

2019

Doctorate thesis

**Advanced Direction-of-Arrival
Estimation for Acoustic Signal
Processing and its Applications**

1208005 Bandhit Suksiri

Advisor Prof. Masahiro Fukumoto

August 2019

Course of Information Systems Engineering
Graduate School of Engineering, Kochi University of Technology

Abstract

Advanced Direction-of-Arrival Estimation for Acoustic Signal Processing and its Applications

Bandhit Suksiri

Recent year have seen rapidly increasing demand for services and systems that depend upon accurate positioning of people and machines. This has led to the development and evolution of numerous positioning systems. In addition to these systems, direction-of-arrival (DOA) estimation in particular plays a critical role in navigation systems for the exploration of sources in widespread applications, such as voice activity detection, human computer interaction, automatic camera steering, robotics and surveillance. Although wide variety of DOA estimation methods have been proposed for the multitude of applications, DOA estimation for acoustic sources however has not been widely investigated, and many significant methods cannot be applied to acoustic circumstances directly. The reason is that the most of DOA estimations are based on wireless communications and characteristics of wireless signal are totally different from the acoustic signal. Furthermore, the presence of reverberation and background noise present challenges that need to be addressed in a realistic environment.

This dissertation therefore aims to bridge a research gap of acoustic source compatibility on recent DOA estimation methods for estimating DOA of the acoustic sources directly and effectively. Our research works focus on the development of new frameworks, suitable theories and extended techniques for estimating acoustic DOAs in the hope of improving efficacy, simplicity, and yet accuracy, to solve practical problem.

Since the conventional signal modeling techniques solely focus on a narrowband signal, which is necessary to consider an acoustic source or equivalence to wideband source for more practical use, especially in case of human speech. Therefore, the first research work of this dissertation begins with an alternative signal modeling for acoustic sources with L-shaped microphone array configuration. The problem of estimating acoustic DOAs is addressed and resolved by declaring terms of temporal and reference frequencies in array manifold matrices of the signal model. In addition to this, the proposed model is now compatible with almost recent narrowband DOA estimation methods, it means that this model enables cutting-edge techniques in the existing narrowband subspace methods to apply acoustic source directly and efficiently.

After the acoustic signal modeling technique has been successfully given, we continue to present DOA estimation method for acoustic sources by using unsupervised learning along with this signal model. The problem of estimating DOA of acoustic signals is addressed by focusing the entire observation subspace in each frequency bin. In addition to this observation process, we employ the Gaussian mixture model with a maximum likelihood estimation algorithm as an interpolation scheme, and the DOA results are estimated easily by interpolating the multi-narrowband DOA results all frequency bins with this interpolation scheme. The performance is evaluated in terms of root-mean-square-error over a range of signal-to-noise ratio (SNR), and the results suggest that the proposed method is a particularly effective method of DOA estimation with a requirement of high computational resources.

In the next part of work, we present an efficient wideband or acoustic 2D DOA estimation, which is an improved version of the previous method. We propose a way to construct a wideband sample cross-correlation matrix without any process of DOA preliminary estimation, such as beamforming technique, by exploiting sample cross-correlation matrices of two different frequencies for all frequency bins along with the

Orthogonal Procrustes analysis. Subsequently, wideband DOAs are estimated by using this wideband matrix along with a recent scheme of estimating DOA in a narrowband subspace method. A contribution of this work is providing an alternative framework for recent narrowband subspace methods to estimating the DOA of wideband sources directly. It means that this framework enables cutting-edge techniques in the existing narrowband subspace methods to implement the wideband direction estimation for reducing the computational complexity and facilitating the estimation algorithm. Theoretical analysis and effectiveness of the proposed method are substantiated through numerical simulations, and the results show that performance of the proposed method performs better than others over a range of SNR with just a few microphones.

In addition to Orthogonal Procrustes analysis as used in the previous work, in the next part of work, we improve accuracy performance of the previous framework for estimating wideband or acoustic DOA by proposing an extension theory of Orthogonal Procrustes analysis with much more efficiently than the original theory. The proposed framework is inspired by the coherent signal subspace technique with further improvement of linear transformation procedure, and the new procedure no longer require any process of DOA preliminary estimation by exploiting unique cross-correlation matrices between received signal and itself on distinct frequencies, along with the higher-order generalized singular value decomposition of the array of this unique matrix. Wideband DOAs are estimated by employing any subspace-based technique for estimating narrowband DOAs, but using the proposed wideband correlation instead of a conventional narrowband correlation matrix. It implies that the proposed framework enables cutting-edge researches in the recent narrowband subspace methods to estimate DOA of the wideband sources directly, which result in the reducing computational complexity and facilitating the estimation algorithm. This thus be a major contribution of the proposed framework as same as the previous framework, but much more efficiently than

the previously. Practical examples are presented to showcase its applicability and effectiveness, and the experimental results show that the performance of fusion methods perform better than others over a range of SNR with just a few sensors, which make it suitable for practical use.

Unlike all previous methods that have focused on estimating wideband or acoustic DOAs, the final part of this work addresses a problem to estimate a variance of sources for all frequency bins and its DOAs simultaneously. Signal model is renovated into a tensor representation of the three features: x-subarray angle, variance-of-frequencies, and z-subarray manifold angle. In particular, complex-valued parallel factor analysis is used as the tensor factorization, and the variance-of-frequencies and its DOAs are now estimated simultaneously via this factorization technique. Effectiveness of proposed method is substantiated through numerical simulations, and it suggests that our method provides a promising alternative for intelligent source localization.

In the end, we believe that our research findings presented in this dissertation will appeal to researchers who wish to develop a sound source based navigation system and improve its robust estimation. We also hope that these findings can be a good alternative for estimating DOA of acoustic sources, especially human speeches and musical sounds.

key words direction-of-arrival (DOA); higher-order generalized singular value decomposition (HOGSVD); wideband sources; sound source; array processing; subspace method; cross-correlation

Contents

Chapter 1	Fundamentals of Position Location and its Recent Issues	1
1.1	Introduction of Positioning System	1
1.2	Basic Methods Used in Positioning	2
1.2.1	Time-of-Arrival Estimation	2
1.2.2	Time-Difference-of-Arrival Estimation	3
1.2.3	Received Signal Strength Indicator	4
1.2.4	Direction-of-Arrival Estimation	5
1.3	Positioning System in Acoustic Signal Processing	7
1.4	Motivations and Objectives of this Dissertation	7
1.5	Dissertation Organization	8
Chapter 2	An Alternative Signal Modeling for Wideband Sources with L-Shaped Microphone Array Configuration	12
2.1	Reviews on Recent Signal Modeling Techniques	13
2.2	Phase Difference Equation	16
2.3	The New Data Model	18
2.4	Example of DOA Estimation Schemes with the New Data Model	20
2.4.1	Extension of 2D MUSIC Method	20
2.4.2	Extension of Wang et al. Method	21
2.4.3	Extension of Dong et al. method	21
2.5	Numerical Simulation	23
2.5.1	Robustness of the Cost Functions	25
2.5.2	RMSE versus SNR	26
2.6	Conclusions	27

2.7	References	28
Chapter 3	Wideband Direction-of-Arrival Estimation by using Unsu-	
	pervised Learning	34
3.1	Reviews on Recent Wideband DOA Estimations	34
3.2	DOA Estimation Scheme	36
3.3	Signal Synthesis	39
3.4	Numerical Simulation	39
3.5	Conclusions	41
3.6	References	43
Chapter 4	An Efficient Wideband Direction-of-Arrival Estimation by	
	using Orthogonal Procrustes Analysis	44
4.1	Robustness Issue in Recent Wideband DOA Estimation	45
4.2	Preliminaries	49
	4.2.1 Data Model	49
	4.2.2 Basic Assumptions	50
	4.2.3 Problem Formulation	52
4.3	Proposed Method	54
	4.3.1 New Optimization Problems and its Solutions	54
	4.3.2 Extension Properties on SVD of the Correlation Matrix	63
	4.3.3 DOA Estimation Scheme	65
4.4	Computational Complexity	68
4.5	Numerical Simulations	70
4.6	Conclusion	78
4.7	References	79

Chapter 5	Extension Theory of Orthogonal Procrustes Analysis for Wideband Direction-of-Arrival Estimation and its Experimentation	85
5.1	Foreword: New Possible DOA Approach via High-Order Signal Subspace	86
5.2	Preliminaries	90
5.2.1	Data Model	90
5.2.2	Redefinition of Model Assumptions	91
5.2.3	Transformation Problem	93
5.3	Proposed Method	94
5.3.1	Problem for Estimating the Transformation Matrix and its Solution	94
5.3.2	Estimation of the Transformation Matrices by HOGSVD	102
5.3.3	DOA Estimation Scheme	107
	DOA Estimation Scheme via MUSIC	109
	DOA Estimation Scheme via ESPRIT	111
5.4	Numerical Simulations	112
5.4.1	Scenario 1: Performance with Respect to Source Types	116
5.4.2	Scenario 2: Performance with Respect to the Number of Micro- phone Elements	117
5.4.3	Scenario 3: Performance with Considering Automatic Pairing	118
5.4.4	Scenario 4: Performance under Reverberation Environment	120
5.4.5	Computational Complexity	122
5.5	Experimental Results	123
5.6	Conclusions	129
5.7	References	131

Chapter 6 Complex-Valued Tensor Factorization for Estimating Di-

Contents

	rection of Wideband Sources and its Variance-	
	of-Frequencies	137
6.1	Problem of Pair Matching Method on Variance-of-Frequency	138
6.2	Proposed Method	139
6.3	Numerical Simulations	143
6.4	Conclusion	145
6.5	References	147
Chapter 7	Conclusion	149
	Acknowledgement	152

List of Figures

- 1.1 Operation of TOA. 3
- 1.2 Operation of TDOA. 4
- 1.3 Operation of DOA. 5
- 1.4 Examples of acoustic source based positioning system; (a) Moving source localization for automatic camera steering, which has found with lots of applications recently such as lecture recording. (b) Passive acoustic detection and computer-based signal processing to locate a shooter in less than a second; it detects small arms fire travelling toward it for bullets passing within approximately 30 meters of the mast-mounted compact array of microphones.(c) Microphone array technology for the “robot ears” in finding victims needing rescue at disaster sites; the three keys are sound source localization technology to estimate where sound is coming from, sound source separation technology to extract the direction from which the sound originates, and automatic speech recognition technology to recognize separated sounds from background noise. (d) A real-time direction-of-arrival estimation for sequential movement events of vehicles. 6
- 1.5 Motivations and objectives of this research works. 8
- 2.1 L-shaped microphone array configuration for 2D DOA estimation. 15
- 2.2 Normalized radiation patterns of the L-shaped arrays respect to DOA angles; (a) $f = f_c$ and (b) $f = f_c/4$ where $M = 8, d = \lambda/2$ 16
- 2.3 Normalized radiation patterns of the L-shaped arrays respect to coordinates; (a) $f = 4f_o$, (b) $f = 2f_o$, (c) $f = f_o$, (d) $f = \frac{f_o}{2}$ where $M = 12, \lambda = \frac{d}{2}$ 17

List of Figures

2.4	Search space results of three types of cost functions along with the proposed array manifold matrices; (a), (b) 2D MUSIC in Eq. (2.13); (c), (d) the Wang et al. method in Eq. (2.16); (e), (f) Dong et al. method in Eq. (2.26), where source frequency in (a), (c), (e) are 3.4 kHz and (b), (d), (f) are 2.26 kHz.	24
2.5	RMSE estimation performance of DOA cost functions along with the proposed array manifold matrices; (a) 2D MUSIC method in Eq. (2.13), (b) the Wang et al. method in Eq. (2.16) and (c) Dong et al. method in Eq. (2.26). Three sources are placed $(\theta_k^{\text{DOA}}, \bar{\phi}_k^{\text{DOA}})$ at $(60^\circ, 30^\circ)$, $(45^\circ, 45^\circ)$ and $(30^\circ, 60^\circ)$	28
3.1	Overviews of wideband DOA estimation by using GMM where the raw DOA datasets are obtained by estimating narrowband DOAs all frequency bins.	38
3.2	Estimation results of the proposed method by (a) Root MUSIC and GMM fusion and (b) MUSIC where three uncorrelated sources are employed.	39
3.3	RMSE performance evaluation; (a) 0° , (b) 10° , (c) 20°	41
3.4	Example of signal synthesis with error magnitude where three uncorrelated human speeches and 16 microphones are used; (a) first source, (b) second source, and (c) third source.	42
4.1	Computational complexities; (a) changing the number of snapshots N , (b) changing the number of frequency bins F , (c) changing the number of microphone elements each subarray M , and (d) changing the number of incident sources K	69

List of Figures

4.2	Performance evaluations of Scenario 2 via computer simulation; (a), (c), (e) RMSE versus SNR, (b), (d), (f) SD versus SNR, where the number of microphone elements each subarray on (a), (b) $M = 6$, (c), (d) $M = 10$, and (e), (f) $M = 14$	74
4.3	Performance evaluations of Scenario 3 via computer simulation; (a) RMSE versus SNR, and (b) SD versus SNR under the reverberation time environment.	75
5.1	RMSE estimation performance versus SNR on Scenario 1; (a) three different human speeches, and (b) three uncorrelated musical sounds where six microphones is employed each subarray.	113
5.2	SD estimation performance versus SNR on Scenario 1; (a) three different human speeches, and (b) three uncorrelated musical sounds where six microphones is employed each subarray.	113
5.3	RMSE estimation performance versus SNR on Scenario 2; three human speeches are employed and the number of microphone elements each subarray on (a) $N = 4$, (b) $N = 8$, and (c) $N = 12$	114
5.4	SD estimation performance versus SNR on Scenario 2; three human speeches are employed and the number of microphone elements each subarray on (a) $N = 4$, (b) $N = 8$, and (c) $N = 12$	114
5.5	RMSE estimation performance versus SNR on Scenario 3 where $M = 8$	115
5.6	SD estimation performance versus SNR on Scenario 3 where $M = 8$	115

List of Figures

5.7 Performance evaluations of Scenario 4; (a) RMSE estimation performance versus SNR, and (b) SD estimation performance versus SNR, where three uncorrelated human speeches are employed along with a reverberant environment, where dimensions of enclosure room is $15 \times 15 \times 5$ meter, a measurement protocol of reverberation time is RT60, and wall absorption coefficients are followed on Table 5.1. 119

5.8 Computational complexities; (a) changing the number of microphone elements each subarray N , and (b) the number of microphone elements including all subarray M where the number of incident sources $K = 3$. . 122

5.9 Photograph of the microphone array system. 127

5.10 Photograph of the experimental environment, floor plan and the room dimensions. 130

6.1 The tensor representation of sample cross-correlation. 139

6.2 Numerical simulation on variance-of-frequency estimation; (a), (b), (c) singular values or estimated variance-of-frequency, (d), (e), (f) actual variance-of-frequency, (g), (h), (i) actual magnitude of DFT, where (a), (d), (g) are the first source, (b), (e), (h) are the second source, (c), (f), (i) are the third source. 144

6.3 Performance evaluations via simulation; (a), (c), (e) RMSE versus SNR and (b), (d), (f) SD versus SNR where the number of microphone elements each subarray on (a), (b) $M = 6$, (c), (d) $M = 8$, and (e), (f) $M = 10$ 146

List of Tables

- 2.1 Cramer-Rao bound of array geometrical structures 14

- 4.1 Performance evaluations of Scenario 1 via a computer simulation; $K = 2, 3$ 71
- 4.2 Performance evaluations of Scenario 1 via a computer simulation; $K = 4$ 72
- 4.3 Performance evaluations of Scenario 1 via a computer simulation; $K = 5$ 73

- 5.1 Wall absorption coefficients at various reverberation time in Scenario 4 . 120
- 5.2 System specification 124
- 5.3 Performance evaluation on Experiment 1; $K = 1, 2$ 125
- 5.4 Performance evaluation on Experiment 1; $K = 3$ 126
- 5.5 Performance evaluation on Experiment 2; $K = 1, 2, 3$ 128
- 5.6 Performance evaluation on Experiment 2; $K = 4$ 129

Chapter 1

Fundamentals of Position

Location and its Recent Issues

Recent year have seen rapidly increasing demand for services and systems that depend upon accurate positioning of people and objects. This has led to the development and evolution of numerous positioning systems. This chapter provides an overview of the main positioning techniques. We introduce the positioning systems that are in use for a variety of applications including in acoustic signal processing. The chapter concludes by reviewing recent positioning techniques and outlining the research objectives.

