\|_2\} \\ \epsilon_{\text{dual}}^{l+1} & \triangleq \sqrt{2M(L+M)}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \|\mathbf{y}^{l+1}\|_2 \end{aligned}$$

Note that in the steps related to the PGD iterations in Algorithm 1, the notation \bar{h} is adopted for updating \bar{h}_R to avoid confusion. Moreover, the number of iterations is limited to 1000 for the outer ADMM.

3. SIMULATION RESULTS

The proposed L1-RLR-TMR method is evaluated and compared to other competitive algorithms in this section. The nML [11], AR [13], and SE-TMR [15] methods are used for comparison. The pilot sequence is constructed 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., $\mathbf{S}\mathbf{S}^H = N_s \mathbf{I}_K$. The SNR and normalized mean square error (NMSE) are respectively defined as $\text{SNR} \triangleq 10 \log_{10} \left(\frac{\|\mathbf{H}\mathbf{S}\|_2^2}{MN_s \sigma^2} \right)$ and

$$\text{NMSE} \triangleq \frac{1}{KN} \sum_{k=1}^K \sum_{n=1}^N \left\| \frac{\hat{\mathbf{h}}_k^{(n)}}{\|\hat{\mathbf{h}}_k^{(n)}\|_2} - \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|_2} \right\|_2^2, \quad \text{where}$$

$\hat{\mathbf{h}}_k^{(n)}$ stands for the k th column of $\hat{\mathbf{H}}$ estimated in the n th Monte Carlo run with \mathbf{h}_k being the actual k th column of \mathbf{H} , and N is the total number of Monte Carlo trials considered as $N = 200$ here. We consider $\lambda = 1$ for the SE-TMR and L1-RLR-TMR methods, and $K = 8$. We set the number of channel clusters and within cluster multipaths for all users to be the same, i.e., $L_1 = \dots = L_K$ and $M_{\text{path}}^{1,1} = \dots = M_{\text{path}}^{1,L_1} = \dots = M_{\text{path}}^{K,1} = \dots = M_{\text{path}}^{1,L_K} = 100$. UL DOAs are generated randomly once and remain the same throughout all Monte Carlo trials, and the channel path gains are distributed as $\mathcal{CN}(0, 1)$. Fig. 1 depicts the NMSE of the methods tested versus SNR for the scenario that $M = 16$, $N_s = 128$, $L_k = 1$ for all users, and the angle spread of 8 degrees within each cluster. It shows that 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 [21]. In Fig. 2, the performance of the methods tested is presented 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 compared to other methods tested. Particularly, Fig. 2 shows that the

Algorithm 1: ADMM implementation for L1-RLR-TMR

Initialization

- 1: Set the initial $\bar{\mathbf{h}}_R^{(0)}$ as (25). In addition, set $\lambda = 1$, $\rho = 10$, $\bar{\rho} = 1$, $c = 1$, and $\kappa = 10$.
- 2: Set $i_{\max} = 10$, the step size $\beta = 0.01$ and the termination threshold $\epsilon = 10^{-3}$ for the PGD. In addition, set the termination thresholds of the outer ADMM as $\epsilon^{\text{abs}} = \epsilon^{\text{rel}} = 10^{-4}$.
- 3: Set

$$\mathbf{Z}_0^0 = \mathbf{I}_M, \mathbf{Z}_R^0 = \bar{\mathbf{H}}_R^0 = \text{unvec} \left\{ \left[\bar{\mathbf{h}}_R^{(0)} \right]_{1:ML} \right\},$$

$$\mathbf{Z}_I^0 = \bar{\mathbf{H}}_I^0 = \text{unvec} \left\{ \left[\bar{\mathbf{h}}_R^{(0)} \right]_{ML+1:2ML} \right\}, \mathbf{w}^- = \bar{\mathbf{F}} \bar{\mathbf{h}}_R^{(0)},$$

$$\mathbf{u}^0 = \mathbf{W} (\mathcal{T}^* (\bar{\mathbf{H}}_R^0 + j \bar{\mathbf{H}}_I^0)), \mathbf{Z}_I^0 = \mathcal{T}(\mathbf{u}^0), \bar{\mathbf{h}}_R^- = \bar{\mathbf{h}}_R^{(0)}.$$
- 4: Initialize Λ^0 and \mathbf{v}^- as all zero matrix and vector, respectively. In addition, set $l = 0$.

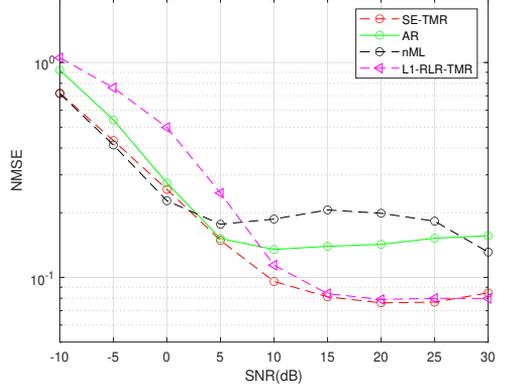
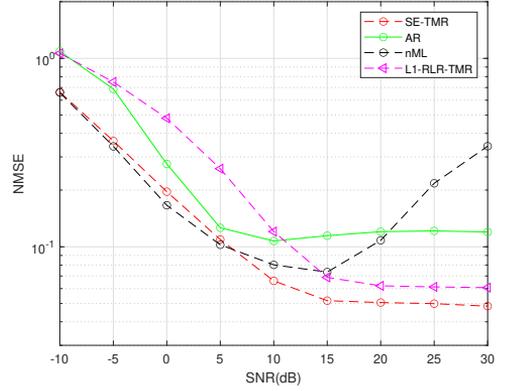
Outer ADMM

- 5: **while** $e_{\text{pri}}^{l+1} \geq \epsilon_{\text{pri}}^{l+1}$ & $e_{\text{dual}}^{l+1} \geq \epsilon_{\text{dual}}^{l+1}$

Inner ADMM

- 6: **for** $i = 1 : i_{\max}$
 - 7: Reset $k = 1$.
 - 8: **while** $\|\tilde{\mathbf{h}}^{(k)} - \tilde{\mathbf{h}}^{(k-1)}\|_2 \geq \epsilon \|\tilde{\mathbf{h}}^{(k-1)}\|_2$
 - 9: $\tilde{\mathbf{h}}^{(k)} = \tilde{\mathbf{h}}^{(k-1)} - \beta \nabla \mathcal{L}_\rho(\tilde{\mathbf{h}}^{(k-1)})$ via (24).
 - 10: $\tilde{\mathbf{H}}_1 \triangleq \text{unvec} \left\{ \left[\tilde{\mathbf{h}}^{(k)} \right]_{1:ML} \right\}$.
 - 11: $\tilde{\mathbf{H}}_2 \triangleq \text{unvec} \left\{ \left[\tilde{\mathbf{h}}^{(k)} \right]_{ML+1:2ML} \right\}$
 - 12: $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}_1 + j \tilde{\mathbf{H}}_2$.
 - 13: **for** $m = 1 : M$
 - 14: $[\tilde{\mathbf{H}}]_{m,:} = \frac{[\tilde{\mathbf{H}}]_{m,:}}{\|[\tilde{\mathbf{H}}]_{m,:}\|_2} \sqrt{c}$
 - 15: **end for**
 - 16: $\tilde{\mathbf{h}}^{(k)} = [\text{vec}\{\Re\{\tilde{\mathbf{H}}\}\}^T, \text{vec}\{\Im\{\tilde{\mathbf{H}}\}\}^T]^T$
 - 17: **if** $\text{rem}(k,10) = 0$
 - 18: $\beta = \frac{\beta}{2}$
 - 19: **end if**
 - 20: **end while**
 - 21: Set $\bar{\mathbf{h}}_R^+ = \tilde{\mathbf{h}}^{(k)}$. Update \mathbf{w}^+ and \mathbf{v}^+ via (22) and (23), respectively.
 - 22: **end for**
 - 23: Set $\bar{\mathbf{h}}_R^{l+1} = \bar{\mathbf{h}}_R^+$. Update \mathbf{u}^{l+1} , \mathbf{Z}^{l+1} , and Λ^{l+1} using (18), (16), and (17), respectively.
 - 24: Construct \mathbf{x}^{l+1} , \mathbf{z}^{l+1} , \mathbf{y}^{l+1} , $\epsilon_{\text{pri}}^{l+1}$, and $\epsilon_{\text{dual}}^{l+1}$.
Then, calculate

$$e_{\text{pri}}^{l+1} = \|\mathbf{x}^{l+1} - \mathbf{z}^{l+1}\|_2, e_{\text{dual}}^{l+1} = \|\rho(\mathbf{z}^{l+1} - \mathbf{z}^l)\|_2$$
 - 25: **if** $e_{\text{pri}}^{l+1} > 2 e_{\text{dual}}^{l+1}$
 - 26: $\rho = 2\rho$ and $\Lambda^{l+1} = \frac{1}{2} \Lambda^{l+1}$
 - 27: **else if** $e_{\text{dual}}^{l+1} > 2 e_{\text{pri}}^{l+1}$
 - 28: $\rho = \frac{\rho}{2}$ and $\Lambda^{l+1} = 2\Lambda^{l+1}$
 - 29: **end if**
 - 30: **end while**
 - 31: Reshape $\bar{\mathbf{H}}$, and recover \mathbf{H} using (6).
-


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

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

performance of L1-RLR-TMR implemented by the ADMM is comparable with that of the SE-TMR implemented using CVX.

4. CONCLUSION

A novel method called L1-RLR-TMR is proposed for estimating UL channels in mmWave multi-user MIMO communications when the BS uses one-bit ADCs. The main idea of the L1-RLR-TMR method is to recover the UL channel as the solution of an optimization problem designed by considering the ℓ_1 regularized logistic regression classification as well as the notion of Toeplitz matrix reconstruction. To make the use of L1-RLR-TMR method practical, we developed an ADMM-based implementation for it with low computational cost. Numerical results validate the efficiency of the L1-RLR-TMR method compared to other competitive methods.

5. REFERENCES

- [1] 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.
- [2] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Millimeter-wave massive MIMO: The next wireless revolution?," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [3] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolio, "Five disruptive technology directions for 5G," *IEEE Comm. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.
- [4] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roth, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [5] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.
- [6] A. Mezghani, F. Antreich, and J. A. Nossek, "Multiple parameter estimation with quantized channel output," in *Proc. Int. ITG Workshop Smart Antennas*, Bremen, Germany, Feb. 2010, pp. 143–150.
- [7] C. Mollen, J. Choi, E. G. Larsson, and R. W. Heath Jr., "Uplink Performance of Wideband Massive MIMO With One-Bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Jan. 2017.
- [8] C. Wen, C. Wang, S. Jin, 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.
- [9] 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.
- [10] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, Jul. 2018.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] M. Esfandiari, S. A. Vorobyov, and R. W. Heath, "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.
- [16] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. 216, 1952.
- [17] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: Data mining, inference, and prediction*. Springer, 2009.
- [18] A. Jung, *Machine learning: The basics*. Springer Nature, 2022.
- [19] "5G channel model for bands up to 100 GHz (v2.3)," Tech. Rep., 2016. [Online]. Available: <http://www.5gworkshops.com/5gcm.html>
- [20] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial-and-frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3393–3406, Jul. 2018.
- [21] M. Grant, S. Boyd, and Y. Ye, CVX: Matlab software for disciplined convex programming. 2008 [Online]. Available: <http://stanford.edu/~boyd/cvx>
- [22] 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.

Publication 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.

© 18 pages, submitted to a journal, April 2023
Reprinted with permission.

A Unified Approach to DOA Estimation in Unknown Noise Fields Using ULA

Majdoddin Esfandiari, Student Member, IEEE
Aalto University, Espoo, Finland

Sergiy A. Vorobyov, Fellow, IEEE
Aalto University, Espoo, Finland

Abstract— The problem of direction-of-arrival (DOA) estimation using uniform linear array in the presence of unknown noise fields is considered, a novel unified approach is developed, and its use for the cases of nonuniform and block-diagonal sensor noise are demonstrated. The approach consists of three phases. First, the noise covariance matrix is estimated using an alternative procedure. Second, a forward-only rooting-based DOA estimator as well as its forward-backward extension are developed for DOA estimation. The DOA estimators take advantage of using second-order statistics of signal subspace perturbation in constructing a weight matrix of a properly designed generalized least squares minimization problem. Despite the fact that these DOA estimators are iterative, only few iterations are sufficient to reach accurate results. The asymptotic performance of these DOA estimators is also investigated. Third, a newly designed DOA selection strategy with reasonable computational cost is developed to select L actual sources out of $2L$ candidates generated at the second phase. Numerical simulations are conducted in order to establish the significant superiority of the proposed approach over the existing state-of-the-art methods, especially in challenging scenarios, in uniform, nonuniform, and block-diagonal sensor noise.

Index Terms—DOA estimation, subspace method, nonuniform noise, block-diagonal noise, generalized least squares (GLS), small sample size.

I. Introduction

DIRECTION-of-arrival (DOA) estimation is an active field of research for decades due to multiple traditional and new important applications and due to significance of source localization in many practical scenarios. The notable current applications of interest are, for example, wireless communication, automotive radar, and sonar where DOA estimation is an essential task [1]–[7].

In the context of DOA estimation, several assumptions can be regarded concerning the structure of the second-order statistics of the observation noise. Most common

(Corresponding author: S. A. Vorobyov).

M. Esfandiari and S. A. Vorobyov are with Department of Information and Communications Engineering, Aalto University, PO Box 15400, 00076 Aalto, Finland (e-mails: majdoddin.esfandiari@aalto.fi; sergiy.vorobyov@aalto.fi).

0018-9251 © IEEE

assumptions are uniform white and nonuniform white. Moreover, a spatially block-correlated noise assumption may be more accurate in some applications [8].

For the uniform white noise assumption, i.e., when the noise powers are identical across all array sensors, many well-known subspace-based methods such as MUSIC [9]–[11], root-MUSIC [12]–[14], and ESPRIT [15], [16] have been developed, and are based on decomposing the signal or its sample covariance matrix (SCM) into two disjoint subspaces of noise and signal. The popularity of the subspace-based methods over, for example, near optimal maximum likelihood (ML) estimators [17]–[19], comes from the fact that they achieve good estimation accuracy along with affordable computational complexity in contrast to the prohibitive computational cost of implementing ML estimators.

In recent years, several works have proposed competitive algorithms to fill the aforementioned performance gap between subspace-based methods and ML estimators including root-swap root-MUSIC [20], enhanced principal-singular-vector utilization for modal analysis (EPUMA) [21], standard ESPRIT using generalized least squares (SE GLS) and unitary ESPRIT using generalized least squares (UE GLS) [22], partial relaxation (PR)-based approaches [23], [24], root-clustering algorithm and root-certificate algorithm [25]. The aim of the aforementioned methods is to achieve an adequate performance in challenging scenarios like scarcity of available data samples, low signal-to-noise ratio (SNR), and/or presence of some correlated or even coherent sources. Noteworthy to mention that most of the subspace-based methods can be revised to use forward-backward averaging (FBA) [28], and forward-backward spatial smoothing (FBSS) [29] techniques for coping with the cases of correlated or coherent signals.

In this context, in [20], a new concept called *root swap* has been introduced and recognized as the main reason behind the collapse of the root-MUSIC algorithm in finding the correct roots in the challenging environments. To remedy this phenomena, the root-swap root-MUSIC was proposed which identifies the correct roots associated with the sources' DOAs by exploiting the deterministic ML (DML) or stochastic ML (SML) [26] objective functions instead of deciding based on the closeness of the absolute values of the roots to unity. In [25], a new criterion for root selection based on the algebraic structure of the noise subspace has been proposed. Using this criterion and the relationship between the source localization problem and the problem of computing the approximate greatest common divisor (GCD) for polynomials, algorithms that learn the number of sources and estimate their locations have been proposed.

Furthermore, in [21], the authors have developed a method called EPUMA via solving a particular generalized least squares (GLS) problem by taking into account the second-order statistics of the estimated signal subspace, and also producing more DOA candidates than the number of sources and then selecting final DOAs

using DML or SML cost functions. Similar to EPUMA, the authors of [22] have extended ESPRIT and unitary ESPRIT to their generalized versions by exploiting the signal subspace perturbation as the source of error in the shift invariance equation (SIE). The aforementioned error has been minimized using the GLS. Adopting the main idea of [22], enhanced ESPRIT-based methods have been developed in [27], which first produces $2L$ DOA candidates for L sources and then selects the final L DOAs. Recently, the PR framework has been introduced in [23] to relax the manifold structure of sources partially, resulting in four methods by applying the PR concept to four previously presented in the literature DOA estimation techniques.

The uniform noise assumption can be, however, violated in some practical scenarios, which appear increasingly more often, and the case of nonuniform white noise has drawn considerable attention in the last two decades. The deterministic ML estimator along with Cramer-Rao bound (CRB) for both deterministic and stochastic source models have been proposed in [30], while [31] have developed the stochastic ML algorithm for the case of nonuniform noise. A simple method has been devised in [32] possessing less computational cost than ML estimators as well as improving the precision of the DOA estimates.

In [33], the authors have developed two iterative methods referred to as iterative ML subspace estimation (IMLSE) and iterative least squares subspace estimation (ILSSE), which estimate the signal subspace and noise covariance matrices based on the ML and least squares (LS) criteria, respectively. Moreover, it has been shown in [34] that the signal and noise subspaces are separable via applying the eigendecomposition (ED) of the so-called reduced covariance matrix for the case of uncorrelated sources, while a rank minimization approach has been suggested for the correlated sources case. The performance degradation caused by correlated and/or coherent sources can also be mitigated via using spatial smoothing [35], and covariance matrix differencing [36]. Recently, a method referred to as non-iterative subspace-based (NISB) [42] method has been proposed where the signal subspace and noise covariance matrix are identified by exploiting a two steps approach with the first step consisting of using the ED of the reduced covariance matrix [34] for estimating noise covariance matrix judiciously, followed by applying the generalized eigendecomposition (GED) of the matrix pair of the SCM and estimated noise covariance matrix from the first step.

In addition, bearing in mind the fact that the desired spatial directions can be modeled via sparse representation, sparse signal reconstruction (SSR) based methods have been developed [37]–[41]. Particularly, the so-called “*hyperparameter-free sparse estimation*” approaches, namely the sparse iterative covariance-based estimation method (SPICE), likelihood-based estimation of sparse parameters (LIKES), sparse learning via iterative minimization (SLIM) and iterative adaptive approach

(IAA) [39] have drawn much attention due to having good statistical properties along with not requiring any hyperparameters to tune.

There are also several colored noise modeling-based ML DOA estimation techniques [43]–[46]. It is known that the analytic concentration of the log-likelihood (LL) function with respect to all noise nuisance parameters is impossible [8]. Moreover, because of the large number of nuisance parameters, the numeric concentration is challenging. In [43] and [44], the noise is modeled using auto-regressive (AR) process, whereas in [45], it is assumed that the noise can be parameterized using a small number of Fourier coefficients. Another ML-based approach that does not require any structural assumptions on the received signals and the noise covariance matrix was proposed in [47]. However, it uses an inconsistent ad hoc estimate of the noise covariance matrix.

In this paper, we propose a novel unified procedure for addressing the DOA estimation problem using uniform linear array (ULA) in the presence of unknown (nonuniform and also block-diagonal) sensor noise that copes with the limitations of existing methods, such as high sensitivity to the presence of correlated sources, requirement of prohibitive computation complexity, poor performance in scenarios with closely located sources, small sample size, and low SNR. The procedure consists of three phases adjusted to each other, and it provides superior performance compared to the existing state-of-the-art methods. In the first phase, the noise covariance matrix is estimated. Indeed, a proper estimation of this matrix gives the necessary information enabling us to facilitate the DOA estimation task by diminishing the detrimental impact of the sensor noise effectively. Although there exist numerous DOA estimators such as MUSIC, ESPRIT, root-MUSIC, ML and their variations, the motivation of the second phase of our unified procedure is to design a novel method that employs the available information about the noise covariance matrix more wisely than it can be done by simply using a straightforward combination of noise covariance matrix estimation with the traditional DOA estimation methods. It leads to estimating unknown DOAs with higher precision. Then the DOA estimator from the second phase is exploited (twice) to generate (double) DOA candidates. Along with the DOA selection strategy from the third phase, it helps us to attain a better performance. At last, the DOA selection strategy selects the best final DOA estimates out of the candidates generated in the second phase. The main contributions of the paper are summarized in the next subsection.

A. Main Contributions

1. A novel unified approach that consists of three key phases is devised for addressing the problem of DOA estimation using ULA in unknown noise fields. The proposed procedure provides reliable estimates in scenarios with small sample size and/or relatively low SNR and/or closely located sources. The joint design of three phases

is matched to the challenges that need to be addressed for developing an effective DOA estimator. Specifically, in unknown noise field, the unknown noise covariance matrix needs to be estimated and then DOA estimator that exploits all the available information needs to be developed. We address the unknown noise covariance estimation problem in the first phase, and the DOA is estimated in the second and third phases as it is detailed in the following main contributions.

2. In the first phase, the nonuniform noise covariance matrix is estimated by alternations via obtaining the noise subspace using the GED of two matrices, followed by updating the noise covariance matrix using LS. The proposed noise covariance matrix estimator is applicable to sensor arrays with arbitrary geometry. Indeed, the proposed noise covariance matrix estimator can be utilized alongside other DOA estimation frameworks than the proposed approach in the second phase described below, which is specific for ULA, but have low computational complexity. We also extend the concept proposed for the nonuniform noise covariance estimation to develop an estimator of the block-diagonal noise covariance matrix.

3. Using the noise covariance matrix estimate, in the second phase, the DOA estimation problem is converted into the uniform noise DOA estimation problem by means of pre-whitening. Then a novel DOA estimation approach is designed by taking into consideration the signal subspace perturbation in conjunction with the use of the discrete Fourier transform (DFT) for both developing a new rooting-based method and also selecting the active equations for generating DOA candidates. The method proposed in phase 2 utilizes the GLS for estimating the coefficients of the desired polynomial. Moreover, to deal with possibly correlated sources, we derive the FB extension of the core DOA estimator. The motivation of phase 2 is to generate double DOA candidates judiciously to assist in improving the resolution of the DOA estimation in conjunction with phase 3, especially in the cases of small sample size and low SNR.

4. In the last phase, a double number of DOA candidates is generated, and a DOA selection strategy is developed and used to pick the final DOAs. The objective of the third phase is to judiciously select the final DOAs by the assist of the conventional beamformer (CB) in tandem with the ML optimization. In particular, the role of the CB is to clean outliers that may exist, while the ML optimization is used to ensure the selection of the best final DOAs.

5. It is demonstrated by conducting extensive numerical simulations for various challenging setups that the proposed DOA estimation procedure (i.e., employing the three phases combined) provides superior estimation accuracy for the uniform, nonuniform, and block-diagonal noise cases compared to the existing state-of-the-art techniques.

B. Paper Organization

The rest of the paper is organized as follows. The signal model and problem formulation are given in Section II. In Section III, alternative methods for estimating the sensor noise covariance matrix are devised. The forward-only subspace-based DOA method as well as its FB extension are developed in Section IV. Moreover, the asymptotic performance of the proposed DOA estimator is studied and a new DOA selection strategy is also designed to select the final DOAs in the section. Numerical simulations are provided in Section V. The paper is concluded in Section VI.

C. Notations

Upper-case and lower-case bold-face letters denote matrices and vectors, respectively, while scalars are denoted by lower-case letters. The expectation, transpose, conjugate, Hermitian transpose, and Moore-Penrose pseudo-inverse are denoted by $\mathbb{E}\{\cdot\}$, $\{\cdot\}^T$, $\{\cdot\}^*$, $\{\cdot\}^H$, and $\{\cdot\}^\dagger$, respectively, while $\|\cdot\|_2$, $\|\cdot\|_F$, and $|\cdot|$ denote the Euclidean norm of a vector, the Frobenius norm of a matrix, and the absolute value of a scalar. If the argument is a set, $|\cdot|$ denotes the set cardinality. The Kronecker product is denoted by \otimes and $\text{trace}\{\cdot\}$ stands for the trace of a square matrix. The $n \times n$ identity and exchange matrices are denoted by \mathbf{I}_n and \mathbf{J}_n , respectively. The $n \times m$ matrix with all elements equal zero and $n \times 1$ zero vector are denoted as $\mathbf{0}_{m \times n}$ and $\mathbf{0}_n$, respectively. The i th entry of the vector $\boldsymbol{\pi}$ is denoted by $[\boldsymbol{\pi}]_i$. The i th row and i th column of the matrix $\mathbf{\Pi}$ are denoted by $[\mathbf{\Pi}]_{i,:}$ and $[\mathbf{\Pi}]_{:,i}$, respectively, while the entry in the intersection of the i th row and j th column is denoted as $[\mathbf{\Pi}]_{i,j}$. The operator $\text{vec}\{\cdot\}$ stacks the columns of a matrix into a long vector, while $\text{unvec}\{\cdot\}$ forms a matrix by splitting the bracketed vector and putting them in the columns of that matrix. The operator $\text{diag}\{\boldsymbol{\pi}\}$ generates a diagonal matrix by plugging the entries of the vector $\boldsymbol{\pi}$ into its main diagonal, while the operator $\mathcal{D}\{\mathbf{\Pi}\}$ creates a diagonal matrix by preserving the main diagonal of the matrix $\mathbf{\Pi}$ and setting all other entries to zero. The block-diagonal matrix is denoted as $\text{bdiag}\{\cdot\}$. The operator $\text{DFT}\{\boldsymbol{\pi}\}$ stands for the DFT of the vector $\boldsymbol{\pi}$, while $\Re\{\cdot\}$ returns the real part of the bracketed argument.

II. SIGNAL MODEL AND UNIFIED SCHEME

Consider a ULA composed of M omni-directional sensors receiving L ($L < M$, and L is assumed to be known) narrowband signals radiated by L sources. It is assumed that the sources are located in the far-field and have distinguished directions, denoted as θ_l , $l = 1, \dots, L$. Then, the signal observed by the sensor array at the time instant t is written as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{s}(t) \triangleq [s_1(t) \dots s_L(t)]^T \in \mathbb{C}^L$ denotes the source signals, $\mathbf{n}(t) \in \mathbb{C}^M$ is the sensor noise vector, the source DOAs are stacked in the vector $\boldsymbol{\theta} \triangleq$

$[\theta_1 \dots \theta_L]^T$, $\mathbf{A}(\boldsymbol{\theta}) \triangleq [\mathbf{a}(\theta_1) \dots \mathbf{a}(\theta_L)]$ denotes the array manifold whose l th column is the steering vector $\mathbf{a}(\theta_l) = [1 \ e^{-j2\pi\sin(\theta_l)d/\lambda} \ \dots \ e^{-j2\pi(M-1)\sin(\theta_l)d/\lambda}]^T \in \mathbb{C}^M$ associated with l th DOA. Here λ denotes the carrier wavelength and $d = \lambda/2$. For notation simplicity, \mathbf{A} is used instead of $\mathbf{A}(\boldsymbol{\theta})$ hereafter, unless the argument of \mathbf{A} is different from $\boldsymbol{\theta}$.

The array covariance matrix can be written as

$$\mathbf{R} \triangleq \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q} \quad (2)$$

where $\mathbf{P} \in \mathbb{C}^{L \times L}$ and $\mathbf{Q} \in \mathbb{R}^{M \times M}$, respectively, stand for the signal and noise covariance matrices defined as

$$\mathbf{P} \triangleq \mathbb{E}\{\mathbf{s}(t)\mathbf{s}^H(t)\}, \quad \mathbf{Q} \triangleq \mathbb{E}\{\mathbf{n}(t)\mathbf{n}^H(t)\}. \quad (3)$$

The critical part of most of the methods designed for the nonuniform and block-diagonal noise cases is to estimate \mathbf{Q} efficiently, and the source DOA estimates are dependent on how precise the estimate of \mathbf{Q} is. Moreover, when the number of available snapshots is small and/or some of the DOAs are closely located to each other and/or SNR is relatively low, the impact of estimated \mathbf{Q} is more notable.

Because \mathbf{R} is unknown in practice, the SCM is considered, and it is given by $\hat{\mathbf{R}} \triangleq \frac{1}{N} \sum_{t=1}^N \mathbf{x}(t)\mathbf{x}^H(t) = \frac{1}{N} \mathbf{X}\mathbf{X}^H$. Here, the matrix signal notation is also used

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (4)$$

with $\mathbf{X} \triangleq [\mathbf{x}(1) \dots \mathbf{x}(N)]$, $\mathbf{S} \triangleq [\mathbf{s}(1) \dots \mathbf{s}(N)]$, $\mathbf{N} \triangleq [\mathbf{n}(1) \dots \mathbf{n}(N)]$, and N being the number of snapshots.

We stress that the proposed framework consists of three phases. The first phase is dedicated to estimating the unknown noise covariance matrix. The second phase employs the estimated noise covariance matrix from the first phase to pre-whiten the data and then generates double DOA candidate estimates. The third phase selects the best DOA estimates from the double DOA candidates generated in the second phase. These are three key phases in designing a DOA estimator in unknown noise fields, and they should be designed jointly to thoroughly exploit all the available information at every phase of the procedure.

III. NOISE COVARIANCE MATRIX ESTIMATION

A. Nonuniform Sensor Noise

Considering the case of nonuniform sensor noise that is spatially and temporally uncorrelated and zero-mean Gaussian, the noise covariance matrix is $\mathbf{Q} = \text{diag}\{\sigma_1^2, \dots, \sigma_M^2\}$ where σ_m^2 , $m = 1, \dots, M$ are the noise variances, which are not necessarily identical, i.e., $\sigma_i^2 \neq \sigma_j^2$ for $i \neq j$. It is desirable that a method estimating \mathbf{Q} would not require any knowledge of the true DOAs, while providing an acceptable accuracy with an affordable computational cost. Moreover, such method should be robust in extreme scenarios. Some examples of extreme scenarios are small sample size and presence of closely located sources. To devise such a method, we begin by multiplying both sides of (2) by $\mathbf{U} \in \mathbb{C}^{M \times (M-L)}$ which

satisfies the following condition

$$\mathbf{A}^H \mathbf{U} = \mathbf{0}_{L \times (M-L)}. \quad (5)$$

It is clear then that the columns of \mathbf{U} constitute a basis for the noise subspace. As \mathbf{A} is unknown, finding \mathbf{U} for the case of nonuniform noise is not as simple as for the uniform noise case, when \mathbf{U} is estimated by just calculating the eigenvectors of $\hat{\mathbf{R}}$. Multiplying (2) by \mathbf{U} , and exploiting (5), we have

$$\mathbf{R}\mathbf{U} = \mathbf{Q}\mathbf{U}. \quad (6)$$

Since \mathbf{R} is unknown, (25) can be expressed for $\hat{\mathbf{R}}$ as

$$\hat{\mathbf{R}}\mathbf{U} \approx \mathbf{Q}\mathbf{U}. \quad (7)$$

As both \mathbf{Q} and \mathbf{U} are unknown, it is natural to use the alternating minimization approach to find them, i.e., alternatively update one matrix at a time while the other is set as the estimate calculated in the previous iteration.

Recalling that the GED of any two square matrices $\{\mathbf{G}, \mathbf{Y}\}$ is defined as

$$\mathbf{G}\mathbf{F} = \mathbf{Y}\mathbf{D} \quad (8)$$

where the columns of \mathbf{F} are the eigenvectors and \mathbf{D} is a diagonal matrix which consists of the eigenvalues. Therefore, given an estimate of \mathbf{Q} in the i th alternation, which is denoted by $\hat{\mathbf{Q}}^{(i)}$, we write (7) in the form of (8) as

$$\hat{\mathbf{R}}\hat{\mathbf{U}}^{(i)} = \hat{\mathbf{Q}}^{(i)}\hat{\mathbf{U}}^{(i)}\hat{\mathbf{\Lambda}}^{(i)} \quad (9)$$

where $\hat{\mathbf{U}}^{(i)}$ is the estimate of \mathbf{U} in the i th alternation. According to (9), $\hat{\mathbf{U}}^{(i)}$ is the collection of $M-L$ eigenvectors corresponding to the $M-L$ smallest eigenvalues obtained by carrying out the GED of the pair of matrices $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}^{(i)}\}$, and $\hat{\mathbf{\Lambda}}^{(i)} \in \mathbb{R}^{(M-L) \times (M-L)}$ is a diagonal matrix which consists of the associated eigenvalues. As multiplying $\hat{\mathbf{U}}^{(i)}$ by a positive scalar does not affect (9), we normalize $\hat{\mathbf{U}}^{(i)}$ to ensure that $\|\hat{\mathbf{U}}^{(i)}\|_F = 1$ in each iteration. This boosts the convergence of the proposed iterative nonuniform noise covariance estimator (see Algorithm 1). Moreover, the introduction of $\hat{\mathbf{\Lambda}}^{(i)}$ is necessary to revise (7) in the form of (8) to come up to the exact equality (9) in the i th alternation. Noteworthy to mention that (9) is inspired by (25), which follows the GED formulation with the identity matrix as the diagonal eigenvalue matrix.

On the other hand, the LS approach based on (7) can be employed to estimate \mathbf{Q} . In doing so, the sub-problem which corresponds to the update \mathbf{Q} in the $(i+1)$ st alternation is given as

$$\hat{\mathbf{Q}}^{(i+1)} = \arg \min_{\mathbf{Q}} \|(\hat{\mathbf{R}} - \mathbf{Q})\hat{\mathbf{U}}^{(i)}\|_F^2. \quad (10)$$

Problem (10) needs to be solved subject to the constraint on \mathbf{Q} to be a diagonal matrix. The cost function of (10)

can be expressed as

$$\begin{aligned}
f(\mathbf{Q}) &\triangleq \left\| (\hat{\mathbf{R}} - \mathbf{Q}) \hat{\mathbf{U}}^{(i)} \right\|_{\mathbb{F}}^2 \\
&= \text{trace} \left\{ \left((\hat{\mathbf{R}} - \mathbf{Q}) \hat{\mathbf{U}}^{(i)} \right) \left((\hat{\mathbf{R}} - \mathbf{Q}) \hat{\mathbf{U}}^{(i)} \right)^H \right\} \\
&= \text{trace} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}}^2 \right\} - \text{trace} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \mathbf{Q} \right\} \\
&\quad - \text{trace} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}} \mathbf{Q} \right\} + \text{trace} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \mathbf{Q}^2 \right\}
\end{aligned} \tag{11}$$

where the properties $\|\mathbf{G}\|_{\mathbb{F}}^2 = \text{trace} \{ \mathbf{G} \mathbf{G}^H \}$, $\text{trace} \{ \mathbf{G} \mathbf{Y} \} = \text{trace} \{ \mathbf{Y} \mathbf{G} \}$, $\hat{\mathbf{R}} = \hat{\mathbf{R}}^H$, and $\mathbf{Q} = \mathbf{Q}^H$ are used. As elaborated in Appendix A, the partial derivative of (11) with respect to \mathbf{Q} is given as

$$\begin{aligned}
\frac{\partial f(\mathbf{Q})}{\partial \mathbf{Q}} &= 2\mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \right\} \mathbf{Q} \\
&\quad - \mathcal{D} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}} \right\}.
\end{aligned} \tag{12}$$

Therefore, by equating (12) to zero and solving the resulting equation for \mathbf{Q} , we find the optimal estimate as

$$\begin{aligned}
\hat{\sigma}_m^{2(i+1)} &= \mathbf{d}_m^T \left(\frac{1}{2} \mathcal{D} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}} \right\} \right. \\
&\quad \left. \times \mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \right\}^{-1} \right) \mathbf{d}_m, \quad m = 1, \dots, M
\end{aligned} \tag{13}$$

where $\mathbf{d}_m \in \mathbb{R}^M$ denotes a vector with one on the m th position and zero elsewhere. The steps of the proposed method for the nonuniform noise covariance matrix estimation are outlined in Algorithm 1, where the iteration loop is terminated when the criteria $|f^{(i+1)} - f^{(i)}| < \epsilon$ is satisfied. Note that in the i th iteration, only the m th diagonal entry of \mathbf{Q} is updated where $m = \text{rem}(i, M) + 1$ with $\text{rem}(a, b)$ being the least non-negative remainder in the division of a by b , while the whole matrix of \mathbf{U} is updated in each iteration. Updating \mathbf{Q} using the aforementioned element-wise manner accelerates the convergence of the algorithm.

Algorithm 1: Nonuniform Noise Covariance Matrix Estimation

1: Compute $\hat{\mathbf{R}} = 1/N \sum_{t=1}^N \mathbf{x}(t) \mathbf{x}^H(t)$. Set $i = 0$ and $\hat{\mathbf{Q}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$.

repeat

2: Carry out the GED of the pair of matrices $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}^{(i)}\}$ to obtain $\hat{\mathbf{U}}^{(i)}$ as the $M - L$ eigenvectors corresponding to the $M - L$ smallest eigenvalues. Then, normalize $\hat{\mathbf{U}}^{(i)}$ as $\hat{\mathbf{U}}^{(i)} = \hat{\mathbf{U}}^{(i)} / \|\hat{\mathbf{U}}^{(i)}\|_{\mathbb{F}}$.

3: Compute $f^{(i)} = \|\hat{\mathbf{R}} - \hat{\mathbf{Q}}^{(i)}\|_{\mathbb{F}}^2$.

4: Set $m = \text{rem}(i, M) + 1$.

5: Construct $\hat{\mathbf{Q}}^{(i+1)} = \text{diag}\{\hat{\sigma}_1^{2(i+1)}, \dots, \hat{\sigma}_m^{2(i+1)}, \dots, \hat{\sigma}_M^{2(i+1)}\}$

where $\hat{\sigma}_m^{2(i+1)}$ is calculated using (13).¹

6: Compute $f^{(i+1)} = \|\hat{\mathbf{R}} - \hat{\mathbf{Q}}^{(i+1)}\|_{\mathbb{F}}^2$.

7: Set $i = i + 1$.

until $|f^{(i+1)} - f^{(i)}| < \epsilon$

¹Note that $\hat{\sigma}_j^{2(i)}$ for $j \neq m$ denote entries of $\hat{\mathbf{Q}}$ from the previous alternation that remain unchanged in the current one.

Remark 1: The estimate of $\hat{\mathbf{Q}}^{(i+1)}$ in (13) is an alternative representation of the power domain (PD) method [32], that is,

$$\hat{\sigma}_m^2 = \frac{(\mathbf{d}_m^T \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{r}}_m + \hat{\mathbf{r}}_m^H \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{d}_m)}{2 \mathbf{d}_m^T \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{d}_m}, \quad m = 1, \dots, M \tag{14}$$

where $\hat{\mathbf{r}}_m$ is the m th column of $\hat{\mathbf{R}}$, and $\mathbf{P}_{\mathbf{A}}^{\perp} \in \mathbb{C}^{M \times M}$ is the orthogonal projection matrix of the signal subspace, i.e., $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I}_M - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$. The main difference between (13) and (14) is the use of $\hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H$ as a proper estimate of $\mathbf{P}_{\mathbf{A}}^{\perp}$ instead of conducting a multidimensional search for finding an ML estimate as in [32]. It is well-known that performing such multidimensional search is a very computationally demanding task. Therefore, the proposed method is advantageous compared to PD technique in the sense that it requires significantly less computations.

Remark 2: The relationship between $\hat{\sigma}_m^{2(i+1)}$ and $\hat{\mathbf{Q}}^{(i)}$ can be obtained via exploiting (9) and (13) as follows

$$\begin{aligned}
\hat{\sigma}_m^{2(i+1)} &= \mathbf{d}_m^T \left(\frac{1}{2} \mathcal{D} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}} \right\} \right. \\
&\quad \left. \times \mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \right\}^{-1} \right) \mathbf{d}_m
\end{aligned} \tag{15}$$

$$= \mathbf{d}_m^T \left(\mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} \hat{\mathbf{\Lambda}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \right\} \mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} (\hat{\mathbf{U}}^{(i)})^H \right\}^{-1} \hat{\mathbf{Q}}^{(i)} \right) \mathbf{d}_m$$

where the properties $\hat{\mathbf{Q}}^{(i)} = (\hat{\mathbf{Q}}^{(i)})^H$ and $\hat{\mathbf{\Lambda}}^{(i)} = (\hat{\mathbf{\Lambda}}^{(i)})^H$ have been used. Since $\hat{\mathbf{R}}$ and $\hat{\mathbf{Q}}^{(i)}$ are positive semi-definite Hermitian and positive definite Hermitian matrices, respectively, the eigenvalues obtained by carrying out the GED of the pair of matrices $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}^{(i)}\}$ are non-negative real numbers [49]. Taking into account this fact, it is straightforward to show that $\hat{\sigma}_m^{2(i+1)}$ is non-negative real number when the diagonal entries of $\hat{\mathbf{Q}}^{(i)}$ are non-negative real numbers. Therefore, initializing the noise covariance matrix as $\hat{\mathbf{Q}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$ guarantees the non-negativity of the diagonal entries of $\hat{\mathbf{Q}}^{(i)}$ for $i \geq 1$.

Remark 3: In terms of convergence, it is worth noting that the update of any given matrix ($\hat{\mathbf{Q}}$ or $\hat{\mathbf{U}}$) may either improve or maintain but cannot worsen the current estimate as the solutions to the corresponding sub-problems are found optimally. Monotone convergence (to a stationary point, but not necessarily to the global minimum) of the proposed noise covariance estimator follows directly from this fact.

B. Block-diagonal Sensor Noise

We aim now to develop an alternative noise covariance matrix estimator for the case where the sensor noise covariance matrix has the block-diagonal structure. We begin with expressing the block-diagonal noise covariance

matrix as

$$\mathbf{Q}_{\text{bdiag}} = \begin{bmatrix} \mathbf{Q}_1 & & 0 \\ & \ddots & \\ 0 & & \mathbf{Q}_q \end{bmatrix} = \text{bdiag}\{\mathbf{Q}_1, \dots, \mathbf{Q}_q\} \quad (16)$$

where $\mathbf{Q}_j \in \mathbb{C}^{n_j \times n_j}$. Inspired by the idea presented in the previous subsection, we rewrite (7) for the block-diagonal noise covariance matrix as

$$\hat{\mathbf{R}}\mathbf{U} \approx \mathbf{Q}_{\text{bdiag}}\mathbf{U} \quad (17)$$

Similar to the nonuniform noise case, (17) suggests to use an alternating approach to estimate $\mathbf{Q}_{\text{bdiag}}$ and \mathbf{U} . Given an estimate of $\mathbf{Q}_{\text{bdiag}}$ in the i th alternation, which is denoted by $\hat{\mathbf{Q}}_{\text{bdiag}}^{(i)}$, we convert (17) into an equality relation that has the form of (8) as

$$\hat{\mathbf{R}}\hat{\mathbf{U}}^{(i)} = \hat{\mathbf{Q}}_{\text{bdiag}}^{(i)}\hat{\mathbf{U}}^{(i)}\hat{\mathbf{\Lambda}}^{(i)} \quad (18)$$

where $\hat{\mathbf{U}}^{(i)}$ is the estimate of \mathbf{U} in the i th alternation. According to (18), $\hat{\mathbf{U}}^{(i)}$ is the collection of $M - L$ eigenvectors corresponding to the $M - L$ smallest eigenvalues obtained by carrying out the GED of the pair of matrices $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}_{\text{bdiag}}^{(i)}\}$, and $\hat{\mathbf{\Lambda}}^{(i)} \in \mathbb{R}^{(M-L) \times (M-L)}$ is a diagonal matrix which consists of the associated eigenvalues. Analogous to the nonuniform noise case, we normalize $\hat{\mathbf{U}}^{(i)}$ in each iteration in order to boost the convergence of the proposed iterative block-diagonal noise covariance estimator (see Algorithm 2).

According to (17), $\mathbf{Q}_{\text{bdiag}}$ can be updated in the $(i + 1)$ st alternation using the LS approach by solving the problem

$$\hat{\mathbf{Q}}_{\text{bdiag}}^{(i+1)} = \arg \min_{\mathbf{Q}_{\text{bdiag}}} \|(\hat{\mathbf{R}} - \mathbf{Q}_{\text{bdiag}})\hat{\mathbf{U}}^{(i)}\|_{\text{F}}^2. \quad (19)$$

The objective function of (19) can be expanded as

$$\begin{aligned} f_{\text{bdiag}}(\mathbf{Q}_{\text{bdiag}}) &\triangleq \|(\hat{\mathbf{R}} - \mathbf{Q}_{\text{bdiag}})\hat{\mathbf{U}}^{(i)}\|_{\text{F}}^2 \\ &= \text{trace} \left\{ \left((\hat{\mathbf{R}} - \mathbf{Q}_{\text{bdiag}})\hat{\mathbf{U}}^{(i)} \right) \left((\hat{\mathbf{R}} - \mathbf{Q}_{\text{bdiag}})\hat{\mathbf{U}}^{(i)} \right)^H \right\} \\ &= \text{trace} \left\{ \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}}^2 \right\} \\ &\quad - \text{trace} \left\{ \hat{\mathbf{R}}\hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H \mathbf{Q}_{\text{bdiag}} \right\} \\ &\quad - \text{trace} \left\{ \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H \hat{\mathbf{R}}\mathbf{Q}_{\text{bdiag}} \right\} \\ &\quad + \text{trace} \left\{ \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H \mathbf{Q}_{\text{bdiag}}^2 \right\} \end{aligned} \quad (20)$$

Defining $\bar{\mathbf{R}}^{(i)} \triangleq \hat{\mathbf{R}}\hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H + \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H\hat{\mathbf{R}}$ and $\bar{\mathbf{V}}^{(i)} \triangleq \hat{\mathbf{U}}^{(i)}(\hat{\mathbf{U}}^{(i)})^H$, we can recast (20) as

$$f_{\text{bdiag}}(\mathbf{Q}_{\text{bdiag}}) = \sum_{j=1}^q \text{trace} \left\{ \bar{\mathbf{V}}_{jj}^{(i)} \mathbf{Q}_j^2 - \bar{\mathbf{R}}_{jj}^{(i)} \mathbf{Q}_j \right\} \quad (21)$$

where $\bar{\mathbf{V}}_{jj}^{(i)}$ and $\bar{\mathbf{R}}_{jj}^{(i)}$ are respectively the j th block on the main diagonal of $\bar{\mathbf{V}}^{(i)}$ and $\bar{\mathbf{R}}^{(i)}$ with the same size as \mathbf{Q}_j , and also the constant term is eliminated. It can be seen from (21) that estimation of \mathbf{Q}_j in the $(i + 1)$ st alternation

only depends on $\bar{\mathbf{V}}_{jj}^{(i)}$ and $\bar{\mathbf{R}}_{jj}^{(i)}$, enabling us to obtain an estimate of $\mathbf{Q}_j, j = 1, \dots, q$ in the $(i + 1)$ st alternation separately. Using the complex-valued derivative properties [50], the partial derivative of (21) with respect to the Hermitian matrix \mathbf{Q}_j is given as

$$\frac{\partial f_{\text{bdiag}}(\mathbf{Q}_{\text{bdiag}})}{\partial \mathbf{Q}_j} = \mathbf{Q}_j^* (\bar{\mathbf{V}}_{jj}^{(i)})^T + (\bar{\mathbf{V}}_{jj}^{(i)})^T \mathbf{Q}_j^T - (\bar{\mathbf{R}}_{jj}^{(i)})^T, \quad j = 1, \dots, q \quad (22)$$

Applying the transposition operator to (22), and then equating the result to zero, we have

$$\bar{\mathbf{V}}_{jj}^{(i)} \mathbf{Q}_j + \mathbf{Q}_j \bar{\mathbf{V}}_{jj}^{(i)} = \bar{\mathbf{R}}_{jj}^{(i)}, \quad j = 1, \dots, q \quad (23)$$

where the Hermitian property of \mathbf{Q}_j 's is also used. The estimates of \mathbf{Q}_j for $j = 1, \dots, q$ obtained by solving (23) are Hermitian matrices but not necessarily are positive semi-definite matrices. Thus, obtaining Hermitian positive semi-definite estimates of \mathbf{Q}_j for $j = 1, \dots, q$ requires modifying (23) slightly. Since the following discussion is true for any $j \in \{1, \dots, q\}$, we only present it for one value of $j = d$ here, and then extend it to all values of $j \in \{1, \dots, q\}$. Let $\mathbf{y} \in \mathbb{C}^{n_d \times 1}$ and $\mu \in \mathbb{R}$ be an eigenvector and the corresponding eigenvalue of the matrix \mathbf{Q}_d , respectively. Therefore, we have

$$\mathbf{Q}_d \mathbf{y} = \mu \mathbf{y}. \quad (24)$$

Left-multiplying and right-multiplying the equation in (23) which corresponds to $j = d$ by \mathbf{y}^H and \mathbf{y} , respectively, and then using (24), we obtain

$$\mu = \frac{\mathbf{y}^H \bar{\mathbf{R}}_{ss}^{(i)} \mathbf{y}}{2\mathbf{y}^H \bar{\mathbf{V}}_{ss}^{(i)} \mathbf{y}}. \quad (25)$$

Given the assumption that $n_s < M - L$ (see Remark 4), the denominator of (25) is positive. Therefore, non-negativity of μ is guaranteed by enforcing $\bar{\mathbf{R}}_{ss}^{(i)}$ to be positive semi-definite. As a result, for having positive semi-definite estimates of \mathbf{Q}_j 's, (23) should be modified as

$$\bar{\mathbf{V}}_{jj}^{(i)} \mathbf{Q}_j + \mathbf{Q}_j \bar{\mathbf{V}}_{jj}^{(i)} = \left[\bar{\mathbf{R}}_{jj}^{(i)} \right]_+, \quad j = 1, \dots, q \quad (26)$$

where $[\cdot]_+$ is an operator which projects the bracketed matrix onto the positive semi-definite cone via carrying out the eigenvalue decomposition and setting all the negative eigenvalues to zero. Applying the vectorization operator to (26), we obtain

$$\mathbf{V}_j^{(i)} \mathbf{q}_j = \bar{\mathbf{r}}_j^{(i)}, \quad j = 1, \dots, q \quad (27)$$

where $\mathbf{V}_j^{(i)} \triangleq \left[\left((\bar{\mathbf{V}}_{jj}^{(i)})^T \otimes \mathbf{I}_{n_j} \right) + \left(\mathbf{I}_{n_j} \otimes \bar{\mathbf{V}}_{jj}^{(i)} \right) \right]$, $\mathbf{q}_j \triangleq \text{vec}\{\mathbf{Q}_j\}$, and $\bar{\mathbf{r}}_j^{(i)} \triangleq \text{vec}\left\{ \left[\bar{\mathbf{R}}_{jj}^{(i)} \right]_+ \right\}$. According to (27), we have

$$\hat{\mathbf{Q}}_j^{(i+1)} = \text{unvec} \left\{ \left(\mathbf{V}_j^{(i)} \right)^\dagger \bar{\mathbf{r}}_j^{(i)} \right\}, \quad j = 1, \dots, q. \quad (28)$$

Analogous to the proposed nonuniform noise covariance matrix estimator, the proposed block-diagonal noise covariance matrix estimator can be initialized by any proper matrix $\hat{\mathbf{Q}}_{\text{bdiag}}^{(0)}$. We stick to $\hat{\mathbf{Q}}_{\text{bdiag}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$. The

steps of the proposed method for the block-diagonal noise covariance matrix estimation are outlined in Algorithm 2. To accelerate the convergence of Algorithm 2, only the j th block of $\mathbf{Q}_{\text{bdiag}}$ is updated in the i th iteration where $j = \text{rem}(i, q) + 1$, while the whole matrix of \mathbf{U} is updated in each iteration.