1.1 Introduction of Positioning System

Positioning systems determine the location of a person or an object either relative to a known position or within a coordinate system. In the last few decades, various positioning systems have been motivated by demand and have been developed. Some of the applications of positioning systems include law enforcement, security, road safety, tracking personnel, vehicles, and other assets, situation awareness, and mobile ad hoc networks. Nowadays, the local positioning in indoor environments with mobile devices is becoming increasingly relevant. Some applications of local positioning system are the location, tracking and guidance of people in indoor environments, as well as new ways of interacting with the objects in their surroundings.

1.2 Basic Methods Used in Positioning

In this section, the fundamental techniques of positioning systems are explained. The first category is based on the propagation time of the signals as an estimate for a distance between target and node. This time-based techniques can be grouped in time-of-arrival (TOA) and time-difference-of-arrival (TDOA) techniques. The second category measures a received signal strength (RSS). The strength of received signal indicates the distance traveled by the signal. The third category is based on the directional property of antenna arrays for detecting a direction-of-arrival (DOA). In fact, TOA, TDOA, RSS, DOA can be used alone or in combinations to improve the qualities of measurements of received signals.

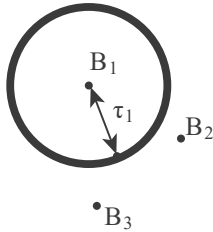
1.2.1 Time-of-Arrival Estimation

Time-of-arrival (TOA) estimation allows the measurement of distance by measuring propagation times in each node. Multiple base nodes collaborate to localize a target node via triangulation. It is assumed that the positions of all base nodes are known. If these nodes are dynamic, a positioning technique is used to allow base nodes to localize their positions. In some circumstances, multiple base nodes may cooperate to find their own position before any attempt to localize a target node.

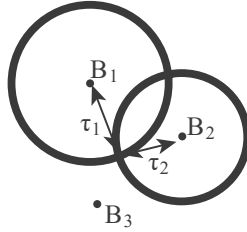
Assuming known positions of base nodes, three base nodes and three measurements of distances are required to localize a target node as illustrated in Fig. 1.1. Using the measurement of distance, the position of a target node is localized within a sphere of radius with the i^{th} receiver at the center of the sphere. The localization of the target node can be carried out either by base nodes using a master station or by the target node itself. TOA has a few drawbacks: (1) it requires all nodes to precisely synchronize, (2) the transmitted signal must be labeled with a time stamp in order to allow the base

1.2 Basic Methods Used in Positioning

Potential target location will lie on a circle with radius R_1, τ_1 .



Potential target location is one of the two points of intersection of the circles with radius R_1, τ_1 and R_2, τ_2 .



Final target location will be the point that passes through the circle with radius R_3, τ_3 .

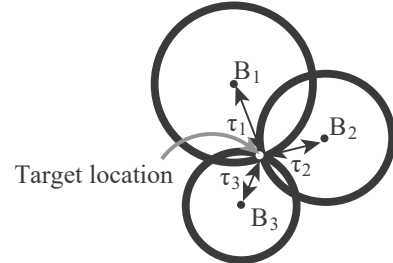


Fig. 1.1 Operation of TOA.

node to determine the time at which the signal was initiated at the target, and (3) the positions of the base nodes should be known.

1.2.2 Time-Difference-of-Arrival Estimation

Time-difference-of-arrival (TDOA) estimation requires the measurement of the difference in time between the signals arriving at two base nodes. Similar to TOA estimation, this method assumes that the positions of base nodes are known; TOA difference at the base nodes can be represented by a hyperbola as shown in Fig. 1.2. A hyperbola is the locus of a point in a plane such that the difference of distances from two fixed points is a constant.

Assuming known positions of base nodes, three base nodes and two TDOA measurements are required to localize a target node. The base station that first receives the signal from the target node is considered as the reference base station as illustrated in Fig. 1.2. TDOA measurements are made with respect to the reference base station.

Although TDOA addresses the first drawback of TOA by removing the requirement of synchronizing the target node clock with base node clocks, TDOA is the time differ-

1.2 Basic Methods Used in Positioning

First sensor that first receives the signal from the target location becomes the reference point.

Potential target location will be on a hyperbola that is formed using one TDOA measurement (1-2) with respect to reference point.

Final target location will be the point of intersection of two hyperbolas, which are formed using two TDOA measurements (1-2) and (1-3) with respect to the reference point.

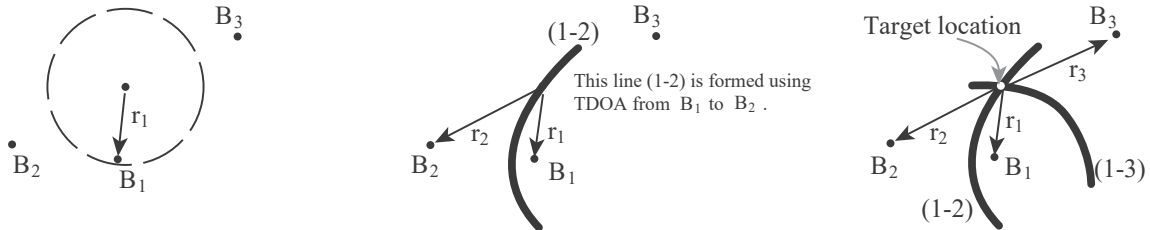


Fig. 1.2 Operation of TDOA.

ence between the end times. Therefore, only base node clocks need to be synchronized to ensure minimum measurement error. The base node clock can be synchronized externally by using a backbone network or internally using timing standards provided at the nodes. The fact that synchronization of target nodes is not required enables many applications for TDOA-based systems. With respect to the second drawback of TOA, since a single TDOA measurement is the difference in the arrival time at the respective base nodes, the transmitted signal from the target node in TDOA need not contain a time stamp. This advantage of TDOA is again exploited by many applications as well.

1.2.3 Received Signal Strength Indicator

In this method, a location estimation is carried out using the received signal strength indicator (RSSI). The strength of the received signal indicates the distance traveled by the signal by assuming that the transmission strength and channel characteristics are known. Similar to the TOA, in RSSI, multiple base nodes collaborate to localize a target node via triangulation.

1.2 Basic Methods Used in Positioning

Potential target location will lie on a line whose direction is determined by peak incoming energy signal using sensor array.

Final target location will be a point that passed through the intersection of two lines whose directions are determined by peak incoming energy signals using sensors arrays at two base nodes.



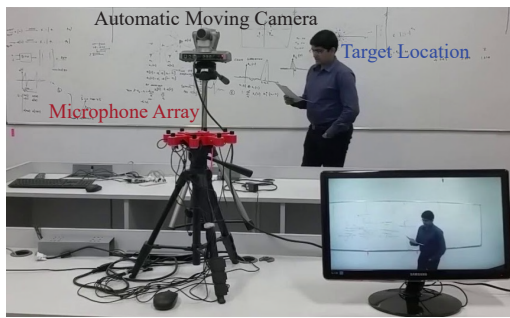
Fig. 1.3 Operation of DOA.

1.2.4 Direction-of-Arrival Estimation

In direction-of-arrival (DOA) estimation, base nodes determine the angle of the arriving signal. The system is required to equip with antenna arrays in order to allow base stations to estimate DOA as shown in Fig. 1.3. Similar to TOA and TDOA estimation, in DOA estimation, the positions of base nodes should be known. Unlike TOA and TDOA, only two base nodes along with two DOA measurements are required; it means that only two base nodes equipped with antenna arrays are sufficient to maintain full localization of a target node. This adds a higher flexibility to DOA estimation techniques compared with TOA or RSSI estimation methods. Furthermore, DOA addresses the first and second drawback of TOA by removing the requirement of synchronizing the target node clock with base node clocks because it does not require the time.

A wide variety of DOA estimation algorithms have been proposed for a multitude of applications such as sonar localization, radar detection, and mobile communication systems. DOA estimation in particular also plays a critical role in navigation systems for the exploration of sound sources in acoustic signal processing, which the next section will address further.

1.2 Basic Methods Used in Positioning



Reference:

[1] Sundar, Harshavardhan; Deepak, Gokul; Sreenivas, Tv; Seelamantula, Chandra Sekhar. (2016). Reverberation-Robust One-Bit TDOA Based Moving Source Localization for Automatic Camera Steering. 10.21437/Inter-speech.2016-575.

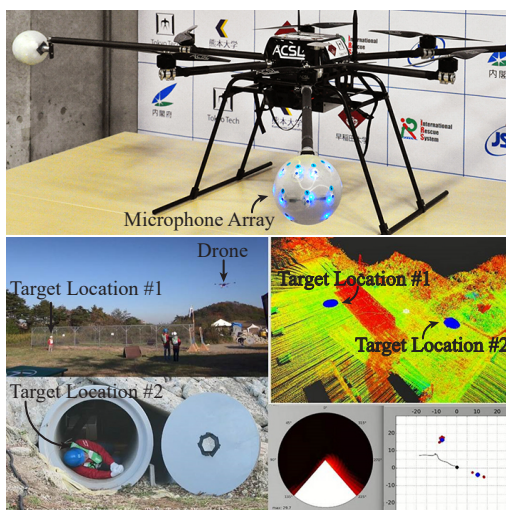
(a)



References:

[2] Raytheon. (2019, June 5) Boomerang III: State-of-the-art Shooter Detection. Retrieved from <http://www.chameleondaq.com/applications.html>.
[3] PVI System. (2019, June 5) Chameleon: Shooter Detection System Testing. Retrieved from <http://www.chameleondaq.com/applications.html>.

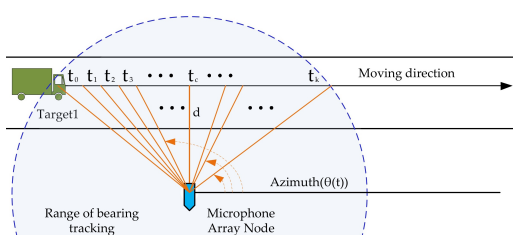
(b)



Reference:

[4] Satoshi Tadokoro . (2017, December 25) Impulsing Paradigm Change through Disruptive Technologies Program (ImPACT): Tough Robotics Challenge. Retrieved from <http://www.jst.go.jp/impact/en/program/07.html>.

(c)



Reference:

[5] Liu, Huawei; Li, Baoqing; Yuan, Xiaobing; Zhou, Qianwei; Huang, Jingchang. 2018. "A Robust Real Time Direction-of-Arrival Estimation Method for Sequential Movement Events of Vehicles." Sensors 18, no. 4: 992.

(d)

Fig. 1.4 Examples of acoustic source based positioning system; (a) Moving source localization for automatic camera steering, which has found with lots of applications recently such as lecture recording. (b) Passive acoustic detection and computer-based signal processing to locate a shooter in less than a second; it detects small arms fire travelling toward it for bullets passing within approximately 30 meters of the mast-mounted compact array of microphones.(c) Microphone array technology for the “robot ears” in finding victims needing rescue at disaster sites; the three keys are sound source localization technology to estimate where sound is coming from, sound source separation technology to extract the direction from which the sound originates, and automatic speech recognition technology to recognize separated sounds from background noise. (d) A real-time direction-of-arrival estimation for sequential movement events of vehicles.

1.3 Positioning System in Acoustic Signal Processing

There are some positioning system, as the one proposed in this dissertation, that benefit from the low propagation speed of sound in air to achieve centimeter accuracy by measuring the DOA of acoustic signals. These acoustic local positioning systems can be easily implemented in portable devices or smart speaker since most of them are equipped with audio recording hardware off-the-shelf. The frequency response of mobile device's microphones and speakers is always below 22 kHz and their audio acquisition sampling rates are not higher than 96 kHz, which is enough for a plenty of acoustic-based applications such as voice activity detection, human computer interaction, automatic camera steering, beamforming, robotics and surveillance as shown in Fig. 1.4.

1.4 Motivations and Objectives of this Dissertation

Historically, measurement of the features of received signals needed for input to position estimation algorithms have relied on four methods: TOA, TDOA, RSSI and DOA. In addition to these methods, DOA estimation has been attracting a great deal of attention in acoustic signal processing and applications as shown in Fig. 1.4. According to the literature reviews, however, there are only few numbers of high efficient techniques and suitable theory when it comes to the acoustic DOA estimation. The reason is that the most of DOA estimations are based on wireless communications and characteristics of wireless signal are totally different from the acoustic signal. Furthermore, the presence of reverberation and background noise present challenges that need to be addressed in a realistic environment.

Therefore, this dissertation focuses on the development of extension of techniques, new framework and suitable theory for estimating acoustic DOAs, which a capable of high accuracy detection for acoustic sources. Given that DOA estimation plays a criti-

1.5 Dissertation Organization

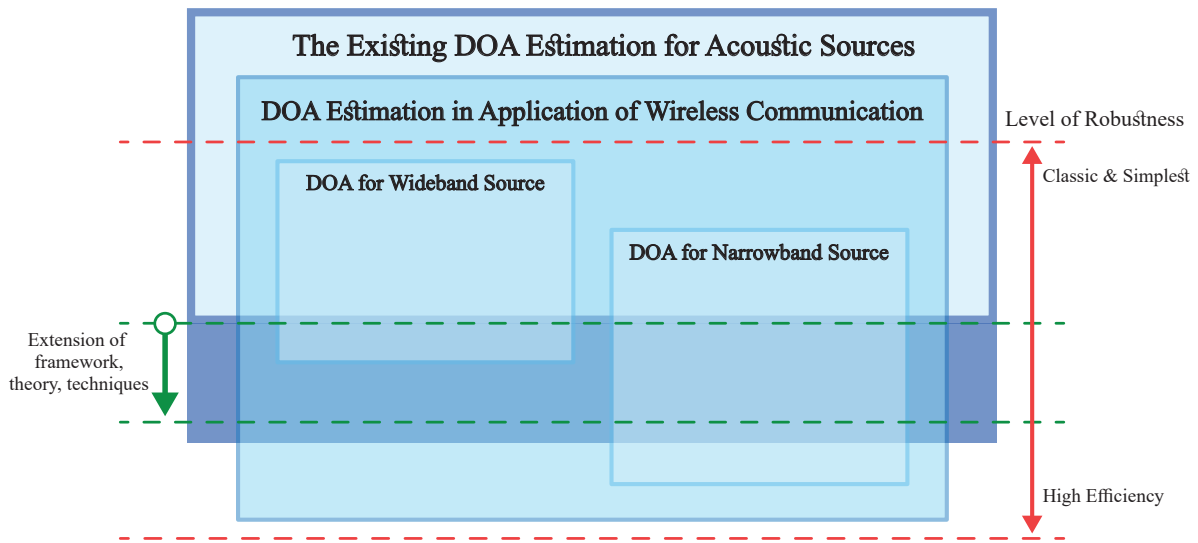


Fig. 1.5 Motivations and objectives of this research works.

cal role in navigation systems for the exploration of sources in widespread applications, especially including in acoustic signal processing recently. DOA estimation for acoustic sources, however, has not been widely investigated, and many significant methods for narrowband DOA estimation cannot be applied to acoustic circumstances directly. Therefore, this dissertation aimed to bridge a research gap of acoustic source compatibility on the recent narrowband and wideband subspace methods to estimate DOA of the acoustic sources directly and effectively. We believe that our findings presented in this dissertation will appeal to researchers who wish to develop a sound source based navigation system and improve its robust estimation. We also hope that this new framework and theory can be a good alternative for estimating DOA of acoustic sources, especially human speeches and musical sounds.

1.5 Dissertation Organization

This dissertation is organized into 7 chapters. Chapter 2 presents an alternative signal modeling for wideband sources (or equivalence to acoustic sources) with L-shaped

1.5 Dissertation Organization

microphone array configuration. The previous approaches focus on a narrowband signal model, which is necessary to consider a wideband source for more practical use, especially in human speeches and musical sounds. The problem of estimating DOA of multi-narrowband signals is addressed and resolved by declaring new terms of temporal and focusing frequencies in array manifold matrices of the new signal model. In addition to this, the proposed model is also compatible with almost recent narrowband DOA estimation methods, it means that this model enables cutting-edge techniques in the existing narrowband subspace methods to apply acoustic source directly and efficiently. Furthermore, performance of the proposed model is evaluated in terms of the root-mean-squared error (RMSE) over a range of the signal-to-noise ratio (SNR) along with the recent DOA estimation methods, and the results demonstrated that the proposed model can efficiently handle wideband sources with acceptable SNR levels.

Chapter 3 presents a novel DOA estimation method for wideband sources by using unsupervised learning. The problem of estimating DOA of wideband signals are addressed by focusing the entire observation subspace in each frequency bin along with employing a multi-narrowband signal model as was previously proposed in Chapter 2. In addition to this observation process, we employ the Gaussian mixture model with a maximum likelihood estimation algorithm as an interpolation scheme, and wideband DOA results are estimated easily by interpolating the multi-narrowband DOA results all frequency bins with this interpolation scheme.

Chapter 4 presents an efficient wideband two-dimensional direction-of-arrival (DOA) estimation for an L-shaped microphone array. We propose a way to construct a wideband sample cross-correlation matrix without any process of DOA preliminary estimation, such as beamforming technique, by exploiting sample cross-correlation matrices of two different frequencies for all frequency bins along with the Orthogonal Procrustes analysis. Subsequently, wideband DOAs can be estimated by using this

1.5 Dissertation Organization

wideband matrix along with a recent scheme of estimating DOA in a narrowband subspace method. Therefore, a contribution of our study is providing an alternative framework for recent narrowband subspace methods to estimating the DOA of wideband sources directly. It means that this framework enables cutting-edge techniques in the existing narrowband subspace methods to implement the wideband direction estimation for reducing the computational complexity and facilitating the estimation algorithm. Theoretical analysis and effectiveness of the proposed method are substantiated through numerical simulations, which are performed in reverberating environments. The results show that performance of the proposed method performs better than others over a range of SNR with just a few microphones. All these advantages make the proposed method a powerful tool for navigation systems based on acoustic signal processing.

After the proposed method in Chapter 4 has been successfully applied to acoustic sources by employing Orthogonal Procrustes analysis. Chapter 5 further extend and propose an efficient framework for estimating DOA of wideband sound sources. We extend a theory of Orthogonal Procrustes analysis for estimating wideband DOAs much more efficiently than the previously. The proposed framework provides an efficient way to construct a wideband cross-correlation matrix from multiple narrowband cross-correlation matrices for all frequency bins via the proposed extension theory of Orthogonal Procrustes analysis. In addition, the proposed framework is inspired by the coherent signal subspace technique with further improvement of linear transformation procedure, and the new procedure no longer require any process of DOA preliminary estimation by exploiting unique cross-correlation matrices between received signal and itself on distinct frequencies, along with the higher-order generalized singular value decomposition (HOGSVD) of the array of this unique matrix. Wideband DOAs are estimated by employing any subspace-based technique for estimating narrowband DOAs, but using the

1.5 Dissertation Organization

proposed wideband correlation instead of the narrowband correlation matrix. It implies that the proposed framework enables cutting-edge researches in the recent narrowband subspace methods to estimate DOA of the wideband sources directly, which result in the reducing computational complexity and facilitating the estimation algorithm. This thus be the major contribution of the proposed framework as same as the contribution on Chapter 4, but much more efficiently than the previously. Practical examples are presented to showcase its applicability and effectiveness, and the experimental results show that the performance of fusion methods perform better than others over a range of SNR with just a few sensors, which make it suitable for practical use.

Unlike all previous methods in Chapters 3 to 5 that have focused on wideband DOA estimation, Chapter 6 proposes a new way to estimate a variance each temporal frequency and DOA of wideband sources simultaneously. Signal model is renovated into a tensor representation of the following features: x-subarray angle, variance-of-frequency, and z-subarray manifold angle. DOAs and variance-of-frequency are now estimated simultaneously by employing a tensor factorization on the proposed signal model. In particular, complex-valued parallel factor analysis is utilized as the tensor factorization. This novel method provides a promising alternative for intelligent source localization.

Finally, Chapter 7 conclude the overall proposed methods. The advantages and discussion of all techniques are given.

Chapter 2

An Alternative Signal Modeling for Wideband Sources with L-Shaped Microphone Array Configuration

This chapter presents an alternative signal modeling for wideband sources with L-shaped microphone array configuration. The previous approaches focus on a small and narrowband of frequency of signal model, which is necessary to consider a wideband source for more practical use. The problem of estimating direction-of-arrival of multi-narrowband signals are addressed and resolved by declaring new terms of temporal and focusing frequencies in array manifold matrices of the new signal model. Additionally, the proposed model is also compatible with almost recent direction-of-arrival estimation methods, such as, the multiple signal classification and the other cross-correlation based techniques. Performance of the proposed model is evaluated in terms of the root-mean-squared error over a range of the signal-to-noise ratio along with the recent direction-of-arrival estimation methods, and the simulation results demonstrated that the proposed

2.1 Reviews on Recent Signal Modeling Techniques

model can efficiently handle wideband sources with acceptable noise levels.

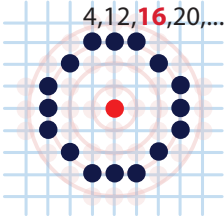
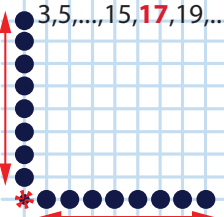
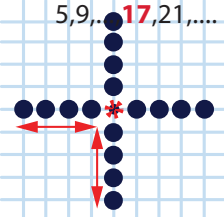
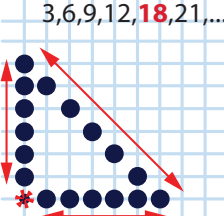
2.1 Reviews on Recent Signal Modeling Techniques

As mentioned in Section 5.1, the fundamental competence of source localization has been extensively implemented in navigation systems for the exploration of sources, which is known as direction-of-arrival (DOA) estimation, and DOA estimation in particular has been widely implemented in many areas [1–6]. The major efficiency drawbacks of the source localization in navigation system are depended on background noise, multiple source locations, and variation in the sound patterns and source frequencies. A number of methods have been introduced to improve the source localization efficiency, including, generalized cross correlation method [7–10], maximum cross-correlation methods [11] and subspace method [12–37, 40, 41]. In addition to this, subspace methods are increasingly utilized for the DOA estimation of multiple sources. In case of the one-dimensional (1D) DOA estimation, the uniform linear array (ULA) structure has been developed, and its effective methods of dealing with 1D DOA estimation can be found in [12–15]. On the contrary, such model is unable to offer simultaneous angles of azimuth and elevation.

Recently, the two-dimensional (2D) DOA estimation using a 2D geometrical structure of sensors has received considerable attention in the array signal processing and widespread applications. Among various array geometrical structures such as the rectangular array, the parallel uniform linear array and circular array, the L-shaped array has attracted a lot of attention. In addition, it was previously found that the structure of L-shaped array is considerably effective for estimating 2D DOAs as shown in Table 2.1, and the literature [18]. Several L-shaped array-based 2D DOA estimation methods have been proposed for estimating 2D DOA of narrowband sources [16–18]. Although

2.1 Reviews on Recent Signal Modeling Techniques

Table 2.1 Cramer-Rao bound of array geometrical structures

Structure	Figure	CRB
Octagon Array (Circle Array)		$\frac{57}{\delta M^3}$
L-Shaped Array		$\frac{60}{\delta M^3}$
Cross Array		$\frac{96}{\delta M^3}$
Triangle Array		$\frac{108}{\delta M^3}$

* $\delta = 2\text{SNR} \left(\frac{2\pi d}{\lambda}\right)^2$, CRB represents the lowest error bound of DOA.

these methods calculate the direction of narrowband sources using noise subspace via eigendecomposition (EVD), which result in a superior multiple sources localization. In particular, these methods also employ s spectrum peak searching algorithm or nonlinear searching algorithm, which result in huge computational costs and not suggestible for real-time applications. Due to its computational complexities, many computationally efficient algorithms for 2D DOA estimation have been proposed in order to overcome this problem. For example, Tayem et al. [20] and Wu et al. [21] initially proposed the

2.1 Reviews on Recent Signal Modeling Techniques

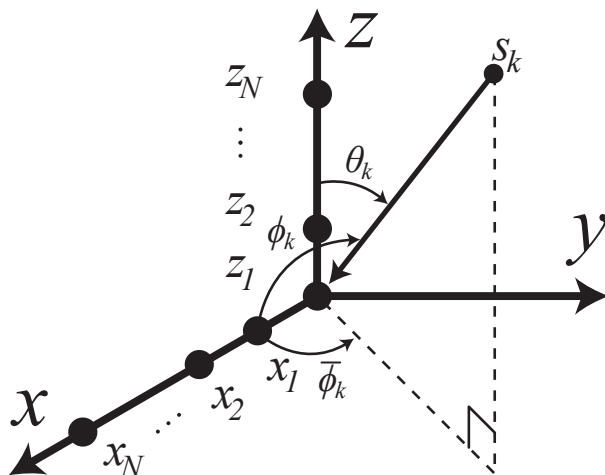


Fig. 2.1 L-shaped microphone array configuration for 2D DOA estimation.

computationally efficient 2D DOA estimation by using the propagator method (PM) in order to avoid the EVD operation; however, the additional angle pair matching algorithm is required for matching azimuth and elevation angles. [19–21]. According to this issue, there are two options to overcome this difficulty: Of course, the first solution is utilizing the angle pair matching algorithm [22]. On the other hand, the automatic pairing 2D DOA estimation were proposed for avoiding this problem [23–35]. According to the recent methods on this issue, only narrowband sources can be used because the previous signal model is based on a narrowband signal.

Therefore, we aim to propose an alternative signal modeling for wideband sources with L-shaped microphone array configuration. The problem of estimating 2D DOA of multi-narrowband signals are addressed and resolved by declaring new terms of temporal and focusing frequencies in array manifold matrices of the new signal model, which the next section will address this matter.