Algorithm 2: Block-diagonal Noise Covariance Matrix Estimation

- 1: Compute $\hat{\mathbf{R}} = 1/N \sum_{t=1}^N \mathbf{x}(t)\mathbf{x}^H(t)$.
- 2: Set $i = 0$, and $\mathbf{Q}_{\text{bdiag}}^{(0)} = \mathcal{D}\{\hat{\mathbf{R}}\}$.
- repeat**
- 3: Carry out the GED of the pair of matrices $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}_{\text{bdiag}}^{(i)}\}$ to obtain $\hat{\mathbf{U}}^{(i)}$ as the collection of $M - L$ eigenvectors corresponding to the $M - L$ smallest eigenvalues. Then, normalize $\hat{\mathbf{U}}^{(i)}$ as $\hat{\mathbf{U}}^{(i)} = \hat{\mathbf{U}}^{(i)} / \|\hat{\mathbf{U}}^{(i)}\|_{\text{F}}$.
- 4: Compute $f_{\text{bdiag}}^{(i)} = \|(\hat{\mathbf{R}} - \hat{\mathbf{Q}}_{\text{bdiag}}^{(i)})\hat{\mathbf{U}}^{(i)}\|_{\text{F}}^2$.
- 5: Construct $\mathbf{V}_j^{(i)}$ and $\bar{\mathbf{r}}_j^{(i)}$ for $j = \text{rem}(i, q) + 1$.
- 6: Calculate $\hat{\mathbf{Q}}_j^{(i+1)}$ using (28).
- 7: Construct $\hat{\mathbf{Q}}_{\text{bdiag}}^{(i+1)} = \text{bdiag}\{\hat{\mathbf{Q}}_1^{(i)}, \dots, \hat{\mathbf{Q}}_j^{(i+1)}, \dots, \hat{\mathbf{Q}}_q^{(i)}\}$.²
- 8: Compute $f_{\text{bdiag}}^{(i+1)} = \|(\hat{\mathbf{R}} - \hat{\mathbf{Q}}_{\text{bdiag}}^{(i+1)})\hat{\mathbf{U}}^{(i)}\|_{\text{F}}^2$.
- 9: Set $i = i + 1$.
- until** $|f_{\text{bdiag}}^{(i+1)} - f_{\text{bdiag}}^{(i)}| < \epsilon$

Remark 4: Noteworthy to mention that there is a limitation to the use of Algorithm 2, that is, the rank of each sub-matrix \mathbf{Q}_j should not exceed $M - L$.

IV. SUBSPACE-BASED DOA ESTIMATION VIA GLS

A. Forward-only Algorithm for DOA Estimation

The received signal can be preprocessed by multiplying (4) by $\mathbf{Q}^{-\frac{1}{2}}$ to enforce the uniform noise. Thus, the received signal becomes

$$\bar{\mathbf{X}} \triangleq \mathbf{Q}^{-\frac{1}{2}} \mathbf{X} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A} \mathbf{S} + \mathbf{Q}^{-\frac{1}{2}} \mathbf{N} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A} \mathbf{S} + \bar{\mathbf{N}} \quad (29)$$

where the columns of $\bar{\mathbf{N}}$ are Gaussian random vectors with zero mean and the covariance matrix is \mathbf{I}_M .

The truncated SVD of $\bar{\mathbf{X}}$ is given as

$$\bar{\mathbf{X}} = \mathbf{U}_s \Sigma_s \mathbf{V}_s^H \quad (30)$$

where $\mathbf{U}_s \in \mathbb{C}^{M \times L}$ and $\mathbf{V}_s \in \mathbb{C}^{N \times L}$ denote respectively the left and right singular vectors associated with the L principal singular values on the diagonal of $\Sigma_s \in \mathbb{R}^{L \times L}$.

According to (29) and (30), the columns of \mathbf{U}_s , denoted as \mathbf{u}_p for $p = 1, \dots, L$, and the columns of $\mathbf{Q}^{-\frac{1}{2}} \mathbf{A}$ span the same column space, i.e., $\text{span}(\mathbf{U}_s) = \text{span}(\mathbf{Q}^{-\frac{1}{2}} \mathbf{A})$. In other words, \mathbf{U}_s and $\mathbf{Q}^{-\frac{1}{2}} \mathbf{A}$ are related as

$$\mathbf{U}_s = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A} \mathbf{G} \quad (31)$$

where $\mathbf{G} \in \mathbb{C}^{L \times L}$ is a non-singular matrix. Multiplying (31) from the left by $\mathbf{Q}^{\frac{1}{2}}$, we obtain

$$\tilde{\mathbf{U}}_s = \mathbf{A} \mathbf{G} \quad (32)$$

²Note that $\hat{\mathbf{Q}}_m^{(i)}$ for $m \neq j$ denote blocks of $\hat{\mathbf{Q}}_{\text{bdiag}}$ from the previous alternation that remain unchanged in the current one.

where $\tilde{\mathbf{U}}_s \triangleq \mathbf{Q}^{\frac{1}{2}} \mathbf{U}_s \in \mathbb{C}^{M \times L}$. We denote the p th column of $\tilde{\mathbf{U}}_s$ as $\tilde{\mathbf{u}}_p$.

Using (32), the m th entry of the p th column of $\tilde{\mathbf{U}}_s$, denoted as $[\tilde{\mathbf{u}}_p]_m$, can be written as

$$[\tilde{\mathbf{u}}_p]_m = [\mathbf{A}]_{m,:} \mathbf{g}_p = \sum_{l=1}^L [\mathbf{g}_p]_l e^{-j2\pi d \sin(\theta_l)(m-1)/\lambda}, \quad p = 1, \dots, L, \quad m = 1, \dots, M \quad (33)$$

where $\mathbf{g}_p \in \mathbb{C}^L$ is the p th column of \mathbf{G} .

Next, the DFT can be applied to each column of $\tilde{\mathbf{U}}_s$ as, for example, in [51], [52]³. Using (33) and the definition of DFT, the k th bin of the DFT of $\tilde{\mathbf{u}}_p$ can be expressed as

$$[\tilde{\mathbf{u}}_p]_k = \sum_{l=1}^L [\mathbf{g}_p]_l \frac{1 - e^{jM\beta_l}}{1 - e^{-j\frac{2\pi k}{M} e^{j\beta_l}}} = \sum_{l=1}^L \frac{\alpha_l}{1 - \gamma_l W_M^k}, \quad k = 1, \dots, M \quad (34)$$

where $\tilde{\mathbf{u}}_p \triangleq \text{DFT}\{\tilde{\mathbf{u}}_p\} = \mathbf{W}_D \tilde{\mathbf{u}}_p$, $p = 1, \dots, L$ and

$$\mathbf{W}_D = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{-j\frac{2\pi}{M}} & \dots & e^{-j\frac{2\pi(M-1)}{M}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\frac{2\pi(M-1)}{M}} & \dots & e^{-j\frac{2\pi(M-1)(M-1)}{M}} \end{bmatrix}. \quad (35)$$

In addition, $\alpha_l \triangleq [\mathbf{g}_p]_l (1 - e^{jM\beta_l})$, $\gamma_l \triangleq e^{j\beta_l}$, $\beta_l \triangleq -2\pi d \sin(\theta_l)/\lambda$, and $W_M^k \triangleq e^{-j\frac{2\pi k}{M}}$ in (34). Unifying the L rational functions into one, we recast (34) as

$$[\tilde{\mathbf{u}}_p]_k = \frac{\sum_{l=1}^L \alpha_l \prod_{\substack{v=1 \\ v \neq l}}^L (1 - \gamma_v W_M^k)}{\prod_{l=1}^L (1 - \gamma_l W_M^k)} \quad (36)$$

where the common denominator is the product of the L denominators of each rational function. Noticing the special structure of (36), the nominator and denominator can be expanded as two polynomials of degrees $L - 1$ and L , respectively, that is,

$$\sum_{l=1}^L \alpha_l \prod_{\substack{v=1 \\ v \neq l}}^L (1 - \gamma_v W_M^k) = \sum_{l=1}^L b_{pl} (W_M^k)^{l-1} = \bar{\mathbf{w}}_k^T \mathbf{b}_p \quad (37)$$

$$\prod_{l=1}^L (1 - \gamma_l W_M^k) = 1 + \sum_{l=1}^L a_l (W_M^k)^l = 1 + \mathbf{w}_k^T \mathbf{a} \quad (38)$$

where $\bar{\mathbf{w}}_k \triangleq [1 \ W_M^k \ (W_M^k)^2 \ \dots \ (W_M^k)^{L-1}]^T$, $\mathbf{w}_k \triangleq [W_M^k \ (W_M^k)^2 \ (W_M^k)^3 \ \dots \ (W_M^k)^L]^T$, $\mathbf{b}_p \triangleq [b_{p1} \ \dots \ b_{pL}]^T$, and $\mathbf{a} \triangleq [a_1 \ \dots \ a_L]^T$.

³It is worth mentioning that the methodology utilized in this work is different from [52].

It can be seen from (38) that estimating \mathbf{a} is the key for finding γ_l 's, $l = 1, \dots, L$, since γ_l 's are the roots of the polynomial defined by the entries of \mathbf{a} as

$$\gamma^L + \sum_{l=1}^L [\mathbf{a}]_l \gamma^{L-l} = 0. \quad (39)$$

Finally, knowing γ_l 's, θ_l 's can be extracted using the relation $\theta_l = \arcsin(-\frac{\beta_l \lambda}{2\pi d})$ where β_l is the phase argument of γ_l as it is defined above.

The objective now is to find an estimate of \mathbf{a} . Multiplying both sides of (36) by the denominator, we have

$$[\bar{\mathbf{u}}_p]_k (1 + \mathbf{w}_k^T \mathbf{a}) = \bar{\mathbf{w}}_k^T \mathbf{b}_p. \quad (40)$$

Piling up all the equations that can be generated for $k = 1, \dots, M$ based on (40), we can write

$$\bar{\mathbf{u}}_p + \text{diag}(\bar{\mathbf{u}}_p) \mathbf{W}_a \mathbf{a} = \bar{\mathbf{W}} \mathbf{b}_p, \quad p = 1, \dots, L \quad (41)$$

where $\mathbf{W}_a \triangleq [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_M]^T \in \mathbb{C}^{M \times L}$ and $\bar{\mathbf{W}} \triangleq [\bar{\mathbf{w}}_1 \bar{\mathbf{w}}_2 \dots \bar{\mathbf{w}}_M]^T \in \mathbb{C}^{M \times L}$.

We suggest to generate two sets of estimates for θ_l 's as the most probable candidates first, and then pick up the best candidates via using a proper selection criteria. In doing so, we introduce the selection matrix $\mathbf{Z}_{\mathcal{I}} \in \mathbb{R}^{|\mathcal{I}| \times M}$. Here \mathcal{I} denotes the set containing the indices of the selected equations. The matrix $\mathbf{Z}_{\mathcal{I}}$ is used to consider different sets of (41) for estimating \mathbf{a} . Since $\mathbf{Z}_{\mathcal{I}}$ is a selection matrix, all the entries of the i th row of $\mathbf{Z}_{\mathcal{I}}$ are zeros except one entry whose index is the i th member of \mathcal{I} . The only nonzero entry of each row is set to 1. The proposed method for selecting two sets of indices to serve as \mathcal{I} will be clarified in the sequel.

Given a particular set \mathcal{I} , the selected subset of (41) can be written as

$$\mathbf{Z}_{\mathcal{I}} \bar{\mathbf{u}}_p + \text{diag}(\mathbf{Z}_{\mathcal{I}} \bar{\mathbf{u}}_p) \mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \mathbf{a} = \mathbf{Z}_{\mathcal{I}} \bar{\mathbf{W}} \mathbf{b}_p, \quad p = 1, \dots, L. \quad (42)$$

To avoid estimating \mathbf{b}_p 's in (42), $\mathbf{B} \in \mathbb{C}^{|\mathcal{I}| \times (|\mathcal{I}| - L)}$ can be obtained using SVD such that $\mathbf{B}^H \tilde{\mathbf{Z}}_{\mathcal{I}} = \mathbf{0}_{(|\mathcal{I}| - L) \times L}$ with $\tilde{\mathbf{Z}}_{\mathcal{I}} \triangleq \mathbf{Z}_{\mathcal{I}} \bar{\mathbf{W}} \in \mathbb{C}^{|\mathcal{I}| \times L}$. Thus, multiplying both sides of (42) by \mathbf{B}^H and using the property $\mathbf{B}^H \tilde{\mathbf{Z}}_{\mathcal{I}} = \mathbf{0}_{(|\mathcal{I}| - L) \times L}$, yields

$$\mathbf{B}^H (\mathbf{Z}_{\mathcal{I}} \bar{\mathbf{u}}_p + \text{diag}(\mathbf{Z}_{\mathcal{I}} \bar{\mathbf{u}}_p) \mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \mathbf{a}) = \mathbf{0}_{(|\mathcal{I}| - L)}, \quad p = 1, \dots, L. \quad (43)$$

The impact of \mathbf{b}_p 's is eliminated in (43) independent of the value of p because of the definition of $\tilde{\mathbf{Z}}_{\mathcal{I}}$, which is independent of p . This property enables us to combine the L sets of linear equations generated via (43). Rearranging the terms in (43), we get the system of linear equations $\mathbf{H}_p \mathbf{a} = \mathbf{h}_p$, $p = 1, \dots, L$, where $\mathbf{H}_p \triangleq \mathbf{B}^H \text{diag}(\mathbf{Z}_{\mathcal{I}} \bar{\mathbf{u}}_p) \mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \in \mathbb{C}^{(|\mathcal{I}| - L) \times L}$ and $\mathbf{h}_p \triangleq -\mathbf{B}^H \mathbf{Z}_{\mathcal{I}} \bar{\mathbf{u}}_p \in \mathbb{C}^{(|\mathcal{I}| - L)}$. Stacking the L matrices \mathbf{H}_p and the L vectors \mathbf{h}_p into a larger matrix \mathbf{H} and a longer vector \mathbf{h} , respectively, we have

$$\mathbf{H} \mathbf{a} = \mathbf{h} \quad (44)$$

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

However, only an estimate of \mathbf{Q} , denoted by $\hat{\mathbf{Q}}$, can be obtained in practice.⁴ Replacing \mathbf{Q} by $\hat{\mathbf{Q}}$, (29) is modified as

$$\hat{\tilde{\mathbf{X}}} \triangleq \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{X} = \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A} \mathbf{S} + \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{N} = \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A} \mathbf{S} + \hat{\tilde{\mathbf{N}}}. \quad (45)$$

Thus, only an estimate of \mathbf{U}_s (denoted by $\hat{\mathbf{U}}_s$), which spans the column space of $\hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A}$, can be obtained via the truncated SVD of $\hat{\tilde{\mathbf{X}}}$ due to the presence of errors. Hence, the truncated SVD of $\hat{\tilde{\mathbf{X}}}$ is given as

$$\hat{\tilde{\mathbf{X}}} = \hat{\mathbf{U}}_s \hat{\Sigma}_s \hat{\mathbf{V}}_s^H \quad (46)$$

where $\hat{\mathbf{U}}_s = [\hat{\mathbf{u}}_1 \dots \hat{\mathbf{u}}_L] \in \mathbb{C}^{M \times L}$ is the matrix of L left singular vectors associated with the L largest singular values on the diagonal of $\hat{\Sigma}_s \in \mathbb{R}^{L \times L}$. Accordingly, $\hat{\mathbf{U}}_s$ is defined as

$$\hat{\mathbf{U}}_s \triangleq \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{U}}_s = [(\hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{u}}_1) \dots (\hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{u}}_L)] \in \mathbb{C}^{M \times L}. \quad (47)$$

Replacing $\tilde{\mathbf{u}}_p$ by $\hat{\tilde{\mathbf{u}}}_p$ in (44) and generating $\hat{\mathbf{H}}$ and $\hat{\mathbf{h}}$, (44) turns into an approximate equality, i.e.,

$$\hat{\mathbf{H}} \mathbf{a} \approx \hat{\mathbf{h}}. \quad (48)$$

The optimal value of \mathbf{a} can now be obtained employing the GLS technique [53], [54]. Using (48), the GLS optimization problem can be formulated as

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} (\hat{\mathbf{H}} \mathbf{a} - \hat{\mathbf{h}})^H \mathbf{W} (\hat{\mathbf{H}} \mathbf{a} - \hat{\mathbf{h}}) \quad (49)$$

where $\mathbf{W} \triangleq (\mathbb{E}\{\hat{\mathbf{e}} \hat{\mathbf{e}}^H\})^{-1} \in \mathbb{C}^{L(|\mathcal{I}| - L) \times L(|\mathcal{I}| - L)}$ and $\hat{\mathbf{e}} \triangleq \hat{\mathbf{H}} \mathbf{a} - \hat{\mathbf{h}} \in \mathbb{C}^{L(|\mathcal{I}| - L)}$. The solution of (49) is given by

$$\hat{\mathbf{a}} = (\hat{\mathbf{H}}^H \mathbf{W} \hat{\mathbf{H}})^{-1} \hat{\mathbf{H}}^H \mathbf{W} \hat{\mathbf{h}}. \quad (50)$$

To use (50), an estimate of \mathbf{W} is required. However, it is clear from the definitions of \mathbf{W} and $\hat{\mathbf{e}}$ that an estimate of \mathbf{W} depends on the unknown vector \mathbf{a} . Thus, it is natural to utilize an iterative scheme to estimate $\hat{\mathbf{a}}$ in one step, followed by estimating $\hat{\mathbf{W}}$ in the other step by employing $\hat{\mathbf{a}}$ obtained in the previous step. The alternation between these two steps is then carried on until a termination criteria is satisfied.

To figure out a principle for finding $\hat{\mathbf{W}}$, we take into consideration the first-order subspace estimation error by expressing $\hat{\mathbf{u}}_p$ as $\hat{\mathbf{u}}_p \triangleq \mathbf{u}_p + \Delta \mathbf{u}_p$ for $p = 1, \dots, L$. Exploiting this definition, we have

$$\begin{aligned} \hat{\mathbf{e}}_p &\triangleq \hat{\mathbf{H}}_p \mathbf{a} - \hat{\mathbf{h}}_p = \mathbf{B}^H (\mathbf{Z}_{\mathcal{I}} \hat{\mathbf{u}}_p + \text{diag}\{\mathbf{Z}_{\mathcal{I}} \hat{\mathbf{u}}_p\} \mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \mathbf{a}) \\ &= \mathbf{B}^H \left(\mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{u}}_p + \text{diag}\{\mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{u}}_p\} \mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \mathbf{a} \right) \\ &= \mathbf{B}^H \left(\mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{u}}_p + \text{diag}\{\mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \mathbf{a}\} \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{u}}_p \right) \\ &= \mathbf{B}^H (\mathbf{I}_{|\mathcal{I}|} + \text{diag}\{\mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \mathbf{a}\}) \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{u}}_p \\ &= \mathbf{C}(\mathbf{a}) \hat{\mathbf{u}}_p, \quad p = 1, \dots, L \end{aligned} \quad (51)$$

where $\mathbf{C}(\mathbf{a}) \triangleq \mathbf{B}^H (\mathbf{I}_{|\mathcal{I}|} + \text{diag}\{\mathbf{Z}_{\mathcal{I}} \mathbf{W}_a \mathbf{a}\}) \mathbf{Z}_{\mathcal{I}} \mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \in \mathbb{C}^{L(|\mathcal{I}| - L) \times M}$. In (51), we also used the property

⁴Algorithm 1 and 2 can be utilized to find $\hat{\mathbf{Q}}$ in the cases on nonuniform and block-diagonal noise, respectively.

$\text{diag}\{\mathbf{x}\}\mathbf{y} = \text{diag}\{\mathbf{y}\}\mathbf{x}$ as well as the equality $\hat{\mathbf{u}}_p = \mathbf{W}_D \hat{\mathbf{u}}_p = \mathbf{W}_D \hat{\mathbf{Q}}_1^{\frac{1}{2}} \hat{\mathbf{u}}_p$.

Using the definition of $\mathbf{C}(\mathbf{a})$, (43) can be recast as

$$\mathbf{C}(\mathbf{a})[\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_L] = \mathbf{C}(\mathbf{a})\mathbf{U}_s = \mathbf{0}_{(|\mathcal{I}|-L) \times L}. \quad (52)$$

According to the definitions of \mathbf{W} , $\hat{\mathbf{e}}$ and $\mathbf{C}(\mathbf{a})$, and using (51) as well as (52), $\hat{\mathbf{e}}$ can be found as

$$\begin{aligned} \hat{\mathbf{e}} &= \text{vec}\{\{\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \dots \hat{\mathbf{e}}_L\}\} = \text{vec}\{\mathbf{C}(\mathbf{a})\hat{\mathbf{U}}_s\} \\ &= \text{vec}\{\mathbf{C}(\mathbf{a})(\mathbf{U}_s + \Delta\mathbf{U}_s)\} = \text{vec}\{\mathbf{C}(\mathbf{a})\Delta\mathbf{U}_s\} \\ &= (\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a}))\text{vec}\{\Delta\mathbf{U}_s\} = (\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a}))\Delta\mathbf{u}_s \end{aligned} \quad (53)$$

where $\Delta\mathbf{U}_s \triangleq [\Delta\mathbf{u}_1 \dots \Delta\mathbf{u}_L] \in \mathbb{C}^{M \times L}$ and $\Delta\mathbf{u}_s \triangleq \text{vec}\{\Delta\mathbf{U}_s\} \in \mathbb{C}^{ML}$. The identity $\text{vec}\{\mathbf{XYZ}\} = (\mathbf{Z}^T \otimes \mathbf{X})\text{vec}\{\mathbf{Y}\}$ has also been used in (53). Moreover, according to the definitions of \mathbf{W} and $\hat{\mathbf{e}}$, and using (53), \mathbf{W} can be expressed as

$$\mathbf{W} = \left[(\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a})) \mathbb{E}\{\Delta\mathbf{u}_s \Delta\mathbf{u}_s^H\} (\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a}))^H \right]^{-1}. \quad (54)$$

Using the first-order perturbation expansion for SVD [22], [55], we can write that

$$\Delta\mathbf{U}_s \approx (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H) \bar{\mathbf{N}} \mathbf{V}_s \Sigma_s^{-1}. \quad (55)$$

Applying the vectorization operator to (55), we have

$$\Delta\mathbf{u}_s = \text{vec}\{\Delta\mathbf{U}_s\} \approx (\Sigma_s^{-1} \mathbf{V}_s^T \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H)) \bar{\mathbf{n}} \quad (56)$$

where $\bar{\mathbf{n}} \triangleq \text{vec}\{\bar{\mathbf{N}}\} \in \mathbb{C}^{MN}$. Inserting (56) into (54) yields

$$\begin{aligned} \mathbf{W} &\approx [(\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a})) (\Sigma_s^{-1} \mathbf{V}_s^T \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H)) \\ &\times \mathbb{E}\{\bar{\mathbf{n}} \bar{\mathbf{n}}^H\} (\mathbf{V}_s^* \Sigma_s^{-1} \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H)) (\mathbf{I}_L \otimes \mathbf{C}^H(\mathbf{a}))]^{-1} \\ &= [(\Sigma_s^{-1} \mathbf{V}_s^T \otimes \mathbf{C}(\mathbf{a})(\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H)) (\mathbf{I}_N \otimes \mathbf{I}_M) \\ &\times (\mathbf{V}_s^* \Sigma_s^{-1} \otimes (\mathbf{I}_M - \mathbf{U}_s \mathbf{U}_s^H) \mathbf{C}^H(\mathbf{a}))]^{-1} \end{aligned} \quad (57)$$

where $\mathbb{E}\{\bar{\mathbf{n}} \bar{\mathbf{n}}^H\} = \mathbf{I}_{MN} = (\mathbf{I}_N \otimes \mathbf{I}_M)$. The property $(\mathbf{X} \otimes \mathbf{Y})(\mathbf{Z} \otimes \mathbf{T}) = (\mathbf{XZ} \otimes \mathbf{YT})$ has been used for deriving (57). Using this property again together with (52), (57) can be further simplified as

$$\begin{aligned} \mathbf{W} &\approx [(\Sigma_s^{-1} \mathbf{V}_s^T \mathbf{V}_s^* \Sigma_s^{-1} \otimes \mathbf{C}(\mathbf{a}) \mathbf{C}^H(\mathbf{a}))]^{-1} \\ &= (\Sigma_s^{-2} \otimes \mathbf{C}(\mathbf{a}) \mathbf{C}^H(\mathbf{a}))^{-1} = (\Sigma_s^2 \otimes (\mathbf{C}(\mathbf{a}) \mathbf{C}^H(\mathbf{a})))^{-1}. \end{aligned} \quad (58)$$

The identities $\mathbf{V}_s^T \mathbf{V}_s^* = \mathbf{I}_L$ and $(\mathbf{X} \otimes \mathbf{Y})^{-1} = (\mathbf{X})^{-1} \otimes (\mathbf{Y})^{-1}$ have also been used here.