2.2 Phase Difference Equation

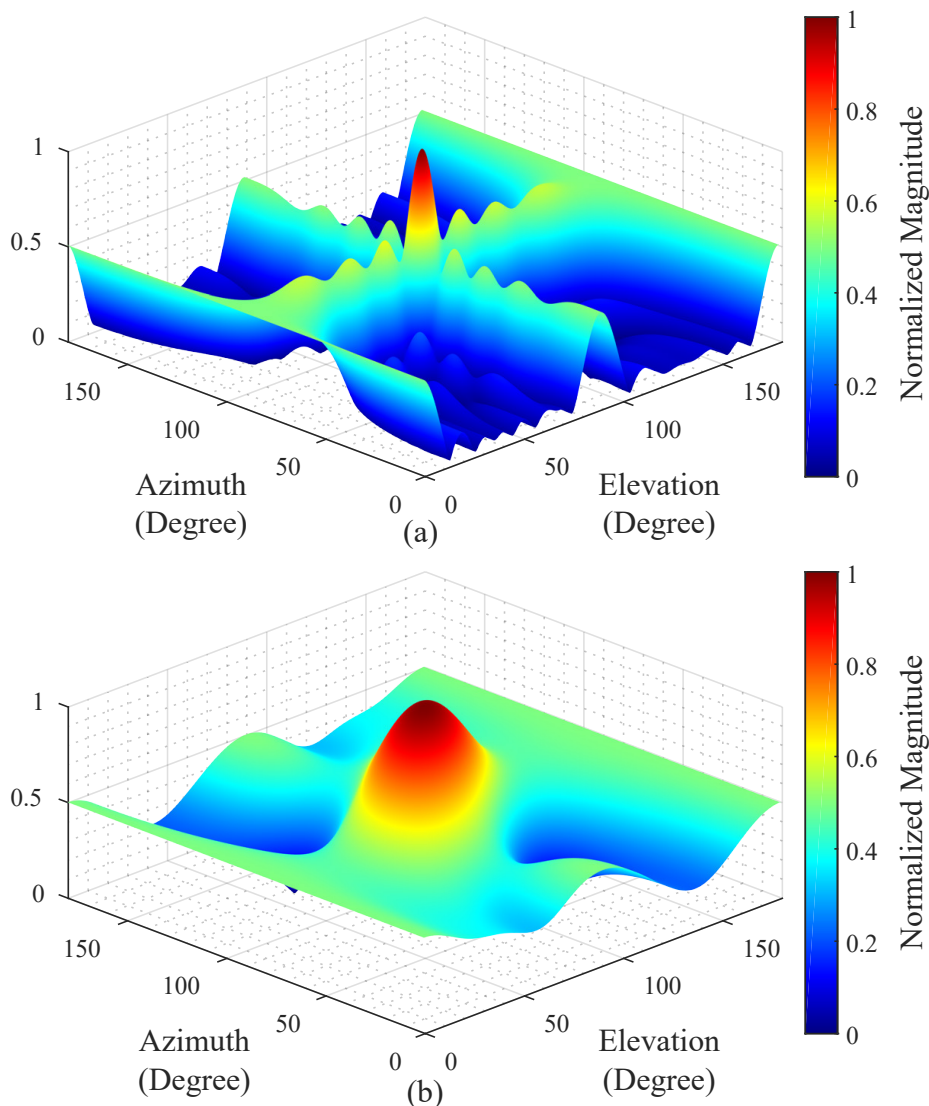


Fig. 2.2 Normalized radiation patterns of the L-shaped arrays respect to DOA angles; (a) $f = f_c$ and (b) $f = f_c/4$ where $M = 8$, $d = \lambda/2$.

2.2 Phase Difference Equation

The traditional difference equation of time-delayed signal is given by Euler's formula as $e^{i2\pi ft}$ where f is source frequency and t is time difference of arrival due to the distance between a source and microphone. Assume that two microphones in x - z plane, the phase difference equation of the x - z plane is defined as

$$q(\phi_k, \theta_k) = e^{\frac{j2\pi f(\epsilon_x \cos \phi_k + \epsilon_z \cos \theta_k)}{c}}, \quad (2.1)$$

2.2 Phase Difference Equation

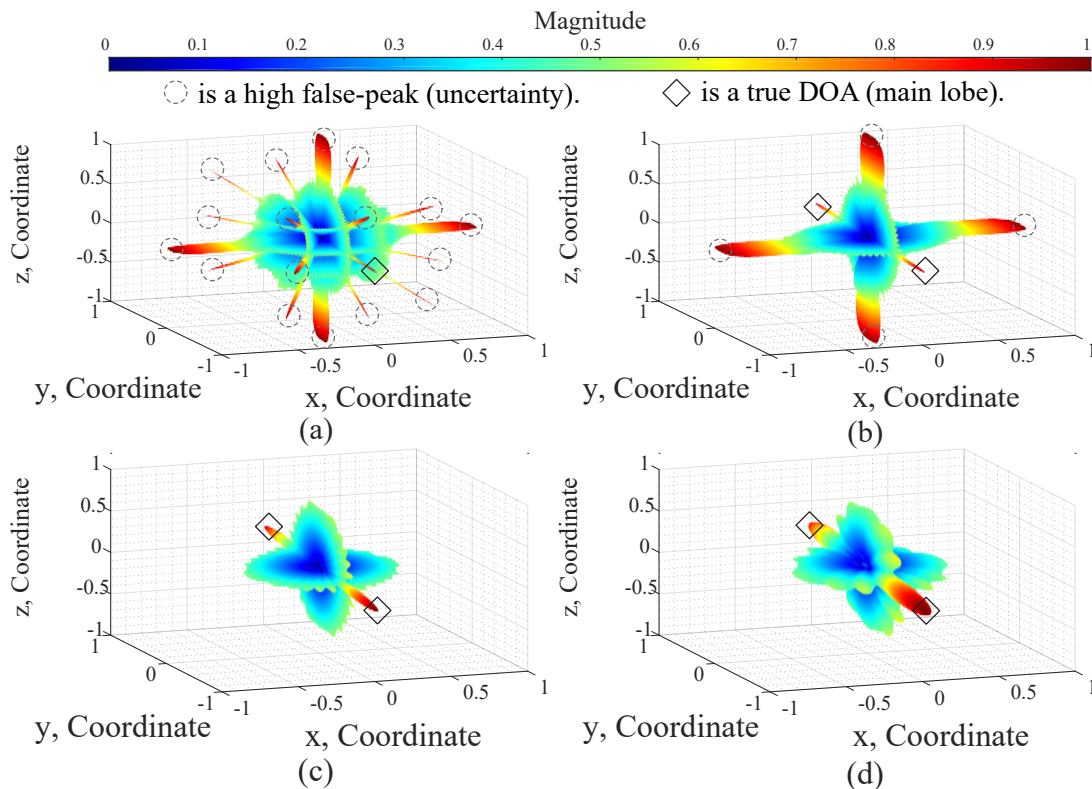


Fig. 2.3 Normalized radiation patterns of the L-shaped arrays respect to coordinates; (a) $f = 4f_o$, (b) $f = 2f_o$, (c) $f = f_o$, (d) $f = \frac{f_o}{2}$ where $M = 12$, $\lambda = \frac{d}{2}$.

where ϕ_k is an angle between x axis and the sources, θ_k is an angle between x axis and the sources, ϵ_x and ϵ_z are spacing of the microphone elements in the x and z axis, and c is the speed of sound. According to the literature [40], the spacing of the microphone elements in any axis should be set to $\frac{\lambda}{2}$ where $\lambda = c/f_c$ is the wavelength and f_c is frequency corresponding to the wavelength. The reason is that to avoid angle confusion caused by grating lobes and the lower power beam-width of main lobe in radiation pattern of array structure.

As the fact that the power beam-width of the grating lobes in the radiation patterns depends on f and f_c , the phase difference function can be redefined as

$$q(\phi_k, \theta_k) = e^{\left(\frac{f}{f_c}\right) \cdot \left(\frac{j2\pi(\epsilon_x \cos \phi_k + \epsilon_z \cos \theta_k)}{\lambda}\right)}. \quad (2.2)$$

Consider the L-shaped array as illustrated in Fig. 2.1, where it consists of two orthogonal

2.3 The New Data Model

M -element ULAs in the x - z plane. The spacing of the microphone elements and the spacing between z_0 and x_1 is also d . The reference point is pinned at the origin of coordinates z_1 . Form Eq. (2.2), the magnitude of the L-shaped array factor represents a radiation pattern and correspond to the following equation;

$$AF = \frac{1}{2M} \left| \sum_{m=1}^M \alpha_m(\phi_k) + \sum_{m=1}^M \beta_m(\theta_k) \right|, \quad (2.3)$$

where

$$\begin{aligned} \boldsymbol{\alpha}(\phi_k) &= \left[1 \quad e^{j\alpha_k} \quad \dots \quad e^{j(M-1)\alpha_k} \right]^T, \\ \boldsymbol{\beta}(\theta_k) &= \left[e^{j\beta_k} \quad e^{j2\beta_k} \quad \dots \quad e^{jM\beta_k} \right]^T, \\ \alpha_k &= \left(\frac{f}{f_c} \right) \cdot \left(\frac{2\pi d \cos \phi_k}{\lambda} \right), \\ \beta_k &= \left(\frac{f}{f_c} \right) \cdot \left(\frac{2\pi d \cos \theta_k}{\lambda} \right). \end{aligned} \quad (2.4)$$

It can be seen from Fig. 2.2 that power beam-width of the grating lobes in Eq. (2.3) is apparently affected by f/f_c . As the ratio of f and f_c decreases from 1 to 1/4, the overall power beam-width decreases significantly; therefore, it suggests that range of source frequency f should be selected to be less than or equal center frequency f_c . In order to avoid the confusion when lower source frequency is chosen, the number of microphone elements M should be increased as well as illustrated in Fig. 2.3.

2.3 The New Data Model

Consider P far-field multi-narrowband sources impinging on the L-shaped arrays as shown in the previous section, the output of each microphone is sampled and decomposed into n snapshots by a short-time Fourier transform (STFT), with each frame containing a set of frequencies f . Therefore, the received signal vector vectors in x and z subarrays are represented as follows:

$$\mathbf{X}(n, f) = \mathbf{A}_x(\phi) \mathbf{S}(n, f) + \mathbf{W}_x(n, f), \quad (2.5)$$

2.3 The New Data Model

$$\mathbf{Z}(n, f) = \mathbf{A}_z(\boldsymbol{\theta}) \mathbf{S}(n, f) + \mathbf{W}_z(n, f), \quad (2.6)$$

$$\mathbf{A}_x(\boldsymbol{\phi}) = \begin{bmatrix} \boldsymbol{\alpha}(\phi_1) & \boldsymbol{\alpha}(\phi_2) & \dots & \boldsymbol{\alpha}(\phi_P) \end{bmatrix}, \quad (2.7)$$

$$\mathbf{A}_z(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\beta}(\theta_1) & \boldsymbol{\beta}(\theta_2) & \dots & \boldsymbol{\beta}(\theta_P) \end{bmatrix}, \quad (2.8)$$

where the remaining components in Eqs. (2.7) and (2.8) are

$$\begin{aligned} \mathbf{X}(n, f) &= \begin{bmatrix} x_1(n, f) & x_2(n, f) & \dots & x_M(n, f) \end{bmatrix}^T, \\ \mathbf{Z}(n, f) &= \begin{bmatrix} z_0(n, f) & z_1(n, f) & \dots & z_{M-1}(n, f) \end{bmatrix}^T, \\ \mathbf{S}(n, f) &= \begin{bmatrix} s_1(n, f) & s_2(n, f) & \dots & s_P(n, f) \end{bmatrix}^T, \\ \mathbf{W}_x(n, f) &= \begin{bmatrix} w_{x_1}(n, f) & w_{x_2}(n, f) & \dots & w_{x_M}(n, f) \end{bmatrix}^T, \\ \mathbf{W}_z(n, f) &= \begin{bmatrix} w_{z_0}(n, f) & w_{z_1}(n, f) & \dots & w_{z_{M-1}}(n, f) \end{bmatrix}^T. \end{aligned} \quad (2.9)$$

Note that $\mathbf{X}(n, f)$ and $\mathbf{Z}(n, f)$ represent sum of received signals in x and z subarrays for all P sources with different angles. $\mathbf{S}(n, f)$ represents source signals. $\mathbf{W}_x(n, f)$ and $\mathbf{W}_z(n, f)$ are an additive noise vectors in x and z subarrays for all microphone elements. $\mathbf{A}_x(\boldsymbol{\phi})$ and $\mathbf{A}_z(\boldsymbol{\theta})$ are the proposed array manifold matrices of x and z subarrays. n is a snapshot and f is source frequency. The relationship between the azimuth angle $\bar{\phi}$, elevation angle θ and x subarray angle ϕ is expressed by Euler's formula as follows;

$$\cos \phi_k = \sin \theta_k \cos \bar{\phi}_k. \quad (2.10)$$

Now, the proposed data model is able to represent the source signals on wideband frequency bands because of Eq. (2.4) on Eqs. (2.7) and (2.8). We can use the proposed data model for estimating multi-narrowband DOAs by employing the recent narrowband DOA estimation, which the next section will address further.

2.4 Example of DOA Estimation Schemes with the New Data Model

In order to verify those Eqs. (2.7) and (2.8) are compatible with most classical subspace-based methods, 2D MUSIC method [16], the Wang et al. [34] and Dong et al. [31] method are chosen and we replace the conventional array manifold matrix to the proposed array manifold matrix, which explains in the next three subsections.

2.4.1 Extension of 2D MUSIC Method

Under the assumptions of data model from [34, 40, 41] and assuming noise free environment, the cross-correlation matrix from the received signals in Eqs. (2.5) and (2.6) can be described as follows:

$$\begin{aligned}
 \mathbf{R}_{yy} &= E \left[\begin{bmatrix} \mathbf{X}(n, f) \\ \mathbf{Z}(n, f) \end{bmatrix} \begin{bmatrix} \mathbf{X}(n, f) \\ \mathbf{Z}(n, f) \end{bmatrix}^H \right] \\
 &= \begin{bmatrix} \mathbf{A}_x(\phi) \\ \mathbf{A}_z(\theta) \end{bmatrix} E \left[\mathbf{S}(n, f) \mathbf{S}(n, f)^H \right] \begin{bmatrix} \mathbf{A}_x(\phi) \\ \mathbf{A}_z(\theta) \end{bmatrix}^H \\
 &= \begin{bmatrix} \mathbf{A}_x(\phi) \\ \mathbf{A}_z(\theta) \end{bmatrix} \mathbf{R}_{ss} \begin{bmatrix} \mathbf{A}_x(\phi) \\ \mathbf{A}_z(\theta) \end{bmatrix}^H
 \end{aligned} \tag{2.11}$$

where $E[\cdot]$ is the statistical expectation, \mathbf{R}_{ss} is the correlation matrix of sources $\mathbf{S}(n, f)$. Applying EVD, \mathbf{R}_{yy} can be represented as the factorization of matrix into a canonical form as follows:

$$\mathbf{R}_{yy} = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H + \mathbf{U}_w \mathbf{\Lambda}_w \mathbf{U}_w^H \tag{2.12}$$

where \mathbf{U}_s and $\mathbf{\Lambda}_s$ are the matrix of eigenvector and eigenvalue in the signal subspace. \mathbf{U}_w and $\mathbf{\Lambda}_w$ are the matrix of eigenvector and eigenvalue in the noise subspace.