However, since the matrix Σ_s is unknown, we replace it with $\hat{\Sigma}_s$ obtained from (46). Eventually, $\hat{\mathbf{W}}$ is expressed as

$$\hat{\mathbf{W}} = (\hat{\Sigma}_s^2 \otimes (\mathbf{C}(\mathbf{a}) \mathbf{C}^H(\mathbf{a})))^{-1}. \quad (59)$$

It can now be seen from (59) that $\hat{\mathbf{W}}$ is a function of \mathbf{a} . As a result, an iterative scheme should be used to estimate $\hat{\mathbf{a}}$ and $\hat{\mathbf{W}}$ via (50) and (59) in alternative manner. The LS solution of (48) can be used as the initial vector for $\hat{\mathbf{a}}$, i.e.,

$$\hat{\mathbf{a}}^{(0)} \triangleq \hat{\mathbf{a}}_{LS} = \hat{\mathbf{H}}^\dagger \hat{\mathbf{h}}. \quad (60)$$

Initializing $\mathbf{C}(\mathbf{a})$ by inserting (60) into the definition of $\mathbf{C}(\mathbf{a})$, enables us to obtain

$$\hat{\mathbf{W}}^{(0)} = (\hat{\Sigma}_s^2 \otimes (\mathbf{C}(\hat{\mathbf{a}}^{(0)}) \mathbf{C}(\hat{\mathbf{a}}^{(0)})^H)^{-1}). \quad (61)$$

Subsequently, a new estimate of $\hat{\mathbf{a}}$ is generated by substituting (61) into (50). The iterations carry on until a proper termination criteria is satisfied.⁵

The remaining problem is still how to determine the members of \mathcal{I} , i.e., the indices of the selected equations of (41) to be employed for estimating DOAs. Given the cardinality $|\mathcal{I}|$, it is reasonable to select those indices which correspond to the entries of $\hat{\mathbf{u}}_1 = \text{DFT}\{\hat{\mathbf{u}}_1\}$ with $|\mathcal{I}|$ largest absolute values. According to (47), $\hat{\mathbf{u}}_1 = \hat{\mathbf{Q}}_1^{\frac{1}{2}} \hat{\mathbf{u}}_1$ with $\hat{\mathbf{u}}_1$ denoting the left singular vector of $\hat{\mathbf{X}}$ which corresponds to the largest singular value. The logic of this choice is rooted in (32), where it is indicated that each $\hat{\mathbf{u}}_p$ can be expressed as a linear combination of the columns of \mathbf{A} . Therefore, picking the indices of $\hat{\mathbf{u}}_1$ with largest absolute values is a sensible choice because of the following three reasons. 1) The structure of the DFT basis is completely matched with the columns of \mathbf{A} , and consequently $\hat{\mathbf{u}}_p$'s, which makes the absolute values of $\hat{\mathbf{u}}_p$'s the best option to be used for selecting the most relevant equations. 2) Choosing the indices with largest absolute values guarantees picking equations with the most contributions. 3) The estimation error of finding $\hat{\mathbf{u}}_1$ is the smallest among $\hat{\mathbf{u}}_p$'s since it associates with the largest singular values, resulting in smaller error. Finally, the steps required for implementing the proposed DOA estimation method are summarized in Algorithm 3.

To provide a theoretical measure for performance of the proposed forward-only DOA estimation method, the asymptotic variance of the l th DOA estimated by the proposed forward-only algorithm is derived under a high SNR assumption in the following proposition.

Proposition 1: The asymptotic variance of the proposed forward-only DOA estimation algorithm, for a particular matrix $\mathbf{Z}_{\mathcal{I}}$, is given as

$$\mathbb{E}\{\Delta\theta_l^2\} \approx \frac{1}{2} \left(\frac{\lambda}{2\pi d \cos(\theta_l)} \right)^2 \frac{\gamma_l^T (\mathbf{H}^H \mathbf{W} \mathbf{H})^{-1} \gamma_l^*}{|\phi_l|^2} \quad (62)$$

where $\gamma_l \triangleq [\gamma_l^{L-1} \dots 1]^T$ and $\phi_l \triangleq L\gamma_l^{L-1} + (L-1)[\mathbf{a}]_1 \gamma_l^{L-2} + \dots + [\mathbf{a}]_{L-1}$.

Proof: See Appendix B. ■

We aim to generate double number of DOA candidates⁶ by running the proposed Algorithm 2 twice for two different values of $|\mathcal{I}|$. Then a proper DOA selection

⁵Although any common termination criteria can be adopted, we observe that performing 3 to 5 iterations are usually sufficient to obtain a precise result. Therefore, in the numerical examples, 5 is opted as the number of iterations for implementing the proposed methods. It should also be pointed out that the discussion provided in Remark 3 is also valid here, hence the monotone convergence is guaranteed.

⁶In fact, the number of DOA candidates can be arbitrary, but from diverse numerical simulations conducted, we find that generating more candidates than $2L$ does not improve the DOA estimation accuracy considerably.

Algorithm 3: Forward-only DOA Estimation

- 1: Compute $\hat{\mathbf{R}} = 1/N \sum_{t=1}^N \mathbf{x}(t)\mathbf{x}^H(t)$ and estimate $\hat{\mathbf{Q}}$ using Algorithm 1.
- 2: Calculate $\hat{\mathbf{X}} = \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{X}$ and construct \mathbf{W}_a as well as $\overline{\mathbf{W}}$.
- 3: For a pre-chosen $|\mathcal{I}|$, determine $\mathbf{Z}_{\mathcal{I}}$ and obtain \mathbf{B} via performing the SVD of the matrix $\hat{\mathbf{Z}}_{\mathcal{I}} = \mathbf{Z}_{\mathcal{I}}\overline{\mathbf{W}}$ so that the condition $\mathbf{B}^H \hat{\mathbf{Z}}_{\mathcal{I}} = \mathbf{0}$ is satisfied.
- 4: Carry out the SVD of $\hat{\mathbf{X}}$ to obtain $\hat{\mathbf{U}}_s$ and $\hat{\Sigma}_s$, where the former contains the L left singular vectors corresponding to the L largest singular values on the diagonal of the latter.
- 5: Utilize (47) for computing $\hat{\mathbf{U}}_s$. Apply DFT on the columns of $\hat{\mathbf{U}}_s$ to obtain $\hat{\mathbf{u}}_p$'s.
- 6: Form $\hat{\mathbf{H}}$ and $\hat{\mathbf{h}}$ using the estimates $\hat{\mathbf{u}}_p$'s instead of \mathbf{u}_p 's.
- 7: Set $i = 0$ and $\hat{\mathbf{a}}^{(0)} = \hat{\mathbf{H}}^\dagger \hat{\mathbf{h}}$. In addition, set the maximum number of iterations $i_{max} = 5$.
- while** $i \leq i_{max}$
- 8: Compute $\mathbf{C}(\hat{\mathbf{a}}^{(i)})$ and $\hat{\mathbf{W}}^{(i)}$.
- 9: Generate a new estimate:

$$\hat{\mathbf{a}}^{(i+1)} = \left(\hat{\mathbf{H}}^H \hat{\mathbf{W}}^{(i)} \hat{\mathbf{H}} \right)^{-1} \hat{\mathbf{H}}^H \hat{\mathbf{W}}^{(i)} \hat{\mathbf{h}}.$$
- 10: set $i = i + 1$.
- end**
- 11: Find the L roots of the polynomial $\gamma^L + \sum_{l=1}^L [\hat{\mathbf{a}}]_l \gamma^{L-l} = 0$, denoted by $\hat{\gamma}_l, l = 1, \dots, L$.
- 12: Obtain the L DOA estimates as $\hat{\theta}_l = \arcsin\left(-\frac{\beta_l \lambda}{2\pi d}\right)$, where β_l is the phase argument of $\hat{\gamma}_l, l = 1, \dots, L$.

strategy should be employed to determine the final DOA estimates. First, consider the following example to get a more through wisdom about how different choices of $|\mathcal{I}|$ affect the DOA estimation accuracy.

Illustrative Example 1: Consider a scenario where the signals of two uncorrelated sources located in $\boldsymbol{\theta} = [-2^\circ, 7^\circ]$ are received by a ULA consisting $M = 8$ sensors with half waveform adjacent distances. The sensor noise covariance matrix is set as $\mathbf{Q} = \text{diag}\{[10, 1.2, 3.5, 18, 2, 8.5, 24, 6.5]\}$, the sample size is $N = 40$, and 2000 Monte Carlo runs are conducted to calculate RMSE defined as

$$\text{RMSE} = 10 \log_{10} \sqrt{\frac{1}{2000L} \sum_{l=1}^L \sum_{i=1}^{2000} (\hat{\theta}_{l,i} - \theta_l)^2}. \quad (63)$$

We also include the deterministic CRB [30] as a benchmark. In Fig. 1, higher accuracy in DOA estimation can be observed for larger $|\mathcal{I}|$. From this observation, we conclude that $|\mathcal{I}| = M - 1$ and $|\mathcal{I}| = M$ are the best choices in the sense of providing the most precise estimates for generating double number of DOA candidates.

B. DOA Selection Strategy

After running Algorithm 3 twice with $|\mathcal{I}| = M - 1$ and $|\mathcal{I}| = M$, $2L$ DOA candidates are generated. The natural question is how to select L final DOA estimates. Two known conventional approaches to DOA selection are based on the CB [14], [56]–[58] and ML cost function minimization [20], [21], [59]. Recently, another method, which has been originally proposed for joint source number detection and DOA estimation [60], has been employed as the DOA selection scheme. It is based on the generalized likelihood ratio (GLR), which extracts

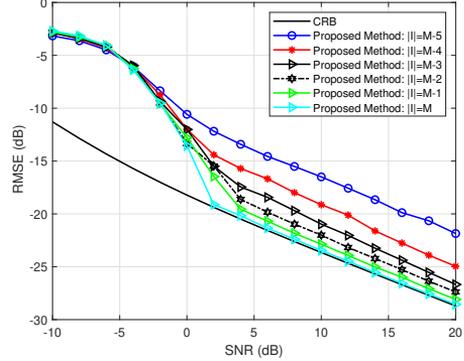


Fig. 1: RMSE performance of the proposed method for different $|\mathcal{I}|$ vs. SNR for $L = 2$ uncorrelated sources with $\boldsymbol{\theta} = [-2^\circ, 7^\circ]$, $M = 8$, and $N = 40$.

the final L DOAs sequentially [27]. The computational complexity of such DOA selection strategy is much lower than that of based on the ML cost function minimization. Besides, the performance provided by the GLR is comparable with that provided by the ML-based methods. Here we aim to design a DOA selection strategy which takes advantage of the three aforementioned approaches, i.e., it uses the CB, deterministic ML cost function, and GLR technique.

In doing so, the following three-step DOA selection strategy is proposed.

Step 1: Denote the vector containing all $2L$ DOA candidates by $\boldsymbol{\theta}_{2L}$. Calculate the CB output⁷ for a proper number of equidistant points to cover the whole interval of interest⁸, i.e., $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Then, find the $(L + 1)$ st peak and define a threshold, denoted by η , as the CB output at the $(L + 1)$ st peak. Afterwards, calculate the CB output for entries of $\boldsymbol{\theta}_{2L}$ and stack those entries with output larger than η in a new vector $\tilde{\boldsymbol{\theta}}$. If the total number of peaks in the CB output is smaller than $(L + 1)$ as well as the number of entries in $\boldsymbol{\theta}$ is either less than L or equal to $2L$, then let $\tilde{\boldsymbol{\theta}}$ contain only the L DOA candidates generated by $|\mathcal{I}| = M$.

Step 2: Determine the first DOA as that entry of $\tilde{\boldsymbol{\theta}}$ which maximizes the GLR, i.e.,

$$\hat{\theta}_1 = \arg \max_{\theta} \frac{\mathbf{a}^H(\theta) \hat{\mathbf{Q}}^{-1} \hat{\mathbf{R}} \hat{\mathbf{Q}}^{-1} \mathbf{a}(\theta)}{\mathbf{a}^H(\theta) \hat{\mathbf{Q}}^{-1} \mathbf{a}(\theta)}, \quad \theta \in \tilde{\boldsymbol{\theta}}. \quad (64)$$

The GLR in (64) is a straightforward unknown noise extension of the GLR in [60] for the case of uniform noise.

Step 3: Denote the remaining entries of $\tilde{\boldsymbol{\theta}}$ as $\bar{\boldsymbol{\theta}}$, and the size of $\bar{\boldsymbol{\theta}}$ as \bar{L} . Then, divide the \bar{L} DOA candidates in $\bar{\boldsymbol{\theta}}$ into $\bar{G} = \frac{\bar{L}}{(L-1)!(L-L+1)!}$ subsets containing $(L - 1)$

⁷The CB output for the direction θ is calculated as $L(\theta) = \mathbf{a}(\theta)^H \hat{\mathbf{R}} \mathbf{a}(\theta)$.

⁸We use 314 equidistant points in this paper.

different DOAs each. Denote these subsets as $\Theta_1, \dots, \Theta_{\bar{G}}$ and associate them with $\mathbf{A}(\Theta_1), \dots, \mathbf{A}(\Theta_{\bar{G}})$. The subset that minimizes the following deterministic ML cost function⁹ determines the $(L-1)$ remaining DOAs

$$\hat{\Theta}_R = \arg \min_{\Theta_S} \text{trace} \left[\left(\mathbf{P}_{\hat{\mathbf{A}}(\Theta_S)}^\perp - \nu_1 \nu_1^H \right) \hat{\mathbf{Q}}^{-\frac{1}{2}} \hat{\mathbf{R}} \hat{\mathbf{Q}}^{-\frac{1}{2}} \right],$$

$$\mathbb{S} \in \{1, \dots, \bar{G}\} \quad (65)$$

where

$$\mathbf{P}_{\hat{\mathbf{A}}(\Theta_S)}^\perp \triangleq \mathbf{I}_M - \tilde{\mathbf{A}}(\Theta_S) \left(\tilde{\mathbf{A}}(\Theta_S)^H \tilde{\mathbf{A}}(\Theta_S) \right)^{-1} \tilde{\mathbf{A}}(\Theta_S)^H. \quad (66)$$

and

$$\tilde{\mathbf{A}}(\Theta_S) \triangleq \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A}(\Theta_S) \quad (67)$$

$$\nu_1 \triangleq \frac{\mathbf{P}_{\hat{\mathbf{A}}(\Theta_S)}^\perp \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{a}(\hat{\theta}_1)}{\|\mathbf{P}_{\hat{\mathbf{A}}(\Theta_S)}^\perp \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{a}(\hat{\theta}_1)\|_2}. \quad (68)$$

In (65), the contribution of previously estimated $\hat{\theta}_1$ on the deterministic ML cost function is isolated in ν_1 thanks to the properties of the orthogonal projection matrix [61]. In addition, $\hat{\mathbf{Q}}^{-\frac{1}{2}} \hat{\mathbf{R}} \hat{\mathbf{Q}}^{-\frac{1}{2}}$ is used instead of $\hat{\mathbf{R}}$ to consider the general case of unknown noise.

Finally, $\hat{\theta}_1$ and $\hat{\Theta}_R$ obtained via (65) form together the final L DOA estimates.

C. FB Extension

FB is a natural extension/improvement of the forward-only-based DOA estimation methods [16], [28], [35], [62]–[64]. The essence of FB is to first transform the observed signal or SCM into a new centro-Hermitian signal matrix or centro-Hermitian covariance matrix, respectively, followed by computationally simplified and more accurate DOA estimation. Higher DOA estimation accuracy is a consequence of decorrelating possibly correlated source pairs and obtaining more accurate SCM estimation.

The FB covariance matrix is given as [13]

$$\mathbf{R}_{\text{FB}} = \frac{1}{2} (\mathbf{R} + \mathbf{J}_M \mathbf{R}^* \mathbf{J}_M)$$

$$= \frac{1}{2} (\mathbf{A} \mathbf{P} \mathbf{A}^H + \mathbf{Q} + \mathbf{J}_M (\mathbf{A}^* \mathbf{P}^* \mathbf{A}^T + \mathbf{Q}^*) \mathbf{J}_M). \quad (69)$$

It can be readily verified that \mathbf{R}_{FB} is centro-Hermitian, i.e., $\mathbf{R}_{\text{FB}} = \mathbf{J}_M \mathbf{R}_{\text{FB}}^* \mathbf{J}_M$. Rearranging (69), we have

$$\mathbf{R}_{\text{FB}} = \frac{1}{2} (\mathbf{A} \mathbf{P} \mathbf{A}^H + \mathbf{Q} + \mathbf{A} \mathbf{D} \mathbf{P}^* \mathbf{D}^H \mathbf{A}^H + \mathbf{J}_M \mathbf{Q}^* \mathbf{J}_M)$$

$$= \mathbf{A} \tilde{\mathbf{P}} \mathbf{A}^H + \frac{1}{2} \tilde{\mathbf{Q}} \quad (70)$$

where $\tilde{\mathbf{P}} \triangleq \frac{1}{2} (\mathbf{P} + \mathbf{D} \mathbf{P}^* \mathbf{D}^H)$, $\tilde{\mathbf{Q}} \triangleq \mathbf{Q} + \mathbf{J}_M \mathbf{Q}^* \mathbf{J}_M$, $\mathbf{D} \triangleq \text{diag} \{ e^{-j(2\pi/\lambda)d(M-1)\sin(\theta_1)}, \dots, e^{-j(2\pi/\lambda)d(M-1)\sin(\theta_L)} \}$.

⁹Since the difference of employing the deterministic ML and stochastic ML is marginal, the deterministic ML is discussed here because it also has lower complexity.

According to (70), a proper centro-Hermitian matrix similar to $\hat{\mathbf{X}}$ defined in (45), can be formed as

$$\hat{\mathbf{X}}_{\text{FB}} = \left[\hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{X} \quad \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{J}_M \mathbf{X}^* \mathbf{J}_M \right] \quad (71)$$

where $\hat{\tilde{\mathbf{Q}}} \triangleq \hat{\mathbf{Q}} + \mathbf{J}_M \hat{\mathbf{Q}}^* \mathbf{J}_M$ is used instead of $\tilde{\mathbf{Q}}$. The matrix $\hat{\mathbf{X}}_{\text{FB}}$ can be decomposed by applying the truncated SVD as $\hat{\mathbf{X}}_{\text{FB}} = \hat{\mathbf{E}}_s \hat{\mathbf{\Pi}}_s \hat{\mathbf{T}}_s^H$ where $\hat{\mathbf{E}}_s = [\hat{\mathbf{e}}_1 \dots \hat{\mathbf{e}}_L] \in \mathbb{C}^{M \times L}$ is composed of L left singular vectors associated with L largest singular values on the diagonal of $\hat{\mathbf{\Pi}}_s \in \mathbb{R}^{L \times L}$. The columns of $\hat{\mathbf{E}}_s$ and the columns of $\hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A}$ span the same vector space. Thus, similar to the forward-only case, the relationship between $\hat{\mathbf{E}}_s$ and $\hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A}$ can be written as $\hat{\mathbf{E}}_s = \hat{\mathbf{Q}}^{-\frac{1}{2}} \mathbf{A} \mathbf{G}$. Multiplying both sides of the latter equation by $\hat{\mathbf{Q}}^{\frac{1}{2}}$, we obtain the FB analog of (32), that is, $\hat{\tilde{\mathbf{E}}}_s \triangleq \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{E}}_s = \mathbf{A} \mathbf{G}$.

Following the same steps as in Subsection A, the DFT of the columns of $\hat{\tilde{\mathbf{E}}}_s$ can be found as

$$\hat{\tilde{\mathbf{E}}}_s = [\hat{\mathbf{e}}_1 \dots \hat{\mathbf{e}}_L] \triangleq \text{DFT}\{\hat{\mathbf{E}}_s\} = \text{DFT}\{\hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{E}}_s\}$$

$$= \left[(\mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{e}}_1) \dots (\mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \hat{\mathbf{e}}_L) \right] \in \mathbb{C}^{M \times L}. \quad (72)$$

Considering a specific value for $|\mathcal{I}|$, a system of linear equations similar to (48) can be formulated as

$$\hat{\mathbf{H}} \mathbf{a} \approx \hat{\mathbf{h}} \quad (73)$$

where $\hat{\mathbf{H}}_p = \mathbf{B}^H \text{diag}\{\mathbf{Z}_T \hat{\mathbf{e}}_p\} \mathbf{Z}_T \mathbf{W}_a \in \mathbb{C}^{(|\mathcal{I}|-L) \times L}$, $\hat{\mathbf{h}}_p = -\mathbf{B}^H \mathbf{Z}_T \hat{\mathbf{e}}_p \in \mathbb{C}^{(|\mathcal{I}|-L)}$, $\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_1^T \dots \hat{\mathbf{H}}_L^T \right]^T \in \mathbb{C}^{L(|\mathcal{I}|-L) \times L}$, $\hat{\mathbf{h}} = \left[\hat{\mathbf{h}}_1^T \dots \hat{\mathbf{h}}_L^T \right]^T \in \mathbb{C}^{L(|\mathcal{I}|-L)}$ with the entries of \mathbf{a} being the coefficients of the polynomial presented in (39). Finally, the GLS solution of (73) is given by

$$\hat{\mathbf{a}} = \left(\hat{\mathbf{H}}^H \hat{\mathbf{W}}_{\text{FB}} \hat{\mathbf{H}} \right)^{-1} \hat{\mathbf{H}}^H \hat{\mathbf{W}}_{\text{FB}} \hat{\mathbf{h}} \quad (74)$$

where $\hat{\mathbf{W}}_{\text{FB}} \approx \hat{\mathbf{\Pi}}_s^2 \otimes (\mathbf{C}_{\text{FB}}(\mathbf{a}) \mathbf{C}_{\text{FB}}(\mathbf{a})^H)^{-1}$, $\mathbf{C}_{\text{FB}}(\mathbf{a}) \triangleq \mathbf{B}^H (\mathbf{I}_{|\mathcal{I}|} + \text{diag}\{\mathbf{Z}_T \mathbf{W}_a \mathbf{a}\}) \mathbf{Z}_T \mathbf{W}_D \hat{\mathbf{Q}}^{\frac{1}{2}} \in \mathbb{C}^{(|\mathcal{I}|-L) \times M}$.

After finding $\hat{\mathbf{a}}$ using (74), the L DOA estimates $\hat{\theta}_l$, $l = 1, \dots, L$ are obtained as $\hat{\theta}_l = \arcsin\left(-\frac{\beta_l \lambda}{2\pi d}\right)$ with β_l denoting the phase argument of $\hat{\gamma}_l$. In addition, $\hat{\gamma}_l$ denotes the l th root of the polynomial defined as $\gamma^L + \sum_{l=1}^L [\hat{\mathbf{a}}]_l \gamma^{L-l} = 0$.

D. Computational Complexity

For implementing the proposed DOA estimation method, Algorithms 1 or 2 should be implemented first that requires $\mathcal{O}(M^2 N)$ flops for calculating $\hat{\mathbf{R}}$ and $\mathcal{O}(M^3)$ flops for computing GED of the pair

of matrices $\{\hat{\mathbf{R}}, \hat{\mathbf{Q}}^{(i)}\}$ in the i th alternation. Thus, the total computational complexity of Algorithm 1 is $\mathcal{O}(I_1(M^3) + M^2N)$ with I_1 being the number of alternations. The computational complexity of performing SVD of $\tilde{\mathbf{X}}$ is $\mathcal{O}(\max(M, N)\min(M, N)^2)$, and the computational complexity of computing \mathbf{B} is $\mathcal{O}(|\mathcal{I}|L^2)$. Applying DFT on the columns of $\tilde{\mathbf{U}}_s$ requires $\mathcal{O}(LM\log_2(M))$ flops. In addition, the main source of computational complexity for calculating $\tilde{\mathbf{W}}$ is the inversion of the matrix $\mathbf{C}(\mathbf{a})\mathbf{C}^H(\mathbf{a})$, which involves $\mathcal{O}((|\mathcal{I}| - L)^3)$ flops, making the computational complexity of computing $\tilde{\mathbf{W}}$ be $\mathcal{O}(L(|\mathcal{I}| - L)^3)$. Determining $\hat{\mathbf{a}}$ requires $\mathcal{O}(L^3(2(M - L) + 1) + L^2(2(M - L)^2 + (M - L)))$ flops. As the proposed DOA estimation algorithm is run twice with $|\mathcal{I}| = M - 1$ and $|\mathcal{I}| = M$, by considering $|\mathcal{I}| \approx M$, the total complexity required to generate $2L$ DOA candidates is $\mathcal{O}(M^2N + I_1(M^3) + (\max(M, N)\min(M, N)^2) + ML^2 + LM\log_2(M) + I_2(M^3L + ML^3 - L^4))$ where I_2 denotes the number of iterations needed for finding the best $\hat{\mathbf{a}}$. At last, the complexity of DOA selection is mainly in *step 3*, which is about $\mathcal{O}(\bar{G}(M^3 + 3M(L - 1)^2 + (L - 1)^3))$. Finally, the total complexity of DOA estimation is reduced to $\mathcal{O}(M^2N + \bar{G}M^3)$ in the case that $\bar{G} \gg \max(I_1, I_2)$ and $M \gg L$. The significant point here is that \bar{G} is approximately four times smaller than the parameter G in [21], resulting in the computational complexity of our method being approximately a quarter of what is required for implementing the EPUMA.