After the matrix \mathbf{U}_w is estimated, we can obtain ϕ_k and θ_k by maximizing the following cost function with Eq. (2.4) as follows:

$$\left\{ \hat{\phi}_k, \hat{\theta}_k \right\}_{k=1}^P = \arg \max_{\phi, \theta} (\ln |f_{\text{MUSIC}}(\phi, \theta)|) \tag{2.13}$$

2.4 Example of DOA Estimation Schemes with the New Data Model

where

$$f_{\text{MUSIC}}(\phi, \theta) = \frac{1}{\begin{bmatrix} \boldsymbol{\alpha}(\phi) \\ \boldsymbol{\beta}(\theta) \end{bmatrix}^H \mathbf{U}_w \mathbf{U}_w^H \begin{bmatrix} \boldsymbol{\alpha}(\phi) \\ \boldsymbol{\beta}(\theta) \end{bmatrix}}. \quad (2.14)$$

2.4.2 Extension of Wang et al. Method

The computationally efficient narrowband subspace-based 2D DOA estimation method without EVD was proposed by Wang et al [34]. This method begin by constructing the cross-correlation matrix \mathbf{R}_{yy} which already described in Eq. (2.11). In summary, the modified cost function with Eq. (2.4) are given by:

$$\left\{ \hat{\phi}_k, \hat{\theta}_k \right\}_{k=1}^P = \arg \min_{\phi, \theta} \left(\begin{bmatrix} \boldsymbol{\alpha}(\phi) \\ \boldsymbol{\beta}(\theta) \end{bmatrix}^H \hat{\Pi} \begin{bmatrix} \boldsymbol{\alpha}(\phi) \\ \boldsymbol{\beta}(\theta) \end{bmatrix} \right), \quad (2.15)$$

or

$$\left\{ \hat{\phi}_k, \hat{\theta}_k \right\}_{k=1}^P = \arg \max_{\phi, \theta} \left(- \begin{bmatrix} \boldsymbol{\alpha}(\phi) \\ \boldsymbol{\beta}(\theta) \end{bmatrix}^H \hat{\Pi} \begin{bmatrix} \boldsymbol{\alpha}(\phi) \\ \boldsymbol{\beta}(\theta) \end{bmatrix} \right), \quad (2.16)$$

where

$$\begin{aligned} \hat{\Pi} &= \hat{\mathbf{Q}} \left(\mathbf{I}_{2M-P} - \hat{\mathbf{P}}^H \left(\hat{\mathbf{P}} \hat{\mathbf{P}}^H + \mathbf{I}_P \right)^{-1} \hat{\mathbf{P}} \right) \hat{\mathbf{Q}}^H, \\ \hat{\mathbf{P}} &= \left(\hat{\mathbf{G}}_1 \hat{\mathbf{G}}_1^H \right)^{-1} \hat{\mathbf{G}}_1 \hat{\mathbf{G}}_2^H, \\ \hat{\mathbf{Q}} &= \begin{bmatrix} \hat{\mathbf{P}}^T & -\mathbf{I}_{2M-P} \end{bmatrix}^T, \\ \mathbf{R}_{yy} &= \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix}. \end{aligned} \quad (2.17)$$

Note that size of matrices \mathbf{G}_1 and \mathbf{G}_2 are $2M \times P$ and $2M \times 2M - P$, respectively.

2.4.3 Extension of Dong et al. method

To estimates the DOA angles without employing 2D peak finding algorithm, Dong et al. proposed the efficient method with only employing 1D peak finding algorithm [31, 32]. This method begin by constructing the cross-correlation matrix \mathbf{R}_{xz} between

2.4 Example of DOA Estimation Schemes with the New Data Model

x and z subarrays as follows:

$$\begin{aligned}\mathbf{R}_{xz} &= E \left[\mathbf{X}(n, f) \mathbf{Z}(n, f)^H \right] \\ &= \mathbf{A}_x(\phi) E \left[\mathbf{S}(n, f) \mathbf{S}(n, f)^H \right] \mathbf{A}_z^H(\theta) \\ &= \mathbf{A}_x(\phi) \mathbf{R}_{ss} \mathbf{A}_z^H(\theta).\end{aligned}\quad (2.18)$$

As the conjugate system property possesses in the proposed array manifold matrix Eqs. (2.7) and (2.8), we can have:

$$\mathbf{J}_M(\mathbf{A}_x(\phi))^* = \mathbf{A}_x(\phi) \tilde{\Phi}_{xr}(\phi), \quad (2.19)$$

$$\mathbf{J}_M(\mathbf{A}_z(\theta))^* = \mathbf{A}_z(\theta) \tilde{\Phi}_{zr}(\theta), \quad (2.20)$$

$$\mathbf{A}_{z2}(\theta) = \mathbf{A}_{z1}(\theta) \Phi_z(\theta), \quad (2.21)$$

where

$$\begin{aligned}\tilde{\Phi}_{xr}(\phi) &= \text{diag} \left(\left[e^{-j(M-1)\alpha_1} \quad \dots \quad e^{-j(M-1)\alpha_P} \right] \right), \\ \tilde{\Phi}_{zr}(\theta) &= \text{diag} \left(\left[e^{-j(M-1)\beta_1} \quad \dots \quad e^{-j(M-1)\beta_P} \right] \right), \\ \Phi_z(\theta) &= \text{diag} \left(\left[e^{j\beta_1} \quad e^{j\beta_2} \quad \dots \quad e^{j\beta_P} \right] \right),\end{aligned}\quad (2.22)$$

$\mathbf{A}_{z1}(\theta)$ and $\mathbf{A}_{z2}(\theta)$ are the first and the last $M - 1$ rows of $\mathbf{A}_z(\theta)$. Then, the new cross-correlation matrix can be reconstructed with the similar scheme on [25, 31], as follows:

$$\begin{aligned}\mathbf{Y} &= \begin{bmatrix} \mathbf{Y}_1, \mathbf{J}_M \mathbf{Y}_2^* \\ \mathbf{Y}_2, \mathbf{J}_M \mathbf{Y}_1^* \end{bmatrix} \\ &= \mathbf{A}_g(\phi, \theta) \mathbf{S}_g(\phi, \theta)\end{aligned}\quad (2.23)$$

where

$$\begin{aligned}\mathbf{A}_g(\phi, \theta) &= \begin{bmatrix} \mathbf{A}_x^T(\phi) & (\mathbf{A}_x(\phi) \tilde{\Phi}_z^*(\theta))^T \end{bmatrix}^T, \\ \mathbf{S}_g(\phi, \theta) &= \begin{bmatrix} \mathbf{R}_{ss} \mathbf{A}_{z1}^H(\theta) & \Phi_z(\theta) \tilde{\Phi}_{xr}(\phi) \mathbf{R}_{ss} \mathbf{A}_{z1}^T(\theta) \end{bmatrix}, \\ \mathbf{Y}_1 &= \mathbf{A}_x(\phi) \mathbf{R}_{ss} \mathbf{A}_{z1}^H(\theta) = \mathbf{R}_{xz}(:, 1 : M - 1), \\ \mathbf{Y}_2 &= \mathbf{A}_x(\phi) \mathbf{R}_{ss} \mathbf{A}_{z2}^H(\theta) = \mathbf{R}_{xz}(:, 2 : M).\end{aligned}\quad (2.24)$$

In the same way of 2D MUSIC method, applying EVD to $\mathbf{R}_{\mathbf{Y}\mathbf{Y}} = \mathbf{Y}\mathbf{Y}^H$, the matrix $\mathbf{R}_{\mathbf{Y}\mathbf{Y}}$ can be factorized as summation of two matrices, including, the matrix of eigenvector and eigenvalue in the signal subspace and noise subspace. After the noise subspace

2.5 Numerical Simulation

matrices is obtained, the new cost function for estimating multi-narrowband DOAs can be expressed as follows:

$$\left\{ \hat{\phi}_k, \hat{\theta}_k \right\}_{k=1}^P = \arg \min_{\phi, \theta} \left(\frac{\mathbf{q}^H(\theta) \mathbf{F}(\phi) \mathbf{q}(\theta)}{\mathbf{q}^H(\theta) \mathbf{q}(\theta)} \right), \quad (2.25)$$

or

$$\left\{ \hat{\phi}_k, \hat{\theta}_k \right\}_{k=1}^P = \arg \max_{\phi, \theta} \left(-\frac{\mathbf{q}^H(\theta) \mathbf{F}(\phi) \mathbf{q}(\theta)}{\mathbf{q}^H(\theta) \mathbf{q}(\theta)} \right), \quad (2.26)$$

where

$$\begin{aligned} \mathbf{q}(\theta) &= \begin{bmatrix} 1 & e^{-j\beta} \end{bmatrix}^T, \\ \mathbf{F}(\phi) &= (\mathbf{I}_2 \otimes \boldsymbol{\alpha}(\phi))^H \mathbf{U}_{w, \mathbf{R}_{YY}} \mathbf{U}_{w, \mathbf{R}_{YY}}^H (\mathbf{I}_2 \otimes \boldsymbol{\alpha}(\phi)), \end{aligned}$$

\otimes is Kronecker product. In order to solve Eq. (2.26) without using 2D search, Dong et al. described the relationship between the eigenvalue and its determinant which can be rewritten as 1D search optimization problem as follows:

$$\hat{\phi}_k = \arg \max_{\phi} \left(\frac{1}{\mathbf{F}(\phi)} \right). \quad (2.27)$$

In the end, $\hat{\theta}_k$ can be estimated via the eigenvector corresponding to the minimum eigenvalue of $\mathbf{F}(\hat{\phi}_k)$;

$$\hat{\theta}_k = \arccos \left(\text{angle} \left(\frac{\mathbf{e}_{\min}^1}{\mathbf{e}_{\min}^2} \right) \cdot \left(\frac{\lambda}{2\pi d} \right) \cdot \left(\frac{f_c}{f} \right) \right). \quad (2.28)$$

where \mathbf{e}_{\min}^1 and \mathbf{e}_{\min}^2 are the first and second elements of the eigenvector corresponding the minimal eigenvalue of $\mathbf{F}(\hat{\phi}_k)$. Those Eqs. (2.27) and (2.28) showed that method can pair the angles automatically without using 2D search.

2.5 Numerical Simulation

In this section, the performance of the proposed array manifold matrix was demonstrated with three types of cost functions; 2D MUSIC method [16], the Wang et al. method [34], and Dong et al. method [31]. We merged the proposed array manifold

2.5 Numerical Simulation

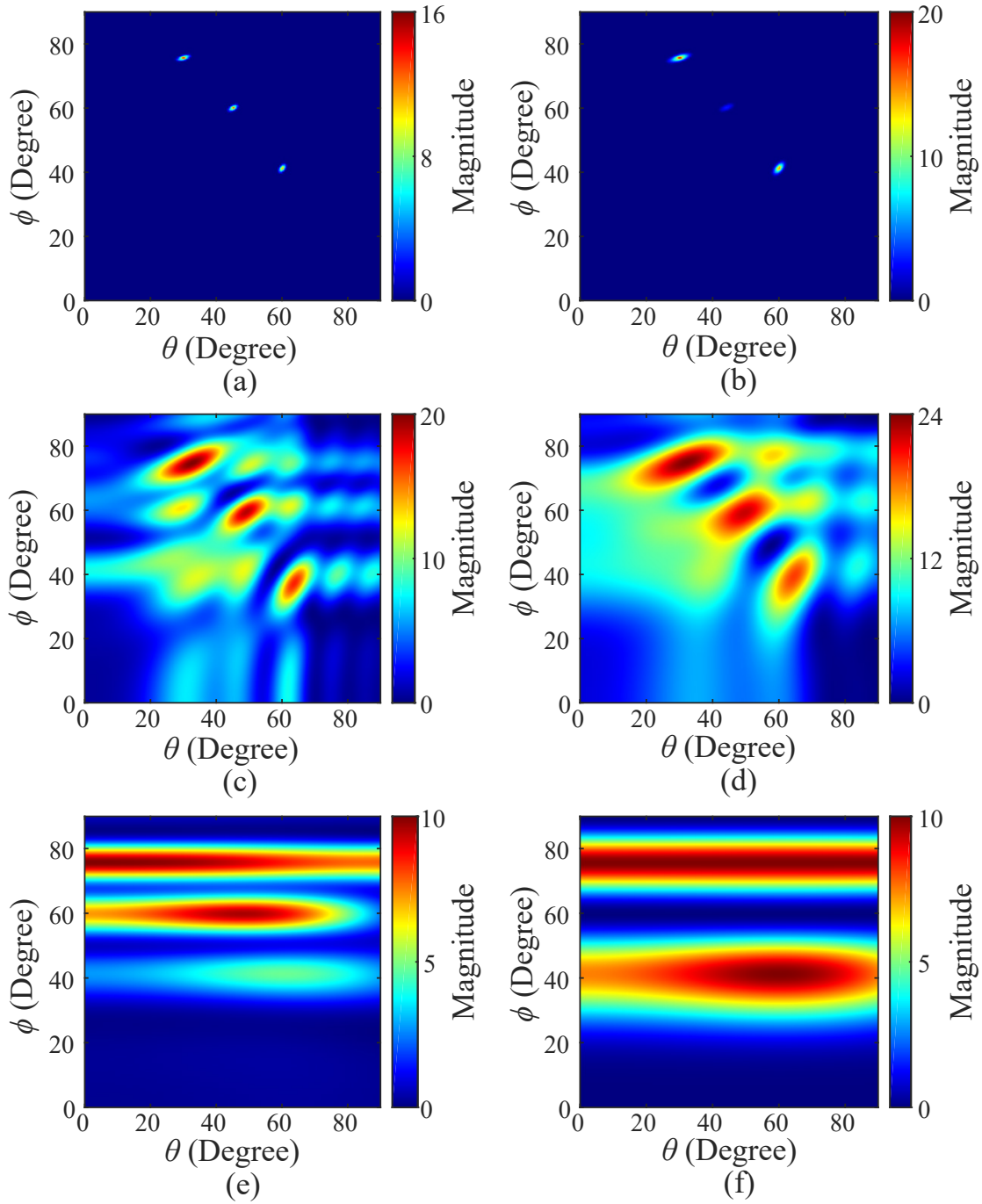


Fig. 2.4 Search space results of three types of cost functions along with the proposed array manifold matrices; (a), (b) 2D MUSIC in Eq. (2.13); (c), (d) the Wang et al. method in Eq. (2.16); (e), (f) Dong et al. method in Eq. (2.26), where source frequency in (a), (c), (e) are 3.4 kHz and (b), (d), (f) are 2.26 kHz.

2.5 Numerical Simulation

matrix with the selected methods for estimating multi-narrowband DOAs as shown in the previous section. All the method was tested by computer simulation and implemented in MATLAB software. The output of each microphone was decomposed into 192,000 snapshots. The sampling frequency was 192 kHz, the center frequency was 3.4 kHz, the number of subarray elements M was 12, the spacing of microphone elements was 5 cm and speed of sound was assumed to be 340 m/s. The 2D search ranges for azimuth and elevation are $[0^\circ, 180^\circ]$ and $[0^\circ, 180^\circ]$ with interval 0.01. The definition of root-mean-square error (RMSE) of the 2D DOA estimation was expressed as

$$\text{RMSE} = \sqrt{\frac{1}{LP} \sum_{l=1}^L \sum_{k=1}^P \left(\left(\bar{\phi}_k^{\text{DOA}} - \bar{\phi}_k^{(l)} \right)^2 + \left(\theta_k^{\text{DOA}} - \theta_k^{(l)} \right)^2 \right)}, \quad (2.29)$$

where L is the number of times of the Monte Carlo experiment. $\bar{\phi}_k^{(l)}$ and $\theta_k^{(l)}$ are the l^{th} estimated azimuth and elevation angles of the k^{th} . $\bar{\phi}_k^{\text{DOA}}$ and θ_k^{DOA} are the true DOA azimuth and elevation angles of the k^{th} .

2.5.1 Robustness of the Cost Functions

Fig. 2.4 showed example of estimation results by employing the three types of cost functions; 2D MUSIC method in Eq. (2.13) as showed in Fig. 2.4 (a), (b). Wang et al. method in Eq. (2.16) as showed in Fig. 2.4 (c), (d). Dong et al. method in Eq. (2.26) as showed in Fig. 2.4 (e), (f). Note that the horizontal axis represents a elevation angle, and the vertical axis represents a azimuth angle. The relationship between the azimuth angle $\bar{\phi}$, elevation angle θ and ϕ is given by Eq. (2.10). Three sources were placed $(\theta_k^{\text{DOA}}, \phi_k^{\text{DOA}})$ at $(30^\circ, 75.52^\circ)$, $(45^\circ, 60^\circ)$ and $(60^\circ, 41.41^\circ)$, where three sources are human voices with containing source frequency range from 100 to 16,000 Hz. Two simulation scenarios were considered. First scenario, Fig. 2.4 (a), (c), (e) are conducted with source frequency f is 3.4 kHz. Second scenario, Fig. 2.4 (b), (d), (f) are conducted with source frequency f is 2.26 kHz.

2.5 Numerical Simulation

As can be seen from Fig. 2.4 (b) and (d) that the proposed array manifold matrix with Eqs. (2.13) and (2.16) can handle the source frequency lower than the reference frequency f_c effectively. Additionally, the surface of cost functions in Fig. 2.4 (d) and (e) spread out around true DOA areas, which may be led to an ineffective DOA estimation. On the one hand, the surface of cost function Eq. (2.16) is influenced by the asymptotic mean-square-error expressions [34], which is shown in Fig. 2.4 (c) and (d) because the computationally intensive and EVD process is no longer used in Wang et al. method. On the other hand, the surface of cost functions of 2D MUSIC method spread out less than Wang et al. method due to utilizing EVD in Eqs. (2.12) and (2.13). In case of Dong et al. method, it can be seen from Eq. (2.26) that surface of cost function is dominated by the Rayleigh quotient problem [31]. However, all confusions can avoid by increasing the number of microphone elements M and employs the constraint of $0 < f \leq f_c$ as shown in Eq. (2.3).

2.5.2 RMSE versus SNR

Three methods were tested for comparison with the proposed array manifold matrix: 2D MUSIC method in Eq. (2.13) as illustrated in Fig. 6.3 (a). Wang et al. method in Eq. (2.16) as illustrated in Fig. 6.3 (b) Dong et al. method in Eqs. (2.27) and (2.28) as illustrated in Fig. 6.3 (c). Note that the horizontal axis represents the signal-to-noise ratio (SNR). This SNR varies from -10 to 40 dB with interval 2 dB. The vertical axis represents RMSE, which was calculated by Eq. (5.76) where 100 Monte Carlo trials have been conducted for every fixed SNR. Three sources were placed $(\theta_k^{\text{DOA}}, \bar{\phi}_k^{\text{DOA}})$ at $(60^\circ, 30^\circ)$, $(45^\circ, 45^\circ)$ and $(30^\circ, 60^\circ)$, where the source frequencies were 1.6, 2.8 and 3.4 kHz, respectively. To estimate the azimuth angle $\bar{\phi}$ and elevation angle θ in each source, multi-narrowband DOA estimates are considered, where the number of sources P was assumed to be 1 in each narrowband. Simulation results in Fig. 6.3 indicated that

2.6 Conclusions

the performance of all the method is significantly dominated by the source frequencies. As the source frequency is 2.8 kHz, all the methods exhibits good performance at SNR range from 0 to 40 dB. Likewise, in case of the source frequency is 1.6 kHz, all the methods exhibits good performance at SNR range from 8 to 40 dB. This problem cause by the interfered power beam-width of the side lobe in L-array structure radiation pattern as discussed in Figs. 2.2 and 2.3. In the end, DOA estimation for multi-narrowband sources can be done by implementing the proposed data model along with the existed narrowband DOA estimations, and the results showed that the fusion methods are able to estimate multi-narrowband DOA sources with acceptable SNRs.

2.6 Conclusions

This chapter presented an alternative signal modeling for wideband sources with L-shaped microphone array configuration. We addressed an efficient way for estimating multi-narrowband DOAs by using the proposed signal model. The problem of estimating multi-narrowband DOAs is resolved by using the proposed signal model along with remodeling the array manifold matrices. Extension of most classical subspace-based methods for multi-narrowband DOA estimation were described and given by employing 2D MUSIC method [16], the Wang et al. method [34], and Dong et al. method [31]. Simulation results indicated that the new array manifold matrices by using the proposed signal model along with the classical subspace-based methods have achieved the multi-narrowband DOA estimation with acceptable noise levels. The modified method provides a potential alternative to intelligent local positioning systems in voice control applications.

2.7 References

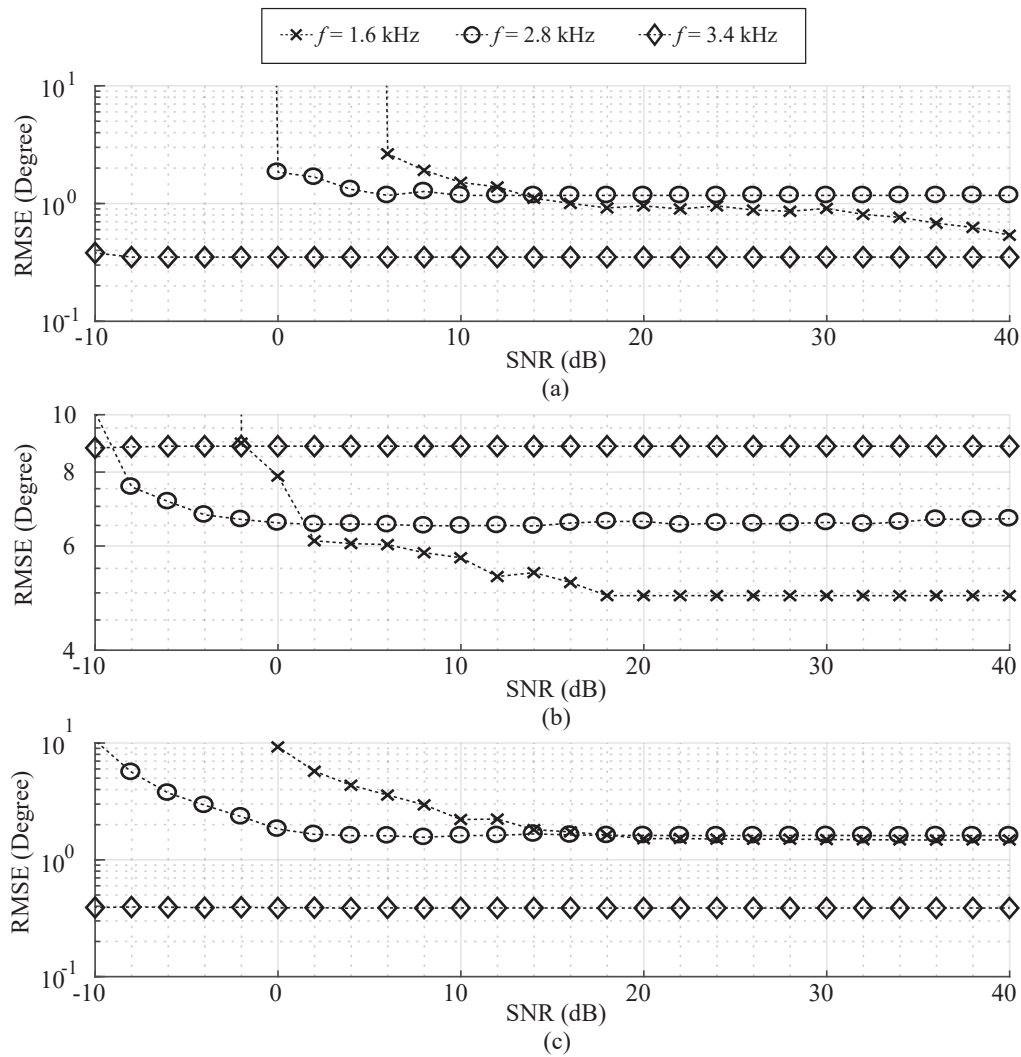


Fig. 2.5 RMSE estimation performance of DOA cost functions along with the proposed array manifold matrices; (a) 2D MUSIC method in Eq. (2.13), (b) the Wang et al. method in Eq. (2.16) and (c) Dong et al. method in Eq. (2.26). Three sources are placed $(\theta_k^{\text{DOA}}, \phi_k^{\text{DOA}})$ at $(60^\circ, 30^\circ)$, $(45^\circ, 45^\circ)$ and $(30^\circ, 60^\circ)$.

2.7 References

- [1] V. R. Reddy, P. Deshpande, and R. Dasgupta, “Robotics Audition using Kinect,” *6th International Conference on Automation, Robotics and Applications (ICARA)*, February 2015.
- [2] M. W. Hansen, J. R. Jensen, and M. G. Christensen, “Localizing Near and Far Field

2.7 References

- Acoustic Sources with Distributed Microhone Arrays,” *48th Asilomar Conference on Signals, Systems and Computers*, November 2014.
- [3] Y. Sasaki, S. Masunaga, S. Thompson, S. Kagami, and H. Mizoguchi, “Sound Localization and Separation for Mobile Robot Tele-Operation by Tri-Concentric Microphone Array,” *Digital Human Symposium 2009*, March 2009.
- [4] J. M. Valinan, F. Michaudb, and J. Rouatb, “Robust Localization and Tracking of Simultaneous Moving Sound Sources using Beamforming and Particle Filtering,” *Robotics and Autonomous Systems*, vol. 55, no. 3, pp. 216–228, March 2007.
- [5] T. Nishiura, M. Nakamura, A. Lee, H. Saruwatari, and K. Shikano, “Talker Tracking Display On Autonomous Mobile Robot With A Moving Microphone Array,” *8th International Conference on Auditory Display*, July 2002.
- [6] J. Huang, T. Supaongprapa, I. Terakura, F. Wang, N. Ohnishi, and N. Sugie, “A Model-Based Sound Localization System and its Application to Robot Navigation,” *Robotics and Autonomous Systems*, vol. 27, no. 4, pp. 199-209, June 1999.
- [7] M. Imran, A. Hussain, N. M. Qazi, and M. Sadiq, “A Methodology for Sound Source Localization and Tracking: Development of 3D Microphone Array for Near-Field and Far-Field Applications,” *13th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, January 2015.
- [8] X. Xiao, S. Zhao, X. Zhong, D. L. Jones, E. S. Chng, and H. Li, “A Learning-Based Approach to Direction of Arrival Estimation in Noisy and Reverberant Environments,” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015.
- [9] I. L. Freire, and J. A. Apolinario, “DoA of Gunshot Signals in a Spatial Microphone Array: Performance of the Interpolated Generalized Cross-Correlation Method,” *Argentine School of Micro-Nanoelectronics Technology and Applications (EAMTA)*, August 2011.