V. SIMULATION RESULTS

The aim of this section is to evaluate the performance of the proposed method and compare it to that of the state-of-the-art algorithms in terms of diverse numerical simulation examples especially for challenging scenarios. Our examples address uniform, nonuniform, and block-diagonal sensor noise cases. For the uniform noise case, the performance of the forward-only and FB versions of the proposed method is compared with that of the unitary root-MUSIC method [13], the root-swap unitary root-MUSIC method [20], the FB EPUMA method [21], and the UE GLS [22]. For the nonuniform noise case, the “NISB+MUSIC” method [42], the “IMLSE+MUSIC” method [33], the SPICE, LIKES and SLIM methods [39] are used for comparison. In addition, for achieving better DOA estimations as well as providing an evidence on the effectiveness of phases 2 and 3 over traditional super-resolution methods, the combinations of both the IMLSE and proposed nonuniform noise covariance estimator (see Algorithm 1) with the unitary root-MUSIC framework are considered for the nonuniform noise case. For the block-diagonal noise case, the approximate technique by Agrawal and Prasad [47], and the combinations of both the “extended IMLSE” and proposed block-diagonal

noise covariance estimator (see Algorithm 2) with the unitary root-MUSIC framework are used for comparison. The uniform stochastic CRB [26], the nonuniform stochastic CRB [30], and stochastic CRB in unknown noise fields [48] are used as the benchmarks in the corresponding examples. The number of trials used for calculating the RMSE is 2000 in all examples. If not further specified, a ULA with $M = 10$ sensors separated by half wavelength collecting $N = 10$ snapshots is considered for the uniform noise examples, while $M = 8$ for the nonuniform noise examples. The SNR is computed as $\text{SNR} = \frac{\sigma_s^2}{\sigma_n^2}$ for the uniform noise case, as $\text{SNR} = \frac{\sigma_s^2}{M} \sum_{m=1}^M \frac{1}{\sigma_m^2}$ for the nonuniform noise case, and as $\text{SNR} = \frac{\sigma_s^2}{M} \sum_{m=1}^M \frac{1}{[\mathbf{Q}_{\text{bdiag}}]_{m,m}}$ for the block-diagonal noise case. Here the powers of different sources are considered to be identical and denoted by σ_s^2 . Moreover, ϵ is set to 10^{-4} in Algorithms 1 and 2.

The first three examples are associated with the uniform sensor noise case. In the first example, three uncorrelated sources located at $\boldsymbol{\theta} = [19^\circ, 34^\circ, 36^\circ]$ are considered. It can be seen in Fig. 2 that the SNR threshold performance of the proposed methods is outstandingly better than that of the other methods tested. As illustrated in Fig. 3, the impact of the number of snapshots on the performance of the methods tested is investigated via setting $\boldsymbol{\theta} = [34^\circ, 38^\circ]$, $\rho = 0.95$ for the fixed SNR = 2 dB, where ρ is the correlation coefficient throughout the paper. It displays that the FB version of the proposed method provides robust estimates even when the number of snapshots is about one order of magnitude smaller than that of the other methods tested. In the next setup, the capability of the methods tested to deal with the scenario of two closely located sources is investigated. In doing so, we regard the setup in which $\boldsymbol{\theta} = [0^\circ, 34^\circ, (34 + \Delta\theta)^\circ]$, $\rho = 0.95$, and SNR = 15 dB with $\Delta\theta$ varying from 0.8° to 6° . It can be seen in Fig. 4 that the FB version of the proposed method has the best performance.

Unless otherwise stated, to study the impact of the nonuniform noise, the sensor noise covariance matrix is set as $\mathbf{Q} = \text{diag}\{[6, 2, 0.5, 2.5, 3, 1, 5.5, 10]\}$ for the follow up examples [33]. The worst noise power ratio (WNPR) used in these examples is given as $\text{WNPR} = \frac{\sigma_{\text{max}}^2}{\sigma_{\text{min}}^2} = \frac{10}{0.5} = 20$. First, Fig. 5 shows the RMSE performance of the methods tested versus SNR for the setup of $\boldsymbol{\theta} = [33^\circ, 36^\circ]$ and $\rho = 0$. As it can be observed, the threshold performance of the proposed methods is substantially better than that of the other methods tested. Moreover, in Fig. 6, the same scenario as that shown in Fig. 5 is considered for the case of correlated sources $\boldsymbol{\theta} = [33^\circ, 38^\circ]$ and $\rho = 0.95$. Fig. 6 demonstrates the superiority of the FB version of the proposed method over the other methods tested. Moreover, Figs. 5 and 6 confirm the necessity of the second and third phases since the combination of the noise covariance matrix estimator and the unitary root-MUSIC does not provide as accurate

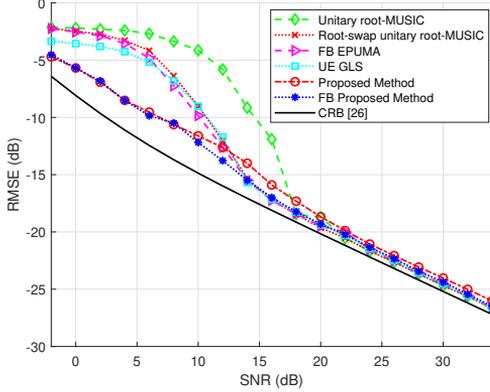


Fig. 2: RMSE vs. SNR for $L = 3$ uncorrelated sources with $\theta = [19^\circ, 34^\circ, 36^\circ]$, $M = 10$, and $N = 10$.

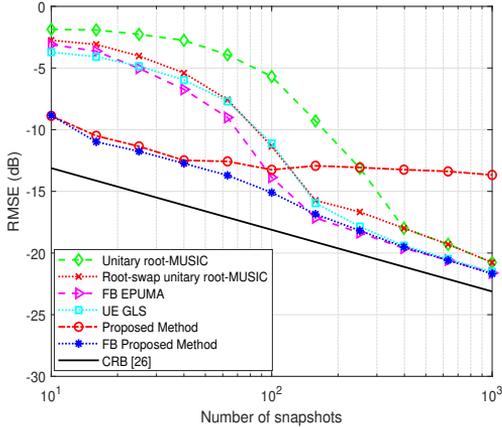


Fig. 3: RMSE vs. the number of snapshots for $L = 2$ correlated sources with $\theta = [34^\circ, 38^\circ]$, $\rho = 0.95$, SNR = 2 dB, and $M = 10$.

results as that of by the proposed method. Additionally, comparing the performance of the combination of the noise covariance estimator and the unitary root-MUSIC with that of the combination of the IMLSE and the unitary root-MUSIC shows that the proposed noise covariance estimator is as accurate as the IMLSE. Fig. 7 depicts how different methods perform depending on the number of snapshots for the scenario of $\theta = [33^\circ, 38^\circ]$, $\rho = 0$, and SNR = 0 dB. Fig. 8 illustrates this dependency also for the scenario of $\theta = [33^\circ, 48^\circ]$, $\rho = 0.95$, and SNR = -4 dB. Based on Figs. 7 and 8, it can be concluded that the reliability of the proposed methods to the scarcity of the number of snapshots is higher for the case of uncorrelated sources compared to the other methods tested, while the FB version of the proposed

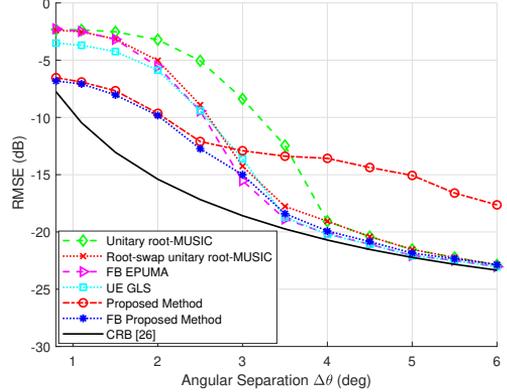


Fig. 4: RMSE vs. the angular separation for $L = 3$ partly correlated sources with $\theta = [0^\circ, 34^\circ, (34 + \Delta\theta)^\circ]$, SNR = 15 dB, $\rho = 0.95$ for the last two directions, $M = 10$, and $N = 10$.

method copes with the correlated sources more efficiently. Moreover, Figs. 9 and 10 show the strengths of the methods tested against the presence of closely located sources for the cases of uncorrelated and correlated signals, respectively. The setup regarded for Fig. 9 is $\theta = [-10^\circ, 34^\circ, (34 + \Delta\theta)^\circ]$, $\rho = 0$, and SNR = 15 dB with $\Delta\theta$ varying from 1° to 12° . It can be observed that the performance of the proposed method almost achieves the CRB. Fig. 10 depicts the results obtained from conducting the same setup as for Fig. 9 with the difference that $\rho = 0.95$ for the last two directions. The superiority of the FB version of the proposed method over other methods tested can be seen to be substantial. To ensure the identifiability of the proposed procedure, Figs. 11 and 12 are dedicated to scenarios where the number of sources is large and close to the number of sensors. Towards this end, the setup regarded for Fig. 11 is $\theta = [21^\circ, -22^\circ, 46^\circ, 0^\circ, -60^\circ, -40^\circ, 65^\circ]$, $\mathbf{Q} = \text{diag}\{[6, 2, 0.5, 2.5, 3, 1, 5.5, 10, 4, 9]\}$, $\rho = 0$, $M = 10$ and $N = 10$. In addition, Fig. 12 depicts the impact of varying WNPR (from 10 to 100) on the performance of the methods tested by considering the setup where $\theta = [34^\circ, 36^\circ, 17^\circ, -2^\circ, -27^\circ, -49^\circ]$, $\rho = 0$, SNR = 20 dB, $M = 8$ and $N = 10$. According to Figs. 11 and 12, the proposed procedure possess the best performance among the methods tested even in the cases when the number of sources is close to the number of sensors and/or the WNPR is large.

For block-diagonal noise case, we assume $M = 20$ and that the noise covariance matrix has the block-diagonal form $\mathbf{Q}_{\text{bdiag}} = \text{bdiag}\{\mathbf{Q}_1, 10, 3, 2, 3, 1, \mathbf{Q}_2, 7, 5, 4, 7, 1, 1, \mathbf{Q}_3\}$ where \mathbf{Q}_1 , \mathbf{Q}_2 , and \mathbf{Q}_3 are the 4×4 , 3×3 , and 2×2 matrices, respectively, constructed as $[\mathbf{Q}_i]_{j,k} = \sigma_i^2 e^{-(j-k)^2 \zeta_i}$. The values $\sigma_1^2 = 5$, $\sigma_2^2 = 3$, $\sigma_3^2 = 1$, $\zeta_1 = 0.7$, $\zeta_2 = 1$,

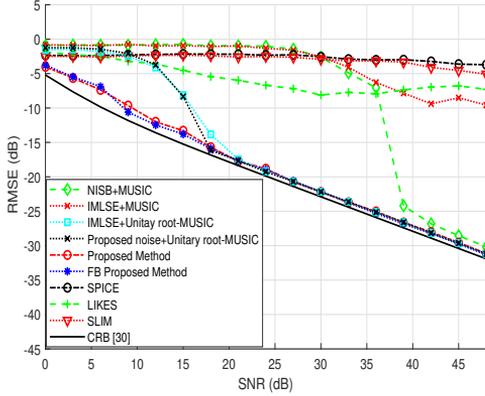


Fig. 5: RMSE vs. SNR for $L = 2$ uncorrelated sources with $\theta = [33^\circ, 36^\circ]$, $M = 8$, and $N = 10$.

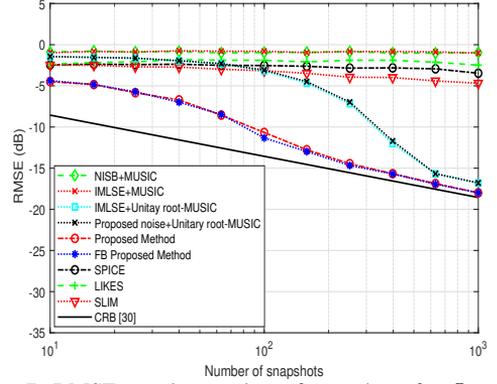


Fig. 7: RMSE vs. the number of snapshots for $L = 2$ uncorrelated sources with $\theta = [33^\circ, 38^\circ]$, SNR = 0 dB, and $M = 8$.

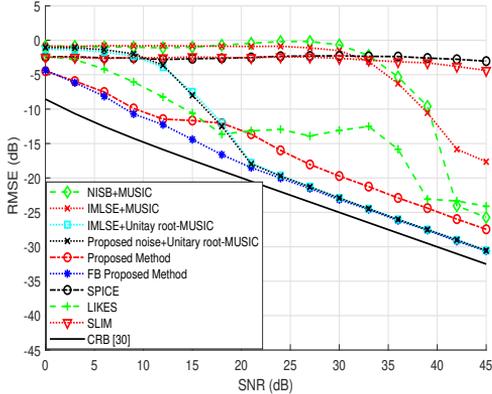


Fig. 6: RMSE vs. SNR for $L = 2$ correlated sources with $\theta = [33^\circ, 38^\circ]$, $\rho = 0.95$, $M = 8$, and $N = 10$.

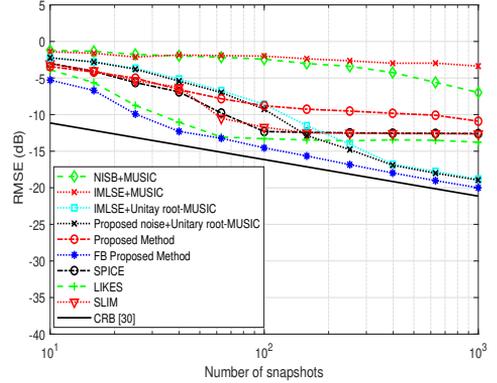


Fig. 8: RMSE vs. the number of snapshots for $L = 2$ correlated sources with $\theta = [33^\circ, 48^\circ]$, $\rho = 0.95$, SNR = -4 dB, and $M = 8$.

and $\zeta_3 = 0.5$ are assumed. Fig. 13 shows the RMSE performance of the methods tested versus SNR for the setup of $\theta = [34^\circ, 36^\circ, -10^\circ]$ and $\rho = 0.95$ for the first two directions. Fig. 14 shows the performance versus the number of snapshots for the scenario of $\theta = [7^\circ, 9^\circ, -18^\circ]$, $\rho = 0.95$ for the first two directions, and SNR = 2 dB. Figs. 13 and 14 confirm the efficiency of the proposed method (Algorithm 2 combined with phases 2 and 3) in coping with the presence of the block-diagonal sensor noise.

To study the convergence of the proposed noise covariance matrix estimators, the cost functions of the proposed nonuniform and block-diagonal noise covariance estimators versus the number of iterations are depicted in Sub-Figures. 15a and 15b, respectively. It can be observed that the cost functions of the proposed noise covariance

matrix estimators converge to their corresponding optimal values relatively fast.

VI. Conclusion and Discussion

A unified procedure for DOA estimation in the presence of unknown noise fields (nonuniform and block-diagonal sensor noise) is introduced. The proposed procedure works in three phases. The goal of the first phase is to devise a robust yet computationally efficient algorithm to estimate the noise covariance matrix in an alternative manner. In each alternation, the noise subspace is estimated using GED first and then the noise covariance estimate is updated as the solution of an LS minimization problem. Compared to the state-of-the-art estimators, the proposed noise covariance estimator is robust in confronting both uncorrelated and correlated sources. After the noise covariance matrix is estimated,

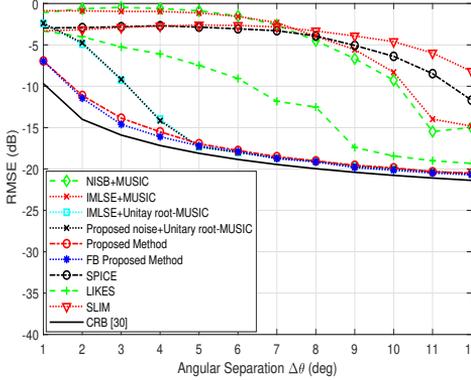


Fig. 9: RMSE vs. the Angular Separation for $L = 3$ uncorrelated sources with $\theta = [-10^\circ, 34^\circ, (34 + \Delta\theta)^\circ]$, SNR = 15 dB, $M = 8$, and $N = 10$.

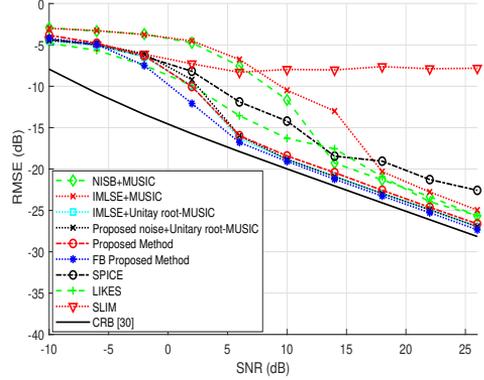


Fig. 11: RMSE vs. SNR for $L = 7$ uncorrelated sources with $\theta = [21^\circ, -22^\circ, 46^\circ, 0^\circ, -60^\circ, -40^\circ, 65^\circ]$, $M = 10$, and $N = 10$.

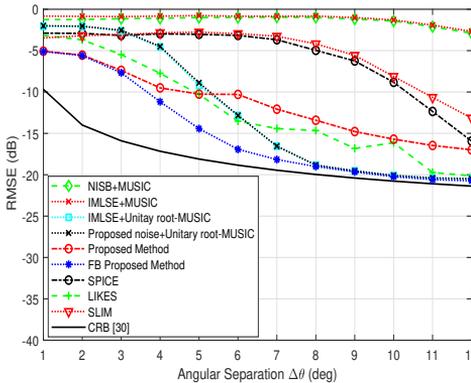


Fig. 10: RMSE vs. the Angular Separation for $L = 3$ partly correlated sources with $\theta = [-10^\circ, 34^\circ, (34 + \Delta\theta)^\circ]$, SNR = 15 dB, $\rho = 0.95$ for the last two directions, $M = 8$, and $N = 10$.

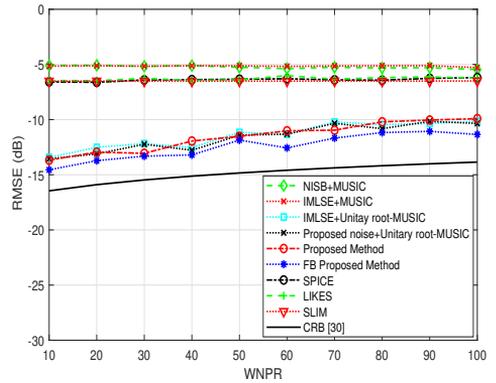


Fig. 12: RMSE vs. WNPR for $L = 6$ uncorrelated sources with $\theta = [34^\circ, 36^\circ, 17^\circ, -2^\circ, -27^\circ, -49^\circ]$, SNR = 20 dB, $M = 8$, and $N = 10$.

$2L$ DOA candidates are estimated using a rooting-based DOA estimation method based on the combination of the GLS and the first-order signal subspace perturbation. The noise covariance estimate obtained in the first phase is used for pre-whitening the array signal. The basic motivation behind generating double number of DOA candidates is to improve the resolution of the DOA estimation approach especially in the cases of small sample size and low SNR. In addition, the forward-only DOA estimation method is extended using FB. Furthermore, the asymptotic performance of both the forward-only and FB versions of the proposed method is studied. In the third phase, the final best L DOA estimates out of $2L$ DOA candidates generated in the second phase are selected using a properly designed DOA selection strategy. This crucial task is accomplished in a novel manner by the

assist of the CB, GLR, and DML. In particular, the role of the CB is to clean outliers that may exist in DOA candidates, while the GLR and DML are used to ensure the selection of the best final DOAs. The limitation of the methodology presented in the second phase is the ULA assumption, and this limitation does not apply to the noise covariance matrix estimators in phase 1 and the DOA selection strategy in phase 3, which can work without any modifications for arrays of arbitrary geometry. Numerical simulation examples are included to show the superiority of the proposed algorithm compared to the state-of-the-art approaches for cases of the uniform, nonuniform, and block-diagonal sensor noise. The remarkable performance improvement is the result of all three phases together, since each sub-algorithm in these three phases is developed in a way to fulfill the corresponding goal with a full use of available information at each phase.

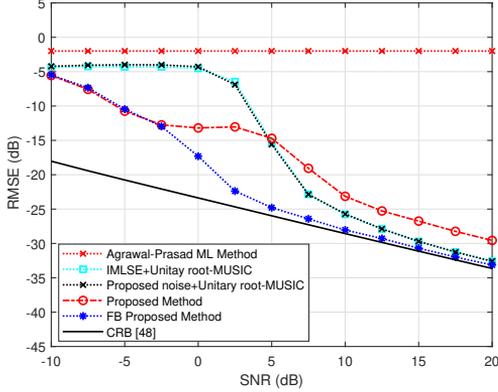


Fig. 13: RMSE vs. SNR for $L = 3$ partly correlated sources with $\theta = [34^\circ, 36^\circ, -10^\circ]$, $\rho = 0.95$, $M = 20$, and $N = 100$.

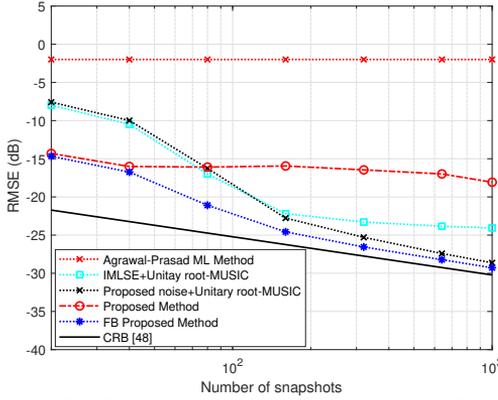
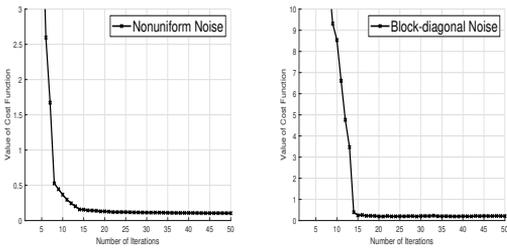


Fig. 14: RMSE vs. the number of snapshots for $L = 3$ partly correlated sources with $\theta = [7^\circ, 9^\circ, -18^\circ]$, $\rho = 0.95$, SNR = 2 dB, and $M = 20$.



(a) Convergence of Algorithm 1 for $M = 8$ and $N = 500$. (b) Convergence of Algorithm 2 for $M = 20$ and $N = 500$.

Fig. 15: Convergence behavior of the proposed noise covariance matrix estimators for $L = 3$ uncorrelated sources with $\theta = [16^\circ, -3^\circ, -16^\circ]$ and SNR = 2 dB.

Appendix A: Proof of (12)

Taking into account the fact that \mathbf{Q} is a real-valued diagonal matrix and also using the derivative properties [65], the partial derivatives of the terms related to \mathbf{Q} in (11) can be found to be

$$\frac{\partial \text{trace} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \mathbf{Q} \right\}}{\partial \mathbf{Q}} = \frac{\partial \text{trace} \left\{ \mathcal{D} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \right\} \mathbf{Q} \right\}}{\partial \mathbf{Q}} = \mathcal{D} \left\{ \hat{\mathbf{R}} \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \right\} \quad (75)$$

$$\frac{\partial \text{trace} \left\{ \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \hat{\mathbf{R}} \mathbf{Q} \right\}}{\partial \mathbf{Q}} = \frac{\partial \text{trace} \left\{ \mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \hat{\mathbf{R}} \right\} \mathbf{Q} \right\}}{\partial \mathbf{Q}} = \mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \hat{\mathbf{R}} \right\} \quad (76)$$

$$\frac{\partial \text{trace} \left\{ \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \mathbf{Q}^2 \right\}}{\partial \mathbf{Q}} = \frac{\partial \text{trace} \left\{ \mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \right\} \mathbf{Q}^2 \right\}}{\partial \mathbf{Q}} = 2\mathcal{D} \left\{ \hat{\mathbf{U}}^{(i)} \hat{\mathbf{U}}^{(i)H} \right\} \mathbf{Q}. \quad (77)$$

Using (75)–(77), the partial derivative of (11) with respect to \mathbf{Q} can be straightforwardly found to be (12).