2.7 References

- [10] C. Knapp, and G. Carter, “The Generalized Correlation Method for Estimation of Time Delay,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320-327, August 1976.
- [11] H. T. Goi, and D. K. Sung, “Maximum Cross-Correlation Method for DOA Estimation with an ESPAR Antenna,” *IEEE 60th Vehicular Technology Conference 2004*, September 2004.
- [12] D. S. Yoo, “Subspace-based DOA Estimation with Sliding Signal-Vector Construction for ULA,” *IET Electronics Letters*, vol. 51, no. 17, pp. 1361-1363, August 2015.
- [13] I. L. Freire, and J. A. Apolinario, “Direction of Arrival Estimation Apparatus for Communication based Train Control System using ESPAR Antenna,” *Third International Conference on Digital Information, Networking, and Wireless Communications (DINWC)*, February 2015.
- [14] S. Marcos, A. Marsal, and M. Benidir, “The Propagator Method for Source Bearing Estimation,” *Signal Processing*, vol. 42, no. 2, pp. 121-138, March 1995.
- [15] R. Schmidt, “Multiple Emitter Location and Signal Parameter Estimation,” *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276-280, March 1986.
- [16] M. G. Porozantidou, and M. T. Chryssomallis, “Azimuth and Elevation Angles Estimation Using 2-D MUSIC Algorithm with an L-Shape Antenna,” *IEEE Antennas and Propagation Society International Symposium (APSURSI)*, July 2010.
- [17] J. E. F. Rio, and M. F. C. Perez, “The Matrix Pencil Method for Two-Dimensional Direction of Arrival Estimation Employing an L-Shaped Array,” *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 11, pp. 1693-1694, November 1997.
- [18] Y. Hua, T. K. Sarkar, and D. D. Weiner, “An L-shaped Array for Estimating 2-D Directions of Wave Arrival,” *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 2, pp. 143-146, February 1991.
- [19] N. Tayem, K. Majeed, and A. A. Hussain, “Two-Dimensional DOA Estimation

2.7 References

- Using Cross-Correlation Matrix With L-Shaped Array,” *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1077-1080, October 2016.
- [20] N. Tayem, and H. M. Kwon, “L-Shape 2-Dimensional Arrival Angle Estimation with Propagator Method,” *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 5, pp. 1622-1630, May 2005.
- [21] Y. Wu, G. Liao, and H. C. So, “A Fast Algorithm for 2-D Direction-of-Arrival Estimation,” *Signal Processing*, vol. 83, no. 8, pp. 1827-1831, August 2013.
- [22] S. Kikuchi, H. Tsuji, and A. Sano, “Pair-Matching Method for Estimating 2-D Angle of Arrival With a Cross-Correlation Matrix,” *IEEE Antennas and Wireless Propagation Letters*, vol. 5, no. 1, pp. 1622-1630, December 2006.
- [23] X. Nie, and P. Wei, “Array Aperture Extension Algorithm for 2-D DOA Estimation with L-Shaped Array,” *Progress In Electromagnetics Research Letters*, vol. 52, pp. 63-69, 2016.
- [24] L. Gan, M. Li, J. Gu, X. Qi, and P. Wei, “Improvement of 2-D Direction Finding Algorithm of Coherent Signals based on Two L-Shape Arrays,” *International Conference on Communications, Circuits and Systems, 2009. ICCCAS 2009*, July 2009.
- [25] D. Liu, and J. Liang, “Joint SVD of Two Cross-Correlation Matrices to Achieve Automatic Pairing in 2-D Angle Estimation Problems,” *IEEE Antennas and Wireless Propagation Letters*, pp. 553-556, November 2007.
- [26] S. Liu, L. Yang, D. Li, and H. Cao, “Subspace Extension Algorithm for 2D DOA Estimation with L-Shaped Sparse Array,” *Multidimensional Systems and Signal Processing*, Springer, 2016.
- [27] X. Luo, P. Ping, and L. Gan, “Two-Dimensional Direction-of-Arrival Estimation Using Two Transform Matrices,” *IEEE 17th International Conference on Computational Science and Engineering (CSE)*, December 2014.
- [28] Q. Cheng, Y. Zhao, and J. Yang, “A novel 2-D DOA Estimation for Coherent

2.7 References

- Signals based on L-Shaped Array,” *International Conference on Wireless Communications & Signal Processing (WCSP)*, October 2012.
- [29] D. Liu, and J. Liang, “L-Shaped Array-bBsed 2-D DOA Estimation using Parallel Factor Analysis,” *8th World Congress on Intelligent Control and Automation (WCICA)*, August 2010.
- [30] F. Harabi, H. Changuel, and A. Gharsallah, “Direction of Arrival Estimation Method using A 2-L Shape Arrays Antenna,” *Progress In Electromagnetics Research*, vol. 69, pp. 145-160, 2007.
- [31] Y. Dong, C. Dong, W. Liu, H. Chen, J. Xu, and G. Zhao, “2-D DOA Estimation for L-Shaped Array With Array Aperture and Snapshots Extension Techniques,” *IEEE Signal Processing Letters*, vol. 24, no. 4, pp. 495-499, April 2017.
- [32] Y. Dong, C. Dong, J. Xu, and G. Zhao, “Computationally Efficient 2-D DOA Estimation for L-Shaped Array with Automatic Pairing,” *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1669-1672, January 2016.
- [33] Y. Dong, C. Dong, Z. Shen, and G. Zhao, “Conjugate Augmented Spatial Temporal Technique for 2-D DOA Estimation With L-Shaped Array,” *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1622-1625, October 2016.
- [34] G. Wang, J. Xin, N. Zheng, and A. Sano, “Computationally Efficient Subspace-Based Method for Two-Dimensional Direction Estimation With L-Shaped Array,” *IEEE Signal Processing Letters*, vol. 59, no. 7, pp. 1669-1672, July 2011.
- [35] J. Xin, and A. Sano, “Computationally Efficient Subspace-Based Method for Direction-of-Arrival Estimation Without Eigendecomposition,” *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 876-893, April 2004.
- [36] Y. S. Yoon, L. M. Kaplan, and J. H. McClellan, “TOPS: New DOA Estimator for Wideband Signals,” *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 1977-1989, June 2006.

2.7 References

- [37] H. Wang, and M. Kaveh, “Coherent Signal-Subspace Processing for the Detection and Estimation of Angles of Arrival of Multiple Wide-Band Sources,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 823-831, August 1985.
- [38] Q. Shen, W. Liu, W. Cui, S. Wu, Y. D. Zhang, and M. G. Amin, “Low-Complexity Direction-of-Arrival Estimation Based on Wideband Co-Prime Arrays,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 9, pp. 1445-1456, September 2015.
- [39] J. Cao, Z. Liu, and Y. Xu, “DOA Estimation for Wideband Sources Using Cross Correlation Transformation,” *19th Annual Wireless and Optical Communications Conference (WOCC)*, May 2010.
- [40] S. A. Zekavat and R. M. Buehrer, *Handbook Of Position Location: Theory, Practice, And Advances* IEEE Press, A John Wiley & Sons, 2012.
- [41] S. Haykin, and K. J. R. Liu, *Handbook Of Array Processing and Sensor Network* IEEE Press, A John Wiley & Sons, 2009.
- [42] J. W. Brown, and R. V. Churchill, *Complex Variables and Applications, Seventh Edition, International Edition*, McGraw-Hill Education (Asia), 2003.

Chapter 3

Wideband Direction-of-Arrival Estimation by using Unsupervised Learning

This chapter presents an alternative direction-of-arrival estimation method for wideband sources by using unsupervised learning. The problem of estimating direction-of-arrival of wideband signals are addressed and resolved by focusing the entire observation subspace in each frequency bin along with employing a multi-narrowband signal model as was previously proposed in Chapter 2. In addition to this observation process, we employ the Gaussian mixture model with a maximum likelihood estimation algorithm as an interpolation scheme, and wideband DOA results are estimated by interpolating the narrowband DOA results all frequency bins with this interpolation scheme. The performance is evaluated in terms of the root-mean-squared error over a range of the signal-to-noise ratio, and it is sufficient for exhibiting wide-band sources angle estimation.

3.1 Reviews on Recent Wideband DOA Estimations

As stated in Section 5.1, the fundamental competence of source localization has been extensively implemented in navigation systems for the exploration of sources,

3.1 Reviews on Recent Wideband DOA Estimations

which is known as direction-of-arrival (DOA) estimation. The main factors reducing the efficiency of source localization in navigation systems are background noise, multiple source locations, and variation in the sound patterns and source frequencies [1]. A number of methods have been introduced in order to improve the source localization efficiency, including, time-difference-of-arrival (TDOA)-based DOA estimation [2], delay and sum beamforming [3] and subspace method [4–6].

Subspace methods are increasingly utilized for the DOA estimation of multiple sources in array signal processing. The conventional subspace method known as multiple signal classification (MUSIC) decomposes the observation space into a signal subspace and a noise subspace by using eigenvalue decomposition [4]. Such conventional methods calculate the directions of sources using the noise subspace, leading to superior multiple-source localization. Nonetheless, only narrow-band sources can be localized.

In order to estimate the DOA for wide-band sources, the coherent signal subspace (CSS) method was introduced [7]. This method constructively focuses the observation subspace at distinct frequency bins to a single subspace associated with a reference frequency, producing a smoothed correlation matrix. Conventional narrow-band DOA estimators such as MUSIC may then be used for DOA estimation. A common shortcoming of CSS methods is the requirement of an initial DOA estimation. It was previously found that the estimation performance is sensitive to the initial conditions [8]. Poor initial conditions can lead to biased estimates.

This research therefore aims to propose a new DOA estimation scheme for wideband sources without any requirement of DOA initial estimation. Unlike the CSS method, in which the signal and noise subspaces in each frequency bin must be arranged into a single viable correlation matrix, the problem of estimating DOA of wideband signals are addressed and resolved by focusing the entire observation subspace in each frequency bin along with employing a multi-narrowband signal model as was previously proposed

3.2 DOA Estimation Scheme

in Chapter 3. In addition to this observation process, we employ the Gaussian mixture model with a maximum likelihood estimation algorithm as an interpolation scheme. After narrow-band DOA estimation is employed in each correlation matrix, the set of estimated DOAs can be classified into groups by the Gaussian mixture model with a maximum likelihood estimation algorithm. Furthermore, the proposed method provides synthesis of signal sources.

3.2 DOA Estimation Scheme

Consider a one-dimensional ULA model in which the microphones are uniformly spaced [1] with longitudinal wave-front sources, and the output of each microphone is uniformly sampled and decomposed into snapshots or frames by a short-time Fourier transform (STFT), with each frame containing a set of frequencies. The received signal vector from P incident sources at τ_k snapshot can be represented as

$$\mathbf{x}(\tau_k, \omega_j) = \mathbf{W}(\boldsymbol{\theta}, \omega_j) \mathbf{s}(\tau_k, \omega_j) + \mathbf{n}(\tau_k, \omega_j), \quad (3.1)$$

where $\tau_k, k = 1, 2, \dots, K$ are snapshots, $\omega_j, j = 1, 2, \dots, J$ are discrete frequencies, $\mathbf{x} \in \mathbb{C}^M$ is the sum of received signal vectors for all sources with different phases or angles, $\mathbf{s} \in \mathbb{C}^P$ is a signal source vector, $\mathbf{n} \in \mathbb{C}^P$ is a noise vector, $\boldsymbol{\theta} \in \mathbb{R}^P$ is an angle vector and $\mathbf{W} \in \mathbb{C}^{M \times P}$ is a phase difference matrix. Each attribute in \mathbf{W} depends on the data model, which was already described in Chapter 2;

$$w_m(\theta, \omega) = \exp\left(\frac{i2\pi\omega(m-1)d \sin \theta}{\omega_\lambda \lambda}\right). \quad (3.2)$$

where d is the spacing of microphone elements, m is the index of microphone elements, θ is the angle and λ is the wavelength. Note that $\omega = 2\pi f$ and $\omega_\lambda = 2\pi f_\lambda$. According to Chapter 2, to avoid confusion caused by many grating lobes and the lower power beam-width of the lobe in ULA radiation, d should be set to $\frac{\lambda}{2}$, where $\lambda = \frac{c}{f_\lambda}$ is the wavelength and f_λ is the frequency corresponding to the wavelength.

3.2 DOA Estimation Scheme

The classical subspace-based methods, called MUSIC and the estimation of signal parameters via a rotational invariance technique (ESPRIT) [4, 5], are the two most popular DOA estimation methods. In this work, the MUSIC method is chosen to estimate the angle sources using a modified phase difference function. The MUSIC spectrum is expressed as follows

$$P_{\text{MUSIC}}(\theta, \omega_j) = \frac{\mathbf{w}(\theta, \omega_j)^H \mathbf{w}(\theta, \omega_j)}{\mathbf{w}(\theta, \omega_j)^H \mathbf{E}(\omega_j) \mathbf{E}(\omega_j)^H \mathbf{w}(\theta, \omega_j)}, \quad (3.3)$$

where $P_{\text{MUSIC}}(\theta, \omega_j)$ is the reciprocal of the ordinary Euclidean distance from $\mathbf{w}(\theta, \omega_j)$ to $\mathbf{E}(\omega_j)$. $\mathbf{w}(\theta, \omega_j)$ is the angle steering vector, which can also be described by (3.2). $\mathbf{E} \in \mathbb{C}^{M \times (M-P)}$ is the set of eigenvectors of noise known as the noise subspace, which is calculated from the covariance matrix of $\mathbf{x}(\boldsymbol{\tau}, \omega_j)$. One previous study suggested that the number of microphones M is larger than number of incident sources P [4]. Prior information of P can be determined from the eigenvalues [4]. The DOAs of signal sources are estimated by locating the peaks of the MUSIC spectrum for each ω_j . Because the MUSIC method has a large computational requirement, the Root MUSIC method has been proposed in order to refine the DOA estimation and to alleviate the algorithm complexity [6]. Root MUSIC with the modified phase difference function can be expressed as follows

$$\theta = \sin^{-1} \left(-\text{angle}(z) \frac{\lambda \omega_\lambda}{2\pi \omega_j d} \right), \quad (3.4)$$

where $\text{angle}(\cdot)$ is the phase angle operator in radians. z is the roots of the polynomial function $f(z, \omega_j)$. The polynomial function $f(z, \omega_j)$ can be represented as $f(z, \omega_j) = \sum_{m=-M+1}^{M-1} z^m C_m(\omega_j)$, where $C_m(\omega_j) = \sum_{a-b=m} V_{a,b}(\omega_j)$ and $\mathbf{V}(\omega_j) = \mathbf{E}(\omega_j) \mathbf{E}(\omega_j)^H$. Among a roots in $f(z, \omega_j)$, the closer a root is to the unit circle, the more likely it corresponds to DOA. The remaining roots that are farther from the unit circle are noises. Note that the unit circle can be considered as the unit complex numbers.

3.2 DOA Estimation Scheme

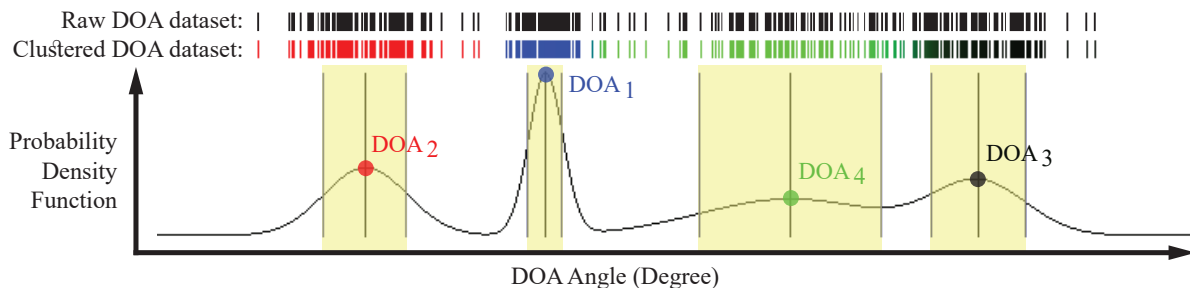


Fig. 3.1 Overviews of wideband DOA estimation by using GMM where the raw DOA datasets are obtained by estimating narrowband DOAs all frequency bins.