Appendix B: Proof of Proposition 1

The proof goes in the same steps as that in [21] (Appendix A), and is included for the sake of completeness. As θ_l and γ_l are related to each other as $\gamma_l = e^{-j2\pi d \sin(\theta_l)/\lambda}$, we perform Taylor's expansion and keep only the terms containing up to the first-order perturbation terms to obtain $\Delta\theta_l \approx -\frac{\lambda}{2\pi d \cos(\theta_l)} \frac{\Delta\gamma_l}{j\gamma_l}$. To enforce $\Delta\theta_l$ to be a real-valued quantity, it is reasonable to define

$$\Delta\theta_l \triangleq \frac{1}{2}(\Delta\theta_l + \Delta\theta_l^*) = \frac{1}{2} \frac{j\lambda}{2\pi d \cos(\theta_l)} (\gamma_l^* \Delta\gamma_l - \gamma_l \Delta\gamma_l^*). \quad (78)$$

Using (78), the variance of $\Delta\theta_l$ can be written as

$$\mathbb{E}\{\Delta\theta_l^2\} \approx \frac{1}{2} \left(\frac{\lambda}{2\pi d \cos(\theta_l)} \right)^2 \times (\mathbb{E}\{|\Delta\gamma_l|^2\} - \Re\{\mathbb{E}\{\Delta\gamma_l^2\}(\gamma_l^*)^2\}). \quad (79)$$

Next we need to find an expression that connects $\Delta\gamma_l$ and $\Delta\mathbf{a} \triangleq \hat{\mathbf{a}} - \mathbf{a}$, as (79) is dependent to the statistics of $\Delta\gamma_l$ which are related to the statistics of $\Delta\mathbf{a}$. Towards this end, using the first-order approximation of (39) when γ is replaced by γ_l , we obtain

$$\gamma_l^T \Delta\mathbf{a} + \phi_l \Delta\gamma_l \approx 0 \quad (80)$$

where $\Delta\mathbf{a} \triangleq [\Delta[\mathbf{a}]_1 \cdots \Delta[\mathbf{a}]_L]^T$, $\gamma_l \triangleq [\gamma_l^{L-1} \cdots 1]^T$, $\phi_l \triangleq L\gamma_l^{L-1} + (L-1)[\mathbf{a}]_1\gamma_l^{L-2} + \cdots + [\mathbf{a}]_{L-1}$. From (80), we obtain $\Delta\gamma_l \approx -\frac{\gamma_l^T \Delta\mathbf{a}}{\phi_l}$. Thus, it can be written that

$$\mathbb{E}\{|\Delta\gamma_l|^2\} \approx \frac{\gamma_l^T \mathbb{E}\{\Delta\mathbf{a} \Delta\mathbf{a}^H\} \gamma_l^*}{|\phi_l|^2}. \quad (81)$$

Let us define

$$f(\mathbf{a}) = (\hat{\mathbf{H}}\mathbf{a} - \hat{\mathbf{h}})^H \mathbf{W} (\hat{\mathbf{H}}\mathbf{a} - \hat{\mathbf{h}}). \quad (82)$$

Since $\hat{\mathbf{a}}$ is the vector that minimizes (82), $f'(\hat{\mathbf{a}})$ can be approximated under the assumption of high SNR using Taylor's expansion as [21], [68]

$$0 = f'(\hat{\mathbf{a}}) \approx f'(\mathbf{a}) + f''(\mathbf{a})\Delta\mathbf{a} \quad (83)$$

where $f'(\mathbf{a})$ and $f''(\mathbf{a})$ denote respectively the first and second derivatives of $f(\mathbf{a})$ with respect to \mathbf{a} , which are given as

$$\begin{aligned} f'(\mathbf{a}) &= 2\hat{\mathbf{H}}^H\mathbf{W}(\hat{\mathbf{H}}\mathbf{a} - \hat{\mathbf{h}}) = 2\hat{\mathbf{H}}^H\mathbf{W}\hat{\mathbf{e}} \\ &= 2\hat{\mathbf{H}}^H\mathbf{W}(\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a}))\Delta\mathbf{u}_s \end{aligned} \quad (84)$$

$$f''(\mathbf{a}) = 2\hat{\mathbf{H}}^H\mathbf{W}\hat{\mathbf{H}}. \quad (85)$$

Combining (83), (84) and (85), we get for high SNR that

$$\begin{aligned} \mathbb{E}\{\Delta\mathbf{a}\Delta\mathbf{a}^H\} &\approx (\mathbf{H}^H\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^H\mathbf{W}(\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a})) \\ &\times \mathbb{E}\{\Delta\mathbf{u}_s\Delta\mathbf{u}_s^H\}(\mathbf{I}_L \otimes \mathbf{C}^H(\mathbf{a}))\mathbf{W}\mathbf{H}(\mathbf{H}^H\mathbf{W}\mathbf{H})^{-1}. \end{aligned} \quad (86)$$

Based on (54) and (86), it can be written that

$$\mathbb{E}\{\Delta\mathbf{a}\Delta\mathbf{a}^H\} \approx (\mathbf{H}^H\mathbf{W}\mathbf{H})^{-1}. \quad (87)$$

Consequently, substituting (87) into (81), we have

$$\mathbb{E}\{|\Delta\gamma_l|^2\} \approx \frac{\gamma_l^T(\mathbf{H}^H\mathbf{W}\mathbf{H})^{-1}\gamma_l^*}{|\phi_l|^2}. \quad (88)$$

The final part is to compute $\mathbb{E}\{\Delta\gamma_l^2\}$, which has the following form

$$\mathbb{E}\{\Delta\gamma_l^2\} \approx \frac{\gamma_l^T\mathbb{E}\{\Delta\mathbf{a}\Delta\mathbf{a}^T\}\gamma_l}{\phi_l^2} \quad (89)$$

where $\mathbb{E}\{\Delta\mathbf{a}\Delta\mathbf{a}^T\} \approx (\mathbf{H}^H\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^H\mathbf{W}(\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a})) \times \mathbb{E}\{\Delta\mathbf{u}_s\Delta\mathbf{u}_s^T\}(\mathbf{I}_L \otimes \mathbf{C}(\mathbf{a})^T)\mathbf{W}^T\mathbf{H}^*(\mathbf{H}^H\mathbf{W}\mathbf{H})^{-T}$. It follows from (56) that

$$\begin{aligned} \mathbb{E}\{\Delta\mathbf{u}_s\Delta\mathbf{u}_s^T\} &\approx (\boldsymbol{\Sigma}_s^{-1}\mathbf{V}_s^T \otimes (\mathbf{I}_M - \mathbf{U}_s\mathbf{U}_s^H))\mathbb{E}\{\bar{\mathbf{n}}\bar{\mathbf{n}}^T\} \\ &\times (\mathbf{V}_s\boldsymbol{\Sigma}_s^{-1} \otimes (\mathbf{I}_M - \mathbf{U}_s^*\mathbf{U}_s^T)). \end{aligned} \quad (90)$$

Since $\mathbb{E}\{\bar{\mathbf{n}}\bar{\mathbf{n}}^T\} = \mathbf{0}_{MN \times MN}$, (90) becomes a zero matrix which gives rise to

$$\mathbb{E}\{\Delta\gamma_l^2\} \approx 0. \quad (91)$$

As a result, by combining (79), (88) and (91), we obtain (62), which completes the proof.

REFERENCES

- [1] H. L. Van Trees, *Optimum Array Processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.
- [2] 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.
- [3] 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.
- [4] K. Upadhy and S. A. Vorobyov, "An array processing approach to pilot decontamination for massive MIMO," in *Proc. IEEE 6th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process., Cancun, Mexico*, Dec. 2015, pp. 453–456.
- [5] K. Upadhy, S. A. Vorobyov, and R. W. Heath, "Low-overhead receiver-side channel tracking for mmWave MIMO," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Alberta, Canada*, Apr. 2018, pp. 3859–3863.

- [6] J. Zhang, D. Rakhimov, and M. Haardt, "Gridless channel estimation for hybrid mmWave MIMO systems via tensor-ESPRIT algorithms in DFT beamspace," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 816–831, Apr. 2021.
- [7] K. Ardah, S. Gherekhloo, A. L. de Almeida, and M. Haardt, "TRICE: A channel estimation framework for RIS-aided millimeter-wave MIMO systems," *IEEE Signal Process. Lett.*, vol. 28, pp. 513–517, Feb. 2021.
- [8] 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.
- [9] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [10] 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. Theo.*, vol. 58, no. 1, pp. 278–301, Jan. 2012.
- [11] P. Vallet, X. Mestre, and P. Loubaton, "Performance analysis of an improved MUSIC DOA estimator," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6407–6422, Dec. 2015.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] M. Haardt and J. A. Nossék, "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.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] M. Shaghghi 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.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] 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.
- [32] D. Madurasinghe, "A new DOA estimator in nonuniform noise," *IEEE Signal Process. Lett.*, vol. 12, no. 4, pp. 337–339, Apr. 2005.
- [33] 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.
- [34] 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 Sensors J.*, vol. 16, no. 24, pp. 8982–8989, Dec. 2016.
- [35] 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.
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.
- [40] 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.
- [41] 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.
- [42] 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.
- [43] 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.
- [44] 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.
- [45] 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.
- [46] 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.
- [47] 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 Trans. Signal Process.*, vol. 48, no. 10, pp. 2743–2749, Oct. 2000.
- [48] A. B. Gershman, P. Stoica, M. Pesavento, and E. Larsson, "The stochastic Cramér-Rao bound for direction estimation in unknown noise fields," *Proc. Inst. Elect. Eng. Radar, Sonar, Navigat.*, vol. 149, no. 1, pp. 2–8, Feb. 2002.
- [49] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templates for the solution of algebraic eigenvalue problems: A practical guide*. Society for Industrial and Applied Mathematics, 2000.
- [50] A. Hjørungnes, *Complex-valued matrix derivatives: With applications in signal processing and communications*. Cambridge University Press, 2011.
- [51] 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.
- [52] 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.
- [53] T. Amemiya, "Generalized least squares theory," *Advanced Econometrics*, Cambridge, MA, USA: Harvard Univ. Press, 1985.
- [54] T. Kariya and H. Kurata, *Generalized Least Squares*. John Wiley & Sons, 2004.
- [55] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.
- [56] 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.
- [57] 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.
- [58] 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.
- [59] 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.
- [60] 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.
- [61] S. M. Kay, *Fundamentals of statistical signal processing*. Prentice Hall PTR, 1993, Chapter 8.
- [62] K.-C. Huang 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.
- [63] A. B. Gershman and P. Stoica, "On unitary and forward-backward MODE," *Digit. Signal Process.*, vol. 9, no. 2, pp. 67–75, Apr. 1999.
- [64] 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.

- [65] K. Petersen and M. Pedersen, "The matrix cookbook, version 20121115," *Technical Univ. Denmark, Kongens Lyngby, Denmark, Tech. Rep.*, vol. 3274, 2012.
- [66] M. Bilodeau and D. Brenner, *Theory of multivariate statistics*. Springer Science & Business Media, 2008.
- [67] 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.
- [68] H. C. So, Y. T. Chan, K. Ho, and Y. Chen, "Simple formulae for bias and mean square error computation [DSP tips and tricks]," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 162–165, Jul. 2013.

Publication 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.

© 10 pages, submitted to a journal, June 2023
Reprinted with permission.

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

Majdoddin Esfandiari, *Student Member, IEEE*, Sergiy A. Vorobyov, *Fellow, IEEE*, and Robert W. Heath Jr., *Fellow, IEEE*

Abstract—Owing to their low power consumption and cost, the use of one-bit analog-to-digital converter (ADC) has been considered as a viable alternative to high resolution counterparts in realizing and commercializing massive multiple-input multiple-output (MIMO) systems. However, the issue of discarding the amplitude information by one-bit quantizers has to be compensated/reversed in order to make their usage plausible. As a result, carefully tailored methods need to be developed for the problems such as one-bit channel estimation and data detection as the conventional ones cannot be used. To address these issues, the problems of one-bit channel estimation and data detection for MIMO orthogonal frequency division multiplexing (OFDM) system that operates over uncorrelated frequency selective channels is investigated in this work. We first develop channel estimators that exploit Gaussian discriminant analysis (GDA) classifier and two approximated versions of it as the so-called *weak classifiers* in an adaptive boosting (AdaBoost) approach. Particularly, the combination of the approximated GDA classifiers with AdaBoost offers the benefit of scalability with the linear order of computations, which is critical in massive MIMO-OFDM systems. We then take advantage of the same idea for proposing the data detectors. Numerical results validate the efficiency of the proposed channel estimators and data detectors compared to other existing methods, where the proposed methods provide comparable performance to that of the state-of-the-art methods but require dramatically lower computational complexities and run times.

Index Terms—One-bit ADC, channel estimation, data detection, massive MIMO-OFDM, frequency selective channel, AdaBoost

I. INTRODUCTION

Utilization of a large number of antennas at the base station (BS) in communication systems have been explored for the purpose of enhancing data rates and network capacity [1], [2]. Massive multiple-input multiple-output (MIMO) communication systems have been demonstrated to offer remarkable advantages, but the hardware cost and high power consumption are two main difficulties (among others), hindering their commercial usage. To address these issues, first analog-to-digital converters (ADCs) have been recognized as one of the parts of the receivers that have high power consumption and expensive price [3], [4]. Then, employing low-resolution (ADCs) has been suggested as a

viable alternative instead of using high-resolution counterparts [5], [6]. However, the use of low-resolution ADCs in multi-user MIMO-OFDM systems poses several challenges in the receiver design. For instance, the non-linearities caused by few bit quantizers may prohibit us from exploiting conventional receivers like zero-forcing (ZF) and minimum mean square error (MMSE) detectors [7]. The reason is that the conventional procedure of isolating narrowband OFDM subcarriers using a discrete Fourier transform (DFT) at the receiver is not valid when low-resolution ADCs are used. Instead, different receiver architectures need to be employed/designed to process the baseband time-domain signals for the tasks such as channel estimation, and/or data detection.

Channel estimation and/or data detection in massive MIMO systems with one-bit ADCs have been explored in several papers, considering the cases of single-carrier (SC) and multi-carriers (MC) signalling. The authors of [8] have revised the non-convex optimization problem of the maximum likelihood (ML) channel estimator and proposed a sub-optimal channel estimator referred to as near-ML (nML). The same methodology has been used to develop the nML-based data detector as well. Convex optimization approaches have been exploited in [9] for estimating MC-OFDM, whereas a data detector has been developed based on a soft-output MMSE algorithm. In [10], the Bussgang decomposition [11] has been employed to develop Bussgang-based minimum mean-squared error (BMMSE) channel estimators and data detectors for both SC and MC-OFDM systems. Analogous to [10], the authors of [12] took advantage of the Bussgang decomposition to estimate the optimal nonzero thresholds in the problem of one-bit quantizer design. Multiple works such as [13]- [16] have considered the problem of joint channel estimation and data detection, where the known pilot sequence is augmented with a portion of detected data to build a longer virtual pilot sequence and subsequently utilize it to refine the channel estimate. For instance, the authors of [13] have developed a bilinear generalized approximate message passing (BiGAMP) method, while the authors of [15] have proposed a variational Bayesian (VB) algorithm to do so.

One interesting idea presented by different researchers is to treat one-bit channel estimation and data detection as binary classification problems, where the output of one-bit ADCs can be viewed as class labels. Moreover, a proper transformation of the known pilots or channel state information (CSI) plays the role of the classification features, while the unknown channel/data vectors act as the corresponding separating hy-

M. Esfandiari and S. A. Vorobyov are with Department of Information and Communications Engineering, Aalto University, PO Box 15400, 00076 Aalto, Finland. (Emails: majdoddin.esfandiari@aalto.fi; sergiy.vorobyov@aalto.fi).

R. W. Heath Jr. is with Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, USA. (Email: rwheathjr@ncsu.edu).

perplanes. For instance, the binary soft-margin support vector machine (SVM) has been considered by the authors of [16] and [17] as a powerful method to estimate the one-bit channel estimation and data detection in SC and MC-OFDM scenarios. Although the soft-margin SVM-based estimators have good properties, their performance relies on careful hyperparameter selection. Deep neural networks (DNN) have been also used for one-bit channel estimation in several works such as [18]-[20]. The main disadvantage of such estimators is that not only a sufficiently large data set is required for the training process, but the offline training procedure needs to be executed carefully. In [21]- [23], several blind/semi-blind learning-based data detectors have been presented for massive MIMO systems that employ one-bit ADCs.

Angular domain channel estimators have been reported in [24]- [31]. In [24] and [25], compressive sensing (CS) techniques have been adopted to recover sparse millimeter wave (mmWave) channels quantized by few-bit ADCs. The authors of [29] have considered the combination of harmonic retrieval methods with a modified expected-maximization GAMP (EM-GAMP) to devise an angular domain one-bit mmWave channel estimation approach called gridless GAMP (GL-GAMP). For such channels, a sparsity enforcing with Toeplitz matrix reconstruction (SE-TMR) method was also presented in [30] recently. Moreover, the authors of [31] have used the Toeplitz matrix reconstruction notion from [30] together with ℓ_1 regularized logistic regression classification method [32] to come up with a novel angular domain channel estimator called ℓ_1 regularized logistic regression with Toeplitz matrix reconstruction (L1-RLR-TMR) for one-bit mmWave systems. They also have employed the alternating direction method of multipliers (ADMM) [33] for solving the optimization problem of L1-RLR-TMR in an efficient manner.

Despite the significance of scalability and efficiency in one-bit massive MIMO-OFDM systems, the existing channel estimators and data detectors may not fulfill the requirement of having low computational complexity in challenging scenarios with large number of unknowns. In other words, there is a gap between the desirable computational complexity and that of the existing methods to the best of our knowledge. Therefore, the objective of this work is to fill the aforementioned gap by proposing one-bit channel estimators and data detectors that have linear order of computations with respect to the system parameters including the number of antennas at BS, the number of users, and the number of OFDM sub-carriers.

In this paper, we develop channel estimation and data detection algorithms for MIMO-OFDM systems that exploit one-bit ADCs at the BS. The channel considered here is a frequency selective channel. Inspired by outstanding properties that classification/learning-based methods have shown in solving one-bit channel estimation and data detection, we design Gaussian discriminant analysis (GDA)-based classification method [34] (known also as linear discriminant analysis (LDA)) and its approximations as so-called weak classifiers, employed in each iteration of an adaptive boosting (AdaBoost)-based scheme [32], [35]. The low computational complexity required for implementation of both GDA-based classifiers and AdaBoost make the proposed algorithms efficient, and easily scalable.

In addition, flexibility in selecting the number of AdaBoost iterations enables us to gain competitive accuracy with low computational complexity.

The main contributions of our work are the following:

- An AdaBoost-based channel estimation approach for one-bit MIMO-OFDM system that operates over uncorrelated frequency selective fading channels is proposed. In each iteration of the AdaBoost-based approach, the GDA classification method along with two efficient approximations are considered as the weak classifiers. These approximate classifiers are derived by manipulating the GDA estimator. The combination of AdaBoost and GDA (and especially its approximations) enables us to estimate the channel in a remarkably efficient and yet precise manner. Specifically, using the approximations of GDA as weak classifiers at the heart of our AdaBoost approach results in having the linear order of computational complexity with respect to the problem dimension. This makes the proposed AdaBoost-based approach a versatile and also powerful tool that can be used in one-bit MIMO-OFDM systems with large number of channel entries. Numerical results validate the efficiency of the proposed AdaBoost-based channel estimator compared to other existing methods. Particularly, the AdaBoost-based channel estimator possesses similar normalized MSE (NMSE) in channel estimation as the SVM-based method of [16], whereas the computational complexity required to implement our method is substantially less than that of the SVM-based method in scenarios with large dimensions.
- We then tailor the main idea of the proposed AdaBoost-based channel estimator to fit the one-bit MIMO-OFDM data detection problem. Analogous to the proposed one-bit channel estimator, we design the data detector as an AdaBoost-based approach with considering GDA and its approximations as the weak classifiers in each iteration. The proposed one-bit data detector has desirable properties like scalability (with linear order of computations) and providing accurate data estimates. These properties are very useful in feasibility of designing one-bit MIMO-OFDM systems with high bandwidth and large number of sub-carriers. Numerical results demonstrate the strength of the proposed AdaBoost-based data detector compared to other existing methods.

The rest of the paper is organized as follows. The considered system model is presented in Section II. A brief review of GDA and AdaBoost are also presented in Section II. The proposed AdaBoost-based one-bit channel estimator and data detector are designed in Section III. Simulation results and the conclusion are presented in Section IV and Section V, respectively.

Notation: Upper-case and lower-case bold-face letters denote matrices and vectors, respectively, while scalars are denoted by lower-case letters. The mathematical expectation, transpose, and inverse of a square matrix are denoted by $\mathbb{E}\{\cdot\}$, $\{\cdot\}^T$, and $(\cdot)^{-1}$, respectively, while $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the Euclidean norm of a vector and the Frobenius norm of a matrix. The Hadamard product is denoted by \odot .

The $n \times n$ identity matrix is denoted by \mathbf{I}_n . The operator $\text{diag}\{\boldsymbol{\pi}\}$ generates a diagonal matrix by plugging the entries of the vector $\boldsymbol{\pi}$ into its main diagonal. The operators $\Re\{\cdot\}$ and $\Im\{\cdot\}$ return respectively the real and imaginary parts of the bracketed argument. The function $\mathbf{1}\{\cdot\}$ is the indicator function that is equal to 1 if its argument is true and 0 otherwise.

II. SYSTEM MODEL AND PRELIMINARIES

A. One-Bit Massive MIMO-OFDM System Model

We assume a massive MIMO system comprising of K users, each equipped with a single-antenna, and an M -antenna base station (BS) where users deploy high-resolution ADCs. Each antenna of the BS converts the real and imaginary components of the received signal from the users separately through a pair of one-bit ADCs. We specifically examine an uplink multiuser OFDM system with N_c sub-carriers that operates over a frequency selective channel. The OFDM symbol in the frequency domain from the k^{th} user is represented by $\mathbf{x}_k^{\text{FD}} \in \mathbb{C}^{N_c \times 1}$. To avoid confusion, we use the notations "TD" and "FD" to distinguish between time and frequency domains, respectively. We add a cyclic prefix (CP) of length N_{cp} and assume that the number of channel taps L_{tap} satisfies the condition $L_{\text{tap}} - 1 \leq N_{\text{cp}} \leq N_c$. It is assumed that L_{tap} is known. Upon removing the CP, the one-bit quantized received signal at the i^{th} antenna of the BS in the time domain can be expressed as follows:

$$\mathbf{y}_i^{\text{TD}} = \mathcal{Q} \left(\sum_{k=1}^K \mathbf{G}_{i,k}^{\text{TD}} \mathbf{F}^H \mathbf{x}_k^{\text{FD}} + \mathbf{n}_i^{\text{TD}} \right) \quad (1)$$

where $\mathbf{F} \in \mathbb{C}^{N_c \times N_c}$ denotes the normalized DFT matrix, and $\mathbf{G}_{i,k}^{\text{TD}}$ is a circulant matrix whose first column is defined by $\mathbf{g}_{i,k}^{\text{TD}} = [(\mathbf{h}_{i,k}^{\text{TD}})^T, 0, \dots, 0]^T$. Here, $\mathbf{h}_{i,k}^{\text{TD}} \in \mathbb{C}^{L_{\text{tap}} \times 1}$ is a vector that contains the L_{tap} channel taps associated with the k^{th} user. The entries of $\mathbf{h}_{i,k}^{\text{TD}}$ are considered to be independent and identically distributed (i.i.d.), generated from the distribution $\mathcal{CN}(0, \frac{1}{L_{\text{tap}}})$. Moreover, $\mathbf{n}_i^{\text{TD}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_c})$ represents additive Gaussian noise at the i^{th} antenna at the BS, whereas the notation $\mathcal{Q}(\cdot) \triangleq \text{sign}(\Re\{\cdot\}) + j\text{sign}(\Im\{\cdot\})$ represents the element-wise one-bit quantizer. The output of the operator $\text{sign}(\cdot)$ is +1 when the argument is a non-negative number, otherwise, the output is -1.

We stress here that because of the nonlinear distortion imposed by one-bit quantizers, different OFDM sub-carriers are not separable by the FFT operation as opposed to the conventional MIMO-OFDM systems. As a result, we are obliged to develop the proposed channel estimators and data detectors based on the wideband time domain representation instead of exploiting the narrowband frequency domain signals associated with each sub-carrier.

B. Binary Classification via GDA

GDA (also known as LDA) is a classification approach that models the training examples associated with each class as samples of a normal distribution. Consider a training set that contains m training examples with n features and two classes

denoted by $\{\mathbf{x}^{(j)}\}_{j=1, \dots, m}$ and $y^{(j)} \in \{1, -1\}_{j=1, \dots, m}$, respectively. GDA assumes that the corresponding training examples $\mathbf{x}^{(j)}$ for each class of $y^{(j)}$ are normally distributed with different means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_{-1}$, respectively, and the same covariance matrix $\boldsymbol{\Sigma}$. Therefore, depending on $y^{(j)}$, the conditional probability density function (PDF) of $\mathbf{x}^{(j)}$ can be given as one of the following equations:

$$p(\mathbf{x}^{(j)} | y^{(j)} = -1) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{x}^{(j)} - \boldsymbol{\mu}_{-1})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}^{(j)} - \boldsymbol{\mu}_{-1}) \right) \quad (2)$$

$$p(\mathbf{x}^{(j)} | y^{(j)} = 1) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{x}^{(j)} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}^{(j)} - \boldsymbol{\mu}_1) \right) \quad (3)$$

To implement binary GDA, we need to estimate $\boldsymbol{\mu}_{-1}$, $\boldsymbol{\mu}_1$, and $\boldsymbol{\Sigma}$ from the training data. The means and the covariance matrix can be estimated as follows [34]

$$\hat{\boldsymbol{\mu}}_{-1} = \frac{\sum_{j=1}^m \mathbf{1}\{y^{(j)} = -1\} \mathbf{x}^{(j)}}{\sum_{j=1}^m \mathbf{1}\{y^{(j)} = -1\}} \quad (4)$$

$$\hat{\boldsymbol{\mu}}_1 = \frac{\sum_{j=1}^m \mathbf{1}\{y^{(j)} = 1\} \mathbf{x}^{(j)}}{\sum_{j=1}^m \mathbf{1}\{y^{(j)} = 1\}} \quad (5)$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{m} \sum_{j=1}^m (\mathbf{x}^{(j)} - \hat{\boldsymbol{\mu}}_{y^{(j)}}) (\mathbf{x}^{(j)} - \hat{\boldsymbol{\mu}}_{y^{(j)}})^T. \quad (6)$$

The decision boundary is then given as

$$\mathbf{h}_{\text{GDA}} = \hat{\boldsymbol{\Sigma}}^{-1} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_{-1}). \quad (7)$$

C. AdaBoost

The objective of AdaBoost is to iteratively train a set of *weak classifiers* on the same data set to create a *strong classifier*. A weak classifier is identified as a classifier whose classification performance is only marginally better than random guessing. A new weak classifier is trained on a weighted version of the training data set, where the weights associated with the misclassified examples in the previous iteration are increased. Given a training set with m examples, AdaBoost learns a weak classifier in the t^{th} iteration which is denoted by $h_t(\mathbf{x})$. The AdaBoost algorithm is outlined in Algorithm 1. Here, $w_j^{(t)}$ is the weight of the j^{th} example at the t^{th} iteration, $\epsilon^{(t)}$ is the weighted error of the t^{th} weak classifier, and $\alpha^{(t)}$ is the weight of the t^{th} weak classifier. Moreover, $Z^{(t+1)}$ is a normalization constant that ensures that the weights sum up to 1. Despite there exists various ways to define the update rule for $w_j^{(t+1)}$, Algorithm 1 employs the exponential function to do so.

In our derivations, we use GDA and its approximate versions as weak classifiers, although there are many linear binary

classifiers available in the literature that can be considered as weak classifiers. The main reason for the aforementioned choice is that these classifiers can be implemented with low computational complexities, particularly when the dimension of the unknown variables scales up.

Algorithm 1 AdaBoost Algorithm

Input: Training set \mathcal{S} , number of weak classifiers T

Output: Final classifier H_{Ada}

Initialize weights $w_j^{(1)} = 1/m$ for $j = 1, 2, \dots, m$

for $t = 1$ to T **do**

Train weak classifier $h^{(t)}(x)$ on the weighted training set (\mathcal{S}, w) .

Compute error as $\epsilon^{(t)} = \sum_{j=1}^m w_j^{(t)} \mathbf{1}(h^{(t)}(\mathbf{x}^{(j)}) \neq y^{(j)})$.

Compute $\alpha^{(t)} = \frac{1}{4} \ln \left(\frac{1-\epsilon^{(t)}}{\epsilon^{(t)}} \right)$.

Update $w_j^{(t+1)} = w_j^{(t)} \exp(\alpha^{(t)} \mathbf{1}(h^{(t)}(\mathbf{x}^{(j)}) \neq y^{(j)}))$, $\forall j$.

Compute $Z^{(t+1)} = \sum_{j=1}^m w_j^{(t+1)}$ and normalize weights as $w_j^{(t+1)} = \frac{w_j^{(t+1)}}{Z^{(t+1)}}$, $\forall j$.

end for

Output $H_{\text{Ada}}(\mathbf{x}) = \sum_{t=1}^T \alpha^{(t)} h^{(t)}(\mathbf{x})$

III. PROPOSED CLASSIFICATION-BASED WIDEBAND CHANNEL ESTIMATION AND DATA DETECTION WITH ONE-BIT ADCS

A. Proposed Classification-Based Channel Estimation

For estimating the frequency selective channels explained in Section II that is utilized in the OFDM system, the frequency domain pilot vector $\mathbf{x}_k^{\text{FD}} \in \mathbb{C}^{N_c \times 1}$ is first transformed into the time domain using the inverse fast Fourier transform (IFFT) operation. The resultant time domain vector is then transmitted by the k^{th} user. The one-bit quantized received signal at the i^{th} antenna of the BS in (1) can be reorganized as

$$\begin{aligned} \mathbf{y}_i^{\text{TD}} &= \mathcal{Q} \left(\sum_{k=1}^K \Phi_k^{\text{TD}} \mathbf{g}_{i,k}^{\text{TD}} + \mathbf{n}_i^{\text{TD}} \right) \\ &= \mathcal{Q} \left(\sum_{k=1}^K \Phi_{k,L_{\text{tap}}}^{\text{TD}} \mathbf{h}_{i,k}^{\text{TD}} + \mathbf{n}_i^{\text{TD}} \right) \\ &= \mathcal{Q} \left(\Phi_{L_{\text{tap}}}^{\text{TD}} \mathbf{h}_i^{\text{TD}} + \mathbf{n}_i^{\text{TD}} \right) \end{aligned} \quad (8)$$

where $\Phi_k^{\text{TD}} \in \mathbb{C}^{N_c \times N_c}$ is a circulant matrix whose first column is $\phi_k^{\text{TD}} \triangleq \mathbf{F}^H \mathbf{x}_k^{\text{FD}}$, $\Phi_{k,L_{\text{tap}}}^{\text{TD}} \in \mathbb{C}^{N_c \times L_{\text{tap}}}$ denotes a matrix which contains only the first L_{tap} columns of Φ_k^{TD} , $\Phi_{L_{\text{tap}}}^{\text{TD}} \in \mathbb{C}^{N_c \times KL_{\text{tap}}}$ and $\mathbf{h}_i^{\text{TD}} \in \mathbb{C}^{KL_{\text{tap}} \times 1}$ respectively concatenate $\Phi_{k,L_{\text{tap}}}^{\text{TD}}$ and $\mathbf{h}_{i,k}^{\text{TD}}$ for $k = 1, \dots, K$ as $\Phi_{L_{\text{tap}}}^{\text{TD}} \triangleq [\Phi_{1,L_{\text{tap}}}^{\text{TD}}, \Phi_{2,L_{\text{tap}}}^{\text{TD}}, \dots, \Phi_{K,L_{\text{tap}}}^{\text{TD}}]$ and $\mathbf{h}_i^{\text{TD}} \triangleq [(\mathbf{h}_{i,1}^{\text{TD}})^T, (\mathbf{h}_{i,2}^{\text{TD}})^T, \dots, (\mathbf{h}_{i,K}^{\text{TD}})^T]^T$.

To simplify our derivations, we use the notation ‘‘R’’ as subscript when scalars, vectors, or matrices are composed of real numbers. Therefore, we transform (8) into the real domain as

$$\mathbf{y}_{i,R}^{\text{TD}} = \text{sign} \left(\Phi_{R,L_{\text{tap}}}^{\text{TD}} \mathbf{h}_{i,R}^{\text{TD}} + \mathbf{n}_{i,R}^{\text{TD}} \right) \quad (9)$$

where

$$\begin{aligned} \mathbf{y}_{i,R}^{\text{TD}} &\triangleq [\Re\{\mathbf{y}_i^{\text{TD}}\}^T, \Im\{\mathbf{y}_i^{\text{TD}}\}^T]^T = [y_{i,R,1}^{\text{TD}}, \dots, y_{i,R,2N_c}^{\text{TD}}]^T \\ &\in \{\pm 1\}^{2N_c \times 1} \end{aligned} \quad (10)$$

$$\begin{aligned} \Phi_R &\triangleq \begin{bmatrix} \Re\{\Phi_{L_{\text{tap}}}^{\text{TD}}\} & -\Im\{\Phi_{L_{\text{tap}}}^{\text{TD}}\} \\ \Im\{\Phi_{L_{\text{tap}}}^{\text{TD}}\} & \Re\{\Phi_{L_{\text{tap}}}^{\text{TD}}\} \end{bmatrix} \\ &= [\phi_{R,1}^{\text{TD}}, \phi_{R,2}^{\text{TD}}, \dots, \phi_{R,2N_c}^{\text{TD}}]^T \in \mathbb{R}^{2N_c \times 2KL_{\text{tap}}} \end{aligned} \quad (11)$$

$$\mathbf{h}_{i,R}^{\text{TD}} \triangleq [\Re\{\mathbf{h}_i^{\text{TD}}\}^T, \Im\{\mathbf{h}_i^{\text{TD}}\}^T]^T \in \mathbb{R}^{2KL_{\text{tap}} \times 1} \quad (12)$$

$$\mathbf{n}_{i,R}^{\text{TD}} \triangleq [\Re\{\mathbf{n}_i^{\text{TD}}\}^T, \Im\{\mathbf{n}_i^{\text{TD}}\}^T]^T \in \mathbb{R}^{2N_c \times 1}. \quad (13)$$

Note that $(\phi_{R,j}^{\text{TD}})^T$ with $j \in \{1, 2, \dots, 2N_c\}$ is the j^{th} row of Φ_R here. Additionally, as suggested by (12), estimating $\{\mathbf{h}_{i,R}^{\text{TD}}\}_{i=1,2,\dots,M}$ is equivalent to estimating $\{\mathbf{h}_{i,R}^{\text{TD}}\}_{i=1,2,\dots,M}$.

We emphasize that binary classification methods can be employed for estimating $\mathbf{h}_{i,R}^{\text{TD}}$ in (9). Here, $\phi_{R,j}^{\text{TD}}$ and $y_{i,R,j}^{\text{TD}}$ with $j \in \{1, 2, \dots, 2N_c\}$ serve as the training examples and the class labels, respectively. In other words, (9)-(12) can be viewed as a binary classification problem with the training set $\mathcal{S}_i = \{\mathbf{x}^{(j)} = \phi_{R,j}^{\text{TD}}, y^{(j)} = y_{i,R,j}^{\text{TD}}\}_{j=1,2,\dots,2N_c}$ and the decision boundary $\mathbf{h}_{i,R}^{\text{TD}}$ based on the definitions provided in the prequel. Hence, we can exploit the GDA classification method as the weak classifier in each iteration of an AdaBoost-based approach for estimating $\mathbf{h}_{i,R}^{\text{TD}}$. The computation of the means and covariance matrix (4)-(6) then should be revised in the t^{th} iteration of the proposed AdaBoost-based approach as

$$\hat{\boldsymbol{\mu}}_{-1}^{(t)} = \sum_{j=1}^{2N_c} \mathbf{1}\{y_{i,R,j}^{\text{TD}} = -1\} w_j^{(t)} \phi_{R,j}^{\text{TD}} \quad (14)$$

$$\hat{\boldsymbol{\mu}}_{1}^{(t)} = \sum_{j=1}^{2N_c} \mathbf{1}\{y_{i,R,j}^{\text{TD}} = 1\} w_j^{(t)} \phi_{R,j}^{\text{TD}} \quad (15)$$

$$\hat{\boldsymbol{\Sigma}}^{(t)} = \sum_{j=1}^{2N_c} w_j^{(t)} (\phi_{R,j}^{\text{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,R,j}^{\text{TD}}}^{(t)}) (\phi_{R,j}^{\text{TD}} - \hat{\boldsymbol{\mu}}_{y_{i,R,j}^{\text{TD}}}^{(t)})^T \quad (16)$$

$$\hat{\mathbf{h}}_{i,R}^{\text{TD},(t)} = (\hat{\boldsymbol{\Sigma}}^{(t)})^{-1} (\hat{\boldsymbol{\mu}}_{1}^{(t)} - \hat{\boldsymbol{\mu}}_{-1}^{(t)}) \quad (17)$$

where $w_j^{(t)}$ represents the weight of the j^{th} training example at the t^{th} iteration.

To implement (17), the inverse of the matrix $\hat{\boldsymbol{\Sigma}}^{(t)}$ should be calculated, which requires the computational complexity of $\mathcal{O}((2KL_{\text{tap}})^3)$. The cubic computational complexity can considerably restrict the time efficiency of implementing (17), especially when the multiplication of K and L_{tap} grows larger. At the same time, as a weak classifier is required to be slightly better than random guesses, the accurate knowledge of the inverse of $\hat{\boldsymbol{\Sigma}}^{(t)}$ is not needed. Thus, it is reasonable to consider approximating (17) to avoid the cubic computational complexity of calculating $(\hat{\boldsymbol{\Sigma}}^{(t)})^{-1}$. Towards this end, two approximations of (17) are introduced in the following.

Approximation 1: As the first approximation, we propose to modify (16) as

$$\hat{\boldsymbol{\Sigma}}_1^{(t)} \triangleq \text{diag} \left\{ \hat{\boldsymbol{\sigma}}_1^{(t)} \right\} \quad (18)$$

where

$$\hat{\sigma}_1^{(t)} = \sum_{j=1}^{2N_c} w_j^{(t)} \left((\phi_{R,j}^{\text{TD}} - \hat{\mu}_{y_{i,R,j}^{\text{TD}}}^{(t)}) \odot (\phi_{R,j}^{\text{TD}} - \hat{\mu}_{y_{i,R,j}^{\text{TD}}}^{(t)}) \right) \quad (19)$$

The essence of this approximation is to set all off-diagonal elements of $\hat{\Sigma}^{(t)}$ in (16) to zero and preserve only its diagonal elements. In other words, only the diagonal elements of the original matrix $\hat{\Sigma}^{(t)}$ in (16) need to be computed as the vector $\hat{\sigma}_1^{(t)}$ in (19), and $\hat{\Sigma}_1^{(t)}$ is defined using $\hat{\sigma}_1^{(t)}$ as (18). Then, (17) is modified as

$$\hat{\mathbf{h}}_{i,R,\text{app1}}^{\text{TD},(t)} \triangleq \left(\hat{\Sigma}_1^{(t)} \right)^{-1} \left(\hat{\mu}_1^{(t)} - \hat{\mu}_{-1}^{(t)} \right) \quad (20)$$

Note that the use of $\hat{\Sigma}_1^{(t)}$ instead of the original $\hat{\Sigma}^{(t)}$ considerably reduces the computational complexity of computing $\hat{\mathbf{h}}_{i,R}^{\text{TD},(t)}$.

Approximation 2: We propose to set $\hat{\Sigma}^{(t)} = \mathbf{I}_{2KL_{\text{tap}}}$ in (17). Then, the modified estimate of $\hat{\mathbf{h}}_{i,R}^{\text{TD},(t)}$ is expressed as

$$\hat{\mathbf{h}}_{i,R,\text{app2}}^{\text{TD},(t)} \triangleq \hat{\mu}_1^{(t)} - \hat{\mu}_{-1}^{(t)} \quad (21)$$

where the weak classifier of (17) is approximated as the distance between the mean vectors of the two classes in (21). We stress here that the later requires substantially less computations compared to that of the former.

The steps of the proposed methods are outlined in Algorithm 2. It should be noted that Algorithm 1 presents the generic procedure of the AdaBoost approach for using weak binary classifiers/learners $h^{(t)}(x)$ to build a strong binary classifier/learner $H_{\text{Ada}}(x)$, whereas we exploit the core idea of AdaBoost to use the weak channel estimates $\hat{\mathbf{h}}_i^{(t)}$ to build a strong channel estimate $\hat{\mathbf{h}}_{i,R}^{\text{TD}}$ in Algorithm 2. We emphasize here the difference of $h^{(t)}(x)$ and $H_{\text{Ada}}(x)$ with $\hat{\mathbf{h}}_i^{(t)}$ and $\hat{\mathbf{h}}_{i,R}^{\text{TD}}$, that is, the former represents binary classifier while the later denotes the separating hyperplane in a binary classification problem.

Note that a normalization step is applied to the output of the AdaBoost-based methods outlined in Algorithm 2. The reason for this is that the estimates provided by these methods only specify the direction of $\hat{\mathbf{h}}_{i,R}^{\text{TD}}$, while the magnitude remains unknown since the one-bit ADCs preserve only the sign of the received signals. Therefore, $\beta \hat{\mathbf{h}}_{i,R}^{\text{TD}}$ for any $\beta > 0$ will yield the same $\mathbf{y}_{i,R}^{\text{TD}}$ as in (10). Here, since we assume that $2KL_{\text{tap}}$ elements of $\hat{\mathbf{h}}_{i,R}^{\text{TD}}$ are independent with variance $1/(2L_{\text{tap}})$, the last normalization step is added to ensure that the channel estimates have squared norm of K .

Remark 1: To ensure the clarity of presentation, we used a loop to estimate $\hat{\mathbf{h}}_{i,R}^{\text{TD}}$ for $i \in \{1, 2, \dots, M\}$ in Algorithm 2. However, it is important to note that these M channel vectors can be estimated in parallel, resulting in a reduction in the overall run time of the channel estimation procedure.

Remark 2: The key feature of AdaBoost that allows us to approximate (17) as (20) and (21) without sacrificing estimation performance is that it can incorporate weak classifiers that are only slightly better than random guessing and combine them to form a strong classifier. The approximation of (20) and (21) are justifiable because they are certainly better than

Algorithm 2 One-bit GDA-AdaBoost Algorithm for Channel Estimation

Input: $\mathcal{S}_i = \{\mathbf{x}^{(j)} = \phi_{R,j}^{\text{TD}}, y^{(j)} = y_{i,R,j}^{\text{TD}}\}_{j=1,2,\dots,2N_c}$ for $i \in \{1, 2, \dots, M\}$ whose elements are defined in (10) and (11), and number of weak classifiers T .

Output: $\hat{\mathbf{h}}_{i,R}^{\text{TD}}$ for $i \in \{1, 2, \dots, M\}$.

for $i = 1$ to M **do**

Initialize weights $w_j^{(1)} = \frac{1}{2N_c}$ for $j \in \{1, 2, \dots, 2N_c\}$.

for $t = 1$ to T **do**

Use the training set \mathcal{S}_i to compute $\hat{\mu}_{-1}^{(t)}, \hat{\mu}_1^{(t)}, \hat{\Sigma}^{(t)}$, and $\hat{\Sigma}_1^{(t)}$ via (14)-(16) and (18), respectively. Then, compute the t^{th} weak classifier as:

one-bit GDA-AdaBoost

$$\mathbf{h}_i^{(t)} = \left(\hat{\Sigma}^{(t)} \right)^{-1} \left(\hat{\mu}_1^{(t)} - \hat{\mu}_{-1}^{(t)} \right)$$

one-bit GDA-AdaBoost-1

$$\mathbf{h}_i^{(t)} = \left(\hat{\Sigma}_1^{(t)} \right)^{-1} \left(\hat{\mu}_1^{(t)} - \hat{\mu}_{-1}^{(t)} \right)$$

one-bit GDA-AdaBoost-2

$$\mathbf{h}_i^{(t)} = \hat{\mu}_1^{(t)} - \hat{\mu}_{-1}^{(t)}.$$

Compute error as

$$\epsilon^{(t)} = \sum_{j=1}^{2N_c} w_j^{(t)} \mathbf{1} \left((\phi_{R,j}^{\text{TD}})^T \mathbf{h}_i^{(t)} \neq y^{(j)} \right).$$

Compute $\alpha^{(t)} = \frac{1}{4} \ln \left(\frac{1-\epsilon^{(t)}}{\epsilon^{(t)}} \right)$.

Update $w_j^{(t+1)} = w_j^{(t)} \exp \left(\alpha^{(t)} \mathbf{1} \left((\phi_{R,j}^{\text{TD}})^T \mathbf{h}_i^{(t)} \neq y^{(j)} \right) \right)$, $\forall j$.

Compute $Z^{(t+1)} = \sum_{j=1}^{2N_c} w_j^{(t+1)}$ and normalize weights as $w_j^{(t+1)} = \frac{w_j^{(t+1)}}{Z^{(t+1)}}$, $\forall j$.

end for

Construct $\hat{\mathbf{h}}_{i,R}^{\text{TD}} = \sum_{t=1}^T \alpha^{(t)} \mathbf{h}_i^{(t)}$, and then normalize as

$$\hat{\mathbf{h}}_{i,R}^{\text{TD}} = \frac{\sqrt{K} \hat{\mathbf{h}}_{i,R}^{\text{TD}}}{\|\hat{\mathbf{h}}_{i,R}^{\text{TD}}\|_2}.$$

end for

random guessing, hence they can be treated as weak classifiers. In this regard, AdaBoost is a powerful approach to build a strong classifier out of weak classifiers with low computational complexity.

B. Proposed Classification-Based Data Detection

In this section, we propose AdaBoost-based methods for one-bit data detection in OFDM systems with frequency selective channels. To begin with, the one-bit quantized received signal at the i^{th} antenna of the BS in (1) can be rewritten as

$$\begin{aligned} \mathbf{y}_i^{\text{TD}} &= \mathcal{Q} \left(\sum_{k=1}^K \mathbf{G}_{i,k}^{\text{TD}} \mathbf{F}^H \mathbf{x}_k^{\text{FD}} + \mathbf{n}_i^{\text{TD}} \right) \\ &= \mathcal{Q} \left(\mathbf{G}_i^{\text{FD}} \mathbf{x}^{\text{FD}} + \mathbf{n}_i^{\text{TD}} \right) \end{aligned} \quad (22)$$

where $\mathbf{G}_i^{\text{FD}} \triangleq [\mathbf{G}_{i,1}^{\text{TD}} \mathbf{F}^H, \dots, \mathbf{G}_{i,K}^{\text{TD}} \mathbf{F}^H] \in \mathbb{C}^{N_c \times KN_c}$ and $\mathbf{x}^{\text{FD}} \triangleq [(\mathbf{x}_1^{\text{FD}})^T, (\mathbf{x}_2^{\text{FD}})^T, \dots, (\mathbf{x}_K^{\text{FD}})^T]^T \in \mathbb{C}^{KN_c \times 1}$. The former represents the pre-estimated/known CSI, while the later is the symbol vectors transmitted over N_c subcarriers by the K users. The objective here is to recover the vector \mathbf{x}^{FD} and then identify the symbols transmitted. Placing all $\{\mathbf{y}_i^{\text{TD}}\}_{i=1,2,\dots,M}$

in a vector as $\mathbf{y}^{\text{TD}} \triangleq [(\mathbf{y}_1^{\text{TD}})^T, (\mathbf{y}_2^{\text{TD}})^T, \dots, (\mathbf{y}_M^{\text{TD}})^T]^T \in \mathbb{C}^{MN_c \times 1}$, we obtain

$$\mathbf{y}^{\text{TD}} = \mathcal{Q}(\mathbf{G}^{\text{FD}} \mathbf{x}^{\text{FD}} + \mathbf{n}^{\text{TD}}) \quad (23)$$

where $\mathbf{G}^{\text{FD}} \triangleq [(\mathbf{G}_1^{\text{FD}})^T, (\mathbf{G}_2^{\text{FD}})^T, \dots, (\mathbf{G}_M^{\text{FD}})^T]^T \in \mathbb{C}^{MN_c \times KN_c}$. The real domain transformation of (23) is given as

$$\mathbf{y}_R^{\text{TD}} = \text{sign}(\mathbf{G}_R^{\text{FD}} \mathbf{x}_R^{\text{FD}} + \mathbf{n}_R^{\text{TD}}) \quad (24)$$

where

$$\mathbf{y}_R^{\text{TD}} \triangleq [\Re\{\mathbf{y}^{\text{TD}}\}^T, \Im\{\mathbf{y}^{\text{TD}}\}^T]^T = [y_{R,1}^{\text{TD}}, \dots, y_{R,2MN_c}^{\text{TD}}]^T \in \{\pm 1\}^{2MN_c \times 1} \quad (25)$$

$$\mathbf{G}_R^{\text{FD}} \triangleq \begin{bmatrix} \Re\{\mathbf{G}^{\text{FD}}\} & -\Im\{\mathbf{G}^{\text{FD}}\} \\ \Im\{\mathbf{G}^{\text{FD}}\} & \Re\{\mathbf{G}^{\text{FD}}\} \end{bmatrix} = [\mathbf{g}_{R,1}^{\text{FD}}, \mathbf{g}_{R,2}^{\text{FD}}, \dots, \mathbf{g}_{R,2MN_c}^{\text{FD}}]^T \in \mathbb{R}^{2MN_c \times 2KN_c} \quad (26)$$

$$\mathbf{x}_R^{\text{FD}} \triangleq [\Re\{\mathbf{x}^{\text{FD}}\}^T, \Im\{\mathbf{x}^{\text{FD}}\}^T]^T \in \mathbb{R}^{2KN_c \times 1} \quad (27)$$

$$\mathbf{n}_R^{\text{TD}} \triangleq [\Re\{\mathbf{n}^{\text{TD}}\}^T, \Im\{\mathbf{n}^{\text{TD}}\}^T]^T \in \mathbb{R}^{2MN_c \times 1}. \quad (28)$$

Here $\{(\mathbf{g}_{R,j}^{\text{FD}})^T\}_{j=1, \dots, 2MN_c}$ is the j^{th} row of \mathbf{G}_R^{FD} .

Analogous to the problem of estimating $\mathbf{h}_{i,R}^{\text{TD}}$ in (9), the problem of estimating \mathbf{x}_R^{FD} in (24) can be treated as a binary classification problem where \mathbf{x}_R^{FD} serves as the separating hyperplane between two classes. Therefore, we can constitute the binary classification training set as $\mathcal{S}_d = \{\mathbf{x}^{(j)} = \mathbf{g}_{R,j}^{\text{FD}}, y^{(j)} = y_{R,j}^{\text{TD}}\}_{j=1, 2, \dots, 2MN_c}$ based on (24)-(26) with the aim of estimating \mathbf{x}_R^{FD} as the corresponding separating hyperplane. Thus, the GDA classification method along with two approximations derived in Subsection III-A can be used as weak classifiers in each iteration of an AdaBoost-based approach for recovering \mathbf{x}_R^{FD} . In this regard, the counterparts of (17), (20), and (21) with respect to \mathbf{x}_R^{FD} are respectively expressed as

$$\hat{\mathbf{x}}_R^{\text{FD},(t)} = \left(\hat{\Sigma}_d^{(t)}\right)^{-1} \left(\hat{\boldsymbol{\mu}}_{d,1}^{(t)} - \hat{\boldsymbol{\mu}}_{d,-1}^{(t)}\right) \quad (29)$$

$$\hat{\mathbf{x}}_{R,\text{app1}}^{\text{FD},(t)} = \left(\hat{\Sigma}_{d,1}^{(t)}\right)^{-1} \left(\hat{\boldsymbol{\mu}}_{d,1}^{(t)} - \hat{\boldsymbol{\mu}}_{d,-1}^{(t)}\right) \quad (30)$$

$$\hat{\mathbf{x}}_{R,\text{app2}}^{\text{FD},(t)} = \hat{\boldsymbol{\mu}}_{d,1}^{(t)} - \hat{\boldsymbol{\mu}}_{d,-1}^{(t)} \quad (31)$$

where

$$\hat{\boldsymbol{\mu}}_{d,-1}^{(t)} = \sum_{j=1}^{2MN_c} \mathbf{1}\{y_{R,j}^{\text{TD}} = -1\} w_j^{(t)} \mathbf{g}_{R,j}^{\text{FD}} \quad (32)$$

$$\hat{\boldsymbol{\mu}}_{d,1}^{(t)} = \sum_{j=1}^{2MN_c} \mathbf{1}\{y_{R,j}^{\text{TD}} = 1\} w_j^{(t)} \mathbf{g}_{R,j}^{\text{FD}} \quad (33)$$

$$\hat{\Sigma}_d^{(t)} = \sum_{j=1}^{2MN_c} w_j^{(t)} (\mathbf{g}_{R,j}^{\text{FD}} - \hat{\boldsymbol{\mu}}_{d,y_{R,j}^{\text{TD}}}^{(t)}) (\mathbf{g}_{R,j}^{\text{FD}} - \hat{\boldsymbol{\mu}}_{d,y_{R,j}^{\text{TD}}}^{(t)})^T \quad (34)$$

$$\hat{\Sigma}_{d,1}^{(t)} = \text{diag} \left\{ \hat{\boldsymbol{\sigma}}_{d,1}^{(t)} \right\} \quad (35)$$

$$\hat{\boldsymbol{\sigma}}_{d,1}^{(t)} = \sum_{j=1}^{2MN_c} w_j^{(t)} \left((\mathbf{g}_{R,j}^{\text{FD}} - \hat{\boldsymbol{\mu}}_{d,y_{R,j}^{\text{TD}}}^{(t)}) \odot (\mathbf{g}_{R,j}^{\text{FD}} - \hat{\boldsymbol{\mu}}_{d,y_{R,j}^{\text{TD}}}^{(t)}) \right). \quad (36)$$

Note that $w_j^{(t)}$ is the weight assigned to the j^{th} training example in the t^{th} iteration. In addition, the notation ‘‘d’’ is used as subscript in (29)-(36) to avoid confusion with channel estimation part’s of equations. Let $\hat{\mathbf{x}}_R^{\text{FD}}$ represents the output of the AdaBoost procedure using either one of the weak classifiers in (29)-(31). A normalization step is needed to match the power of the estimated signal with that of the actual transmitted signal.¹ Then, we have

$$\bar{\mathbf{x}}_R^{\text{FD}} = \frac{\sqrt{KN_c} \hat{\mathbf{x}}_R^{\text{FD}}}{\|\hat{\mathbf{x}}_R^{\text{FD}}\|_2} = [\bar{x}_{R,1}^{\text{FD}}, \bar{x}_{R,2}^{\text{FD}}, \dots, \bar{x}_{R,2KN_c}^{\text{FD}}]^T. \quad (37)$$

The final step is to detect the transmitted symbols by mapping the elements of $\bar{\mathbf{x}}_R^{\text{FD}}$ to one member of the transmitted signal constellation set denoted by \mathcal{F} by solving the following optimization problem symbol-by-symbol:

$$\hat{x}_{R,k}^{\text{FD}} = \arg \min_{x \in \mathcal{F}} |x - (\bar{x}_{R,k}^{\text{FD}} + j \bar{x}_{R,k+KN_c}^{\text{FD}})| \quad \text{for } k = 1, 2, \dots, KN_c \quad (38)$$

where $\hat{x}_{R,k}^{\text{FD}}$ is the k^{th} entry of the final estimate, i.e., $\hat{\mathbf{x}}_R^{\text{FD}} \triangleq [\hat{x}_{R,1}^{\text{FD}}, \hat{x}_{R,2}^{\text{FD}}, \dots, \hat{x}_{R,KN_c}^{\text{FD}}]^T$. The steps of the proposed data detection methods are listed in Algorithm 3.

We emphasize that the last loop in Algorithm 3 is only included for the sake of presentation clarity, and the symbol-by-symbol detection can be executed concurrently. It is also worth noting that post-processing can be performed for refining the outputs of (38) as have been suggested in [8] and [16]. The former has exploited the ML criterion to select the final data symbol from a properly designed data candidate set [8], whereas the latter has resorted to a *minimum weighted Hamming distance*-based criterion [36] to pick up the refined data symbol from a data candidate set. Despite the efficiency of the aforementioned post-processing data refinement, we do not use it here and the simulation results for the proposed methods are provided without considering the post-processing in the next section.

Remark 3: One of the advantages of the proposed AdaBoost-based algorithms is that the sufficient number of weak classifiers for obtaining a reasonable accuracy is of order of a few tens. In other words, increasing T more than some value usually does not change the performance of the proposed AdaBoost-based algorithms dramatically. We propose to set $T = 40$ for the proposed channel estimator and data detector, as using these values has been found to be effective in reaching accurate results. The impact of using different values of T for channel estimation and data detection will be examined in the next section though.

IV. SIMULATION RESULTS

In this section, numerical result that demonstrate the efficiency as well as superiority of the proposed wideband channel estimators and data detectors compared to other existing techniques are presented. In terms of computational complexity and run time, the AdaBoost-based methods are highly efficient, particularly when considering one-bit large-scale MIMO-OFDM systems. We use $T = 40$ for the proposed

¹Such normalization is also used in [8] and [16] for example.

Algorithm 3 One-bit GDA-AdaBoost Algorithm for data detection

Input: $\mathcal{S}_d = \{\mathbf{x}^{(j)} = \mathbf{g}_{R,j}^{\text{FD}}, y^{(j)} = y_{R,j}^{\text{TD}}\}_{j=1,2,\dots,2MN_c}$ whose elements are defined in (25) and (26), and number of weak classifiers T .

Output: $\hat{\mathbf{x}}_R^{\text{FD}}$.

Initialize $w_j^{(1)} = \frac{1}{2MN_c}$ for $j \in \{1, 2, \dots, 2MN_c\}$.

for $t = 1$ to T **do**

Use the training set \mathcal{S}_d to compute $\hat{\boldsymbol{\mu}}_{d,-1}^{(t)}, \hat{\boldsymbol{\mu}}_{d,1}^{(t)}, \hat{\boldsymbol{\Sigma}}_d^{(t)}$, and $\hat{\boldsymbol{\Sigma}}_{d,1}^{(t)}$ via (32)-(35), respectively. Then, compute the t^{th} weak classifier as:

one-bit GDA-AdaBoost
 $\mathbf{x}_d^{(t)} = \left(\hat{\boldsymbol{\Sigma}}_d^{(t)}\right)^{-1} \left(\hat{\boldsymbol{\mu}}_{d,1}^{(t)} - \hat{\boldsymbol{\mu}}_{d,-1}^{(t)}\right)$

one-bit GDA-AdaBoost-1
 $\mathbf{x}_d^{(t)} = \left(\hat{\boldsymbol{\Sigma}}_{d,1}^{(t)}\right)^{-1} \left(\hat{\boldsymbol{\mu}}_{d,1}^{(t)} - \hat{\boldsymbol{\mu}}_{d,-1}^{(t)}\right)$

one-bit GDA-AdaBoost-2
 $\mathbf{x}_d^{(t)} = \hat{\boldsymbol{\mu}}_{d,1}^{(t)} - \hat{\boldsymbol{\mu}}_{d,-1}^{(t)}$.

Compute error as

$$\epsilon^{(t)} = \sum_{j=1}^{2MN_c} w_j^{(t)} \mathbf{1} \left((\mathbf{g}_{R,j}^{\text{FD}})^T \mathbf{x}_d^{(t)} \neq y^{(j)} \right).$$

Compute $\alpha^{(t)} = \frac{1}{4} \ln \left(\frac{1-\epsilon^{(t)}}{\epsilon^{(t)}} \right)$.

Update $w_j^{(t+1)} = w_j^{(t)} \exp \left(\alpha^{(t)} \mathbf{1} \left((\mathbf{g}_{R,j}^{\text{FD}})^T \mathbf{x}_d^{(t)} \neq y^{(j)} \right) \right)$, $\forall j$.

Compute $Z^{(t+1)} = \sum_{j=1}^{2MN_c} w_j^{(t+1)}$ and normalize weights as $w_j^{(t+1)} = \frac{w_j^{(t+1)}}{Z^{(t+1)}}$, $\forall j$.

end for

Construct $\hat{\mathbf{x}}_R^{\text{FD}} = \sum_{t=1}^T \alpha^{(t)} \mathbf{x}_d^{(t)}$, and then normalize as $\hat{\mathbf{x}}_R^{\text{FD}} = \frac{\sqrt{KN_c} \hat{\mathbf{x}}_R^{\text{FD}}}{\|\hat{\mathbf{x}}_R^{\text{FD}}\|_2}$. Denote the k^{th} entry of $\hat{\mathbf{x}}_R^{\text{FD}}$ as $\hat{x}_{R,k}^{\text{FD}}$ for $k \in \{1, 2, \dots, 2KN_c\}$.

for $k' = 1$ to KN_c **do**

Solve the optimization problem (38) to detect $\hat{x}_{R,k'}^{\text{FD}}$.

end for

Construct $\hat{\mathbf{x}}_R^{\text{FD}} = [\hat{x}_{R,1}^{\text{FD}}, \hat{x}_{R,2}^{\text{FD}}, \dots, \hat{x}_{R,KN_c}^{\text{FD}}]^T$.

AdaBoost-based channel estimators and data detectors, unless otherwise stated. For channel estimation figures, orthogonal pilots are employed analogous to those suggested in [5, Eq. (23)]. In addition, quadrature phase shift keying (QPSK) constellations are used as the frequency domain symbols in data detection figures. The hyperparameter C is set to 1 for SVM-based channel estimator and data detector of [16]. Furthermore, the modified finite Newton (MFN) method [37] is used for implementing the ℓ_2 -SVMs as it is one of the most efficient algorithms [16]. Performance of different channel estimators and data detectors are compared in terms of normalized MSE (NMSE) and bit-error-rate (BER), respectively. The former is defined as

$$\text{NMSE} = \frac{\mathbb{E}\{\|\mathbf{H} - \hat{\mathbf{H}}\|_F^2\}}{KM}$$

where $\mathbf{H} \triangleq [\mathbf{h}_1^{\text{TD}}, \mathbf{h}_2^{\text{TD}}, \dots, \mathbf{h}_M^{\text{TD}}]$ and $\hat{\mathbf{H}} \triangleq [\hat{\mathbf{h}}_1^{\text{TD}}, \hat{\mathbf{h}}_2^{\text{TD}}, \dots, \hat{\mathbf{h}}_M^{\text{TD}}]$. The block-fading interval is divided into two parts, where the first part and second part are used for channel estimation and data detection, respectively.

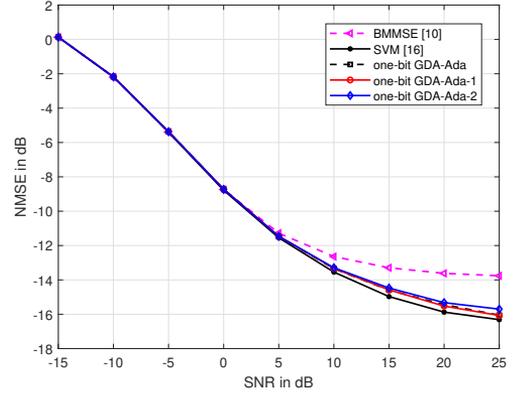


Fig. 1. Performance comparison of different channel estimators with $K = 2$, $M = 16$, $N_c = 256$, and $L_{\text{tap}} = 8$.

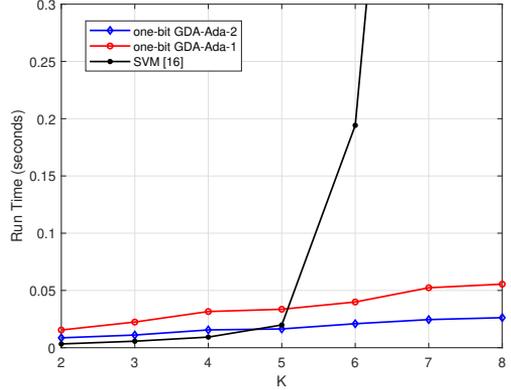


Fig. 2. Average run time comparison of the proposed one-bit GDA-Ada-1 and one-bit GDA-Ada-2 with SVM in estimating channel between users and one antenna of the BS vs. the number of users K , considering the scenario where $N_c = 512$, and $L_{\text{tap}} = 16$.

Noteworthy to mention that the performance of one-bit GDA-Ada method (when the covariance matrix has to be computed exactly) is not reported in data detection figures as its computational complexity is higher than that of the proposed one-bit GDA-Ada-1 and one-bit GDA-Ada-2 methods for achieving similar performance.

In Fig. 1, the NMSE of the proposed AdaBoost-based channel estimators are compared with those of BMMSE [10] and SVM [16]. It can be observed that the performance of BMMSE is worse than other methods tested, while SVM possesses the best performance. The AdaBoost-based channel estimators are very marginally outperformed by SVM that has no effect on the follow up data detection. The one-bit GDA-Ada-2 method provides the worst performance among the proposed Adaboost-based channel estimators.

Despite having comparable channel estimation performance, the proposed one-bit GDA-Ada-1 and one-bit GDA-Ada-2 require substantially lower computational complexity compared to that of the SVM-based method as depicted in Fig. 2. We compare the required average run time for estimating channel between users and one antenna of the BS (i.e., average run time for estimating \mathbf{h}_i^{TD} 's). Although the average run times for performing the channel estimation task are comparable for all methods tested when $K \leq 5$, the SVM-based channel estimator needs much higher computational complexity than those of one-bit GDA-Ada-1 and one-bit GDA-Ada-2 when $K > 5$. We stress here that this advantage of the proposed methods is rooted in using low computation demanding techniques as weak classifiers in Algorithm 2. Moreover, the computational complexity order of different channel estimators tested is listed in Table I, where $\kappa(\cdot)$ represents a super-linear function.

TABLE I: Order of Computational Complexity for Different Channel Estimators.

Method	Complexity
BMMSE	$\mathcal{O}(M^2 K L_{\text{tab}} N_c)$
SVM-based	$\mathcal{O}(M K L_{\text{tab}} N_c \kappa(N_c))$
one-bit GDA-Ada	$\mathcal{O}(T M \max\{(K L_{\text{tab}})^{2.373}, (K L_{\text{tab}})^2 N_c\})$
one-bit GDA-Ada-1	$\mathcal{O}(T M K L_{\text{tab}} N_c)$
one-bit GDA-Ada-2	$\mathcal{O}(T M K L_{\text{tab}} N_c)$

In Fig. 3, a performance comparison of different proposed AdaBoost-based channel estimators is presented for the values of $T = 1$, $T = 10$, and $T = 40$. It can be seen that the NMSE associated with $T = 1$ is the worst, whereas the difference between $T = 10$ and $T = 40$ is negligible for the channel estimation task. Combining the results of Figs. 2 and 3 shows that the run time superiority of the proposed one-bit GDA-Ada-1 and one-bit GDA-Ada-2 over SVM studied in Fig. 2 can be even more considerable by setting $T = 10$ instead of $T = 40$ without sacrificing performance.

Fig. 4 compares the NMSE of the proposed AdaBoost-based channel estimators with SVM for $N_c = 256$ and $N_c = 1024$, where the NMSEs of the methods tested are decreased for about 4 dB at high SNRs by increasing N_c from 256 to 1024. Analogous to Fig. 1, the proposed AdaBoost-based channel estimators possess quite similar performance to the performance of the SVM-based channel estimator.

Fig. 5 compares the one-bit GDA-Ada-1 and one-bit GDA-Ada-2 data detectors with the SVM data detector for both cases of estimated CSI and perfect CSI. It should be noted here that the estimated CSI of each method is found by their corresponding channel estimators. It can be seen in Fig. 5 that the BERs of the proposed AdaBoost-based data detectors are quite close to the BER of the SVM data detector for both cases of estimated CSI and perfect CSI. Moreover, Fig. 5 shows that the performance of the one-bit GDA-Ada-1 method is slightly better than that of the one-bit GDA-Ada-2 method at high SNRs.

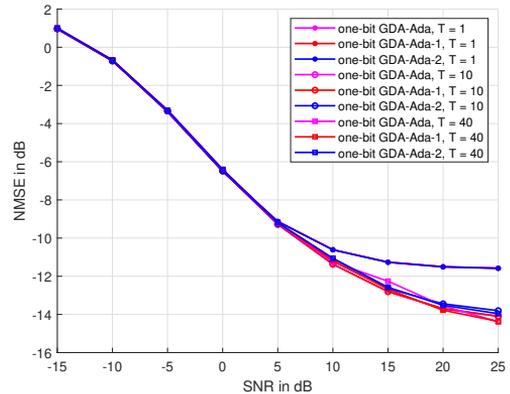


Fig. 3. NMSE comparison of the proposed Adaboost-based channel estimators for different values of T with $K = 4$, $M = 32$, $N_c = 512$, and $L_{\text{tap}} = 16$.

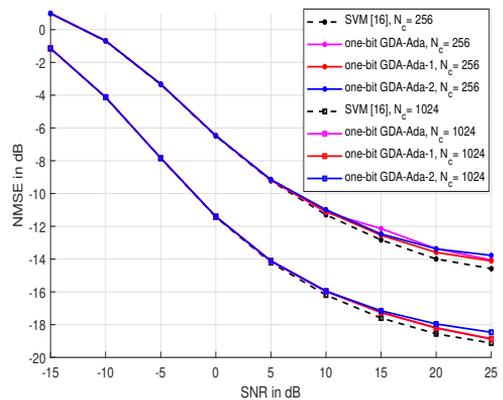


Fig. 4. Performance comparison between the AdaBoost-based channel estimators and SVM with $K = 2$, $M = 32$, $L_{\text{tap}} = 16$, and $N_c \in \{256, 1024\}$.

An average run time comparison for implementing the one-bit GDA-Ada-1, one-bit GDA-Ada-2, and SVM data detectors is presented in Fig. 6, where $K \in \{2, 3, \dots, 8\}$. For $K \leq 7$, the average run time of the SVM data detector is comparable with the average run times of the one-bit GDA-Ada-1 and one-bit GDA-Ada-2, whereas the average run time of the SVM data detector is substantially higher than those of the other two methods when $K > 7$. Fig. 6 also shows that the average run time of the one-bit GDA-Ada-2 method is approximately half of the average run time of the one-bit GDA-Ada-1 method. The computational complexity order of different data detectors tested is listed in Table II.

The impact of choosing different T on the performance of the proposed AdaBoost-based data detectors is investigated in Fig. 7. Contrary to the comparison presented in Fig. 3 for the channel estimation task, it is observed in Fig. 7 that the BER improves by increasing T from 10 to 40. Moreover,

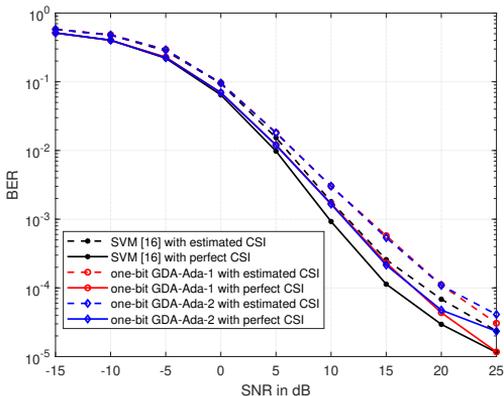


Fig. 5. Performance comparison of different data detectors with $K = 2$, $M = 16$, $N_c = 256$, $L_{\text{tap}} = 8$, and QPSK modulation.

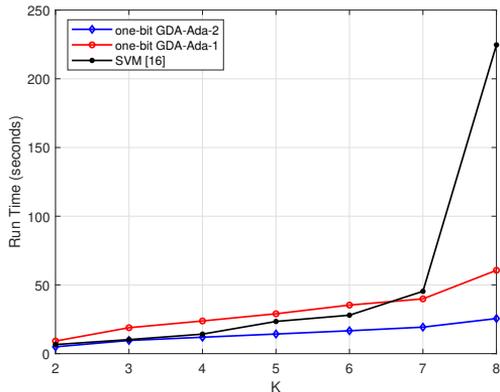


Fig. 6. Average run time for implementing different data detectors with various K , $N_c = 512$, $M = 32$, and $L_{\text{tap}} = 8$.

TABLE II: Order of Computational Complexity for different data detectors.

Method	Complexity
SVM-based	$\mathcal{O}(MKN_c^2\kappa(MN_c))$
one-bit GDA-Ada	$\mathcal{O}(T\max\{(KN_c)^{2.373}, MK^2N_c^3\})$
one-bit GDA-Ada-1	$\mathcal{O}(TMKN_c^2)$
one-bit GDA-Ada-2	$\mathcal{O}(TMKN_c^2)$

the performance gap of using $T = 1$ with other values of T is significant. We stress here that setting T larger than 40 marginally improves the data detection performance based on our observations, hence $T = 40$ is a proper choice.

V. CONCLUSION

In this paper, we have found out and demonstrated that the GDA classifier/approximated GDA classifier together with the AdaBoost technique result in developing efficient and reliable channel estimators and data detectors, specifically

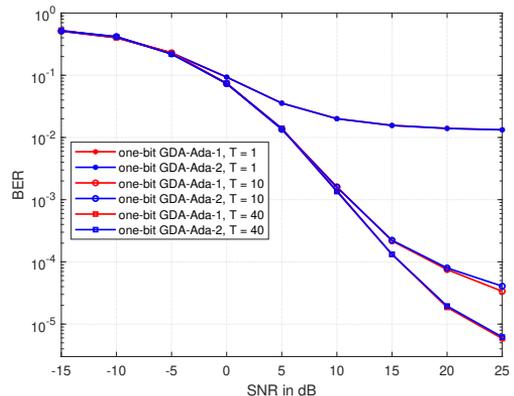


Fig. 7. The impact of different T on the performance of the proposed AdaBoost-based data detectors with $K = 4$, $N_c = 512$, $M = 32$, $L_{\text{tap}} = 16$, and QPSK modulation.

in large scale scenarios such as MIMO-OFDM systems that operate over frequency selective channels. It was shown that two of the proposed AdaBoost-based channel estimators and data detectors named one-bit GDA-Ada-1 and one-bit GDA-Ada-2 require dramatically lower run time compared to those of the SVM-based methods, while maintaining the accuracy. Numerical results were presented to showcase the efficiency and robustness of the proposed methods in large scale MIMO-OFDM systems.

REFERENCES

- [1] 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.
- [2] 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.
- [3] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.
- [4] B. Murmann, "Energy limits in A/D converters," *Proc. IEEE Faible Tension Faible Consommation*, 2013, pp. 1–4.
- [5] 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.
- [6] 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.
- [7] J. Liu, Z. Luo, and X. Xiong, "Low-resolution ADCs for wireless communication: A comprehensive survey," *IEEE Access*, vol. 7, pp. 91291–91324, 2019.
- [8] J. Choi, J. Mo, and R. W. Heath, "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, May 2016.
- [9] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, Jun. 2016.
- [10] 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.
- [11] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. 216, 1952.
- [12] 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.

- [13] 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.
- [14] P. Sun, Z. Wang, R. W. Heath, 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.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] 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.
- [32] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: Data mining, inference, and prediction*. Springer, 2009.
- [33] 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.
- [34] [online]: https://sgfin.github.io/files/notes/CS229_Lecture_Notes.pdf
- [35] T. Hastie, R. Saharon, Z. Ji, and Z. Hui, "Multi-class AdaBoost," *Statistics and its Interface*, vol. 2, no. 3, pp. 349–360, 2009.
- [36] 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.
- [37] S. S. Keerthi and D. DeCoste, "A modified finite newton method for fast solution of large scale linear SVMs," *J. Mach. Learn. Res.*, vol. 6, pp. 341–361, Mar. 2005.



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**