While the Root MUSIC method generates a large set of DOAs at each frequency and angle, unsupervised machine learning has been suggested as a potential solution to solves classification problem [9]. In this research work, Gaussian mixture model (GMM) with a maximum likelihood estimation algorithm is utilized to cluster local maxima of angle observations into P clusters or sources, where each observation belongs to a cluster with the parametric probability density function (PDF) represented as a weighted sum of Gaussian component densities [9]. The reason we chose GMM as unsupervised learning is due to its remarkable of flexible in terms of cluster covariance, can accommodates mixed membership and simple calculation. The PDF can be written as a linear superposition of Gaussian in the form

$$p(\boldsymbol{\theta}) = \sum_{p=1}^P \pi_p \mathcal{N}(\mu_p, \boldsymbol{\Sigma}_p), \quad (3.5)$$

where $\boldsymbol{\theta}$ is the set of DOAs obtained by the Root MUSIC method Eq.(3.4) without a frequency space, $\pi_p, p = 1, 2, \dots, P$ are the mixture weights, $\mathcal{N}(\mu_p, \boldsymbol{\Sigma}_p)$ is the multivariate normal distribution or multivariate Gaussian distribution, $\boldsymbol{\Sigma}_p, p = 1, 2, \dots, P$ are DOA covariance matrices and $\mu_p, p = 1, 2, \dots, P$ are the estimated DOA. The idea of estimating wideband DOAs via GMM is illustrated in Remark that the estimated number of sources P is defined by the maximum of the likelihood function of the observation in GMM algorithm, which is obtained by changing the P free parameters in Eq.(3.5).

3.3 Signal Synthesis

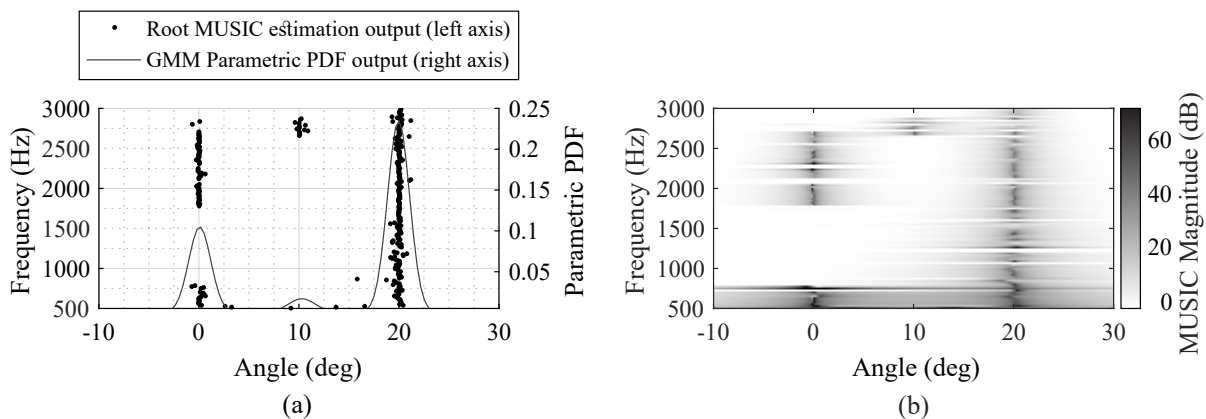


Fig. 3.2 Estimation results of the proposed method by (a) Root MUSIC and GMM fusion and (b) MUSIC where three uncorrelated sources are employed.

3.3 Signal Synthesis

Once the DOAs of the P incident signals have been estimated, $\mathbf{s}(\tau_k, \omega_j)$ can be estimated for synthesis of the signal sources by using Eq.(3.1) as follows:

$$\mathbf{s}(\tau_k, \omega_j) = (\mathbf{W}(\boldsymbol{\theta}_{\text{DOA}}, \omega_j))^{-1} (\mathbf{x}(\tau_k, \omega_j) - \mathbf{n}(\tau_k, \omega_j)) \quad (3.6)$$

where $\boldsymbol{\theta}_{\text{DOA}}$ is the DOA of sources, which is obtained from μ_p in Eq.(3.5). The sources in the time domain can be estimated by the inverse STFT of $\mathbf{s}(\tau_k, \omega_j)$.

3.4 Numerical Simulation

The proposed system was tested by computer simulation considering a 16-sensor ULA. The statistical performance was evaluated by performing 100 Monte Carlo runs for each scenario. Three far-field uncorrelated sources with the same power were placed at 0° , 10° and 20° . Source frequencies were 1, 1.75 and 2.5 kHz, respectively. The output of each microphone was decomposed into 192000 snapshots. The sampling frequency was thus 192 kHz, the spacing of microphone elements was 5 cm and speed of sound was assumed to be 340 m/s.

3.4 Numerical Simulation

Three other methods were tested for comparison: the generalized cross-correlation phase transform (GCC-PHAT) [2], the conventional CSS method [7] and a modification of the total least-squares CSS method (CSS-MTLS) [8]. For the proposed method, 218 frequency bins were utilized in the range from $\frac{\omega_\lambda}{4}$ to ω_λ , where ω_λ was 3.4 kHz. The conventional CSS method and CSS-MTLS used 435 frequency bins in the range from $f_0 - \frac{BW}{2}$ to $f_0 + \frac{BW}{2}$, where the central frequency f_0 was 3.4 kHz, bandwidth BW was 5.1 kHz and the preliminary approximate DOA was 10° . Both the CSS method and CSS-MTLS were recursively processed five times to update the focusing DOA and the preliminary approximate DOA. For GCC-PHAT, only two microphones were selected with a maximum spacing of microphone elements of 75 cm.

Figures 3.2(a) and 3.2(b) show examples of the proposed method output for a 16-sensor ULA model with three human voice sources at angles of 0° , 10° and 20° . The example indicates that there are three sources because one can see three peaks at the correct DOAs. Note that the magnitude of the parametric PDF is not necessarily proportional to the power of the signal sources.

Figure 6.3 compares the performance of the four methods in terms of the root-mean-square error (RMSE) of the sources at 0° , 10° and 20° , over the signal-to-noise ratio (SNR) range from -10 to 40 dB. The results indicate that the performance of CSS-MTLS is not significantly dominated by the SNR, whereas the other three methods are significantly dominated by the SNR. As the SNR decreases to less than -3 dB, CSS-MTLS exhibits better performance than the proposed method and the other two methods. The proposed method could not resolve the issue caused by higher variance in the GMM. When the SNR is as high as -3 dB, the proposed method begins to exhibit a smaller RMSE than CSS-MTLS and the other two methods. The results indicate that the proposed method exhibits high performance in the whole range of SNRs in terms of RMSE while CSS-MTLS exhibits high performance at high SNRs. The RMSE

3.5 Conclusions

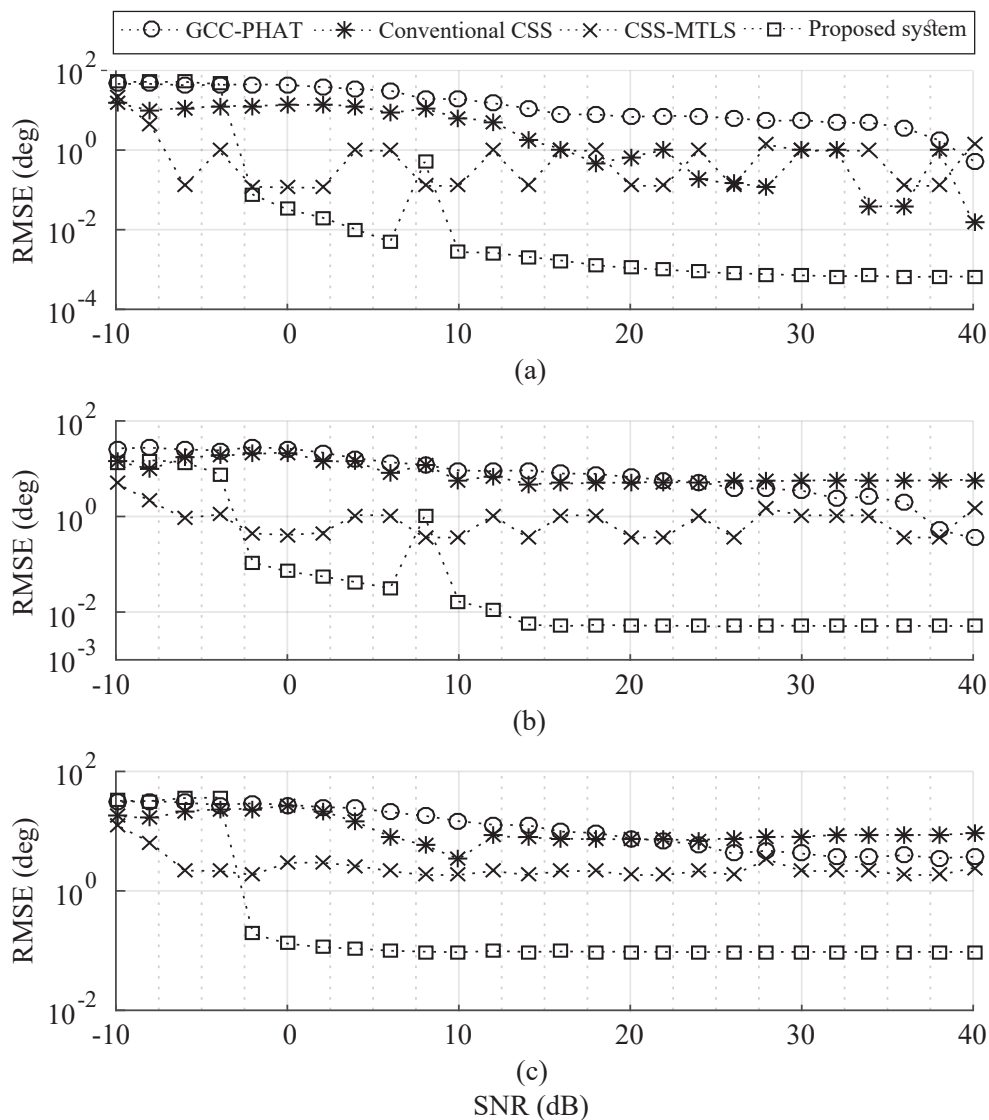


Fig. 3.3 RMSE performance evaluation; (a) 0° , (b) 10° , (c) 20° .

performances for the conventional CSS method and GCC-PHAT are similar.

3.5 Conclusions

This chapter presented an alternative DOA estimation method for wide-band sources by using a GMM with a maximum likelihood estimation algorithm for multiple frequencies and source angle estimation. The problem of estimating direction-of-arrival of wideband signals have been addressed and resolved by focusing the entire observation

3.5 Conclusions

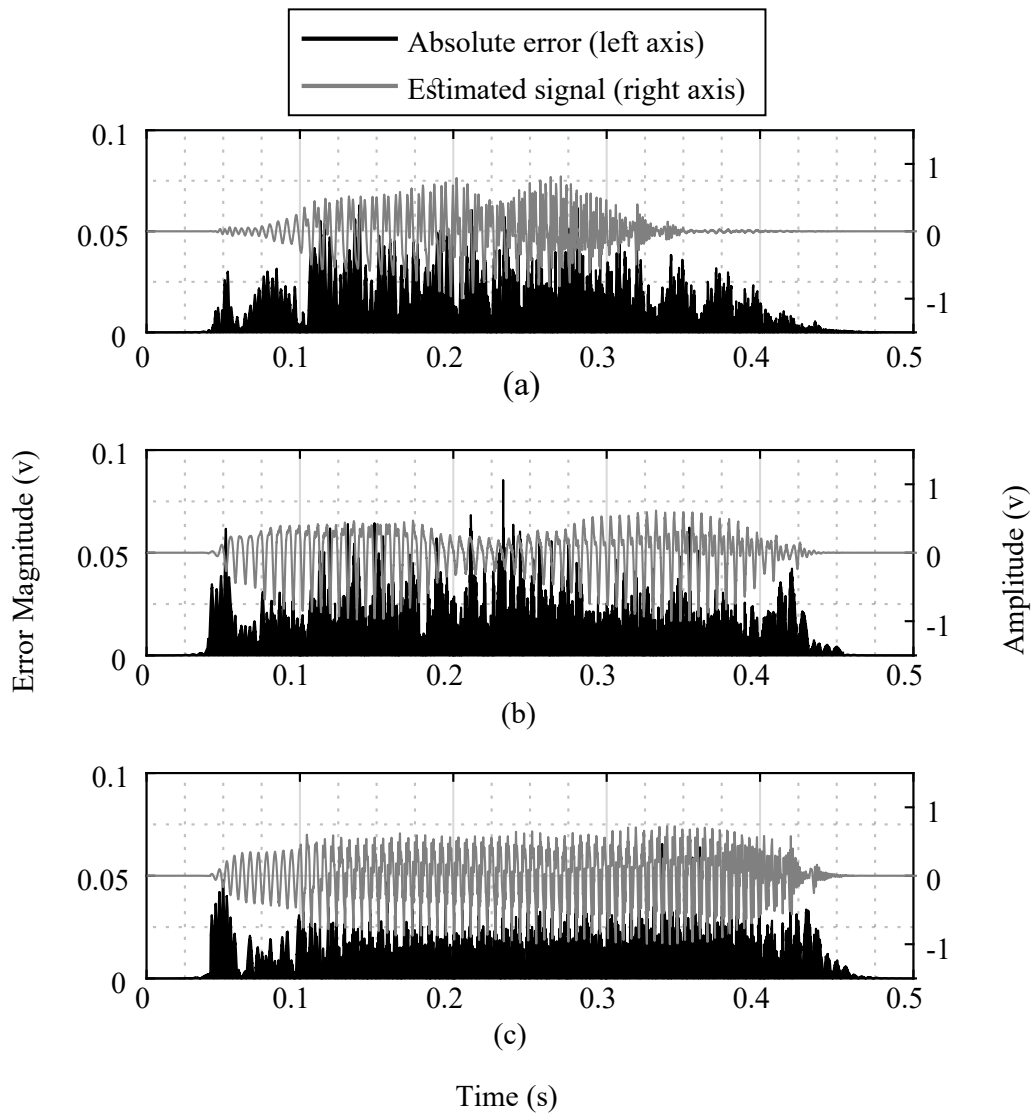


Fig. 3.4 Example of signal synthesis with error magnitude where three uncorrelated human speeches and 16 microphones are used; (a) first source, (b) second source, and (c) third source.

subspace in each frequency bin along with employing a multi-narrowband signal model as was previously proposed in Chapter 3. The performance is evaluated in terms of the root-mean-squared error over a range of the signal-to-noise ratio, and the results suggested that the proposed method may be a particularly effective method of DOA estimation. Furthermore, the proposed method enables the synthesis of signal sources and provides an alternative to intelligent source localization systems.

3.6 References

- [1] S. A. Zekavat and R.M. Buehrer: Handbook of Position Location: Theory, Practice, and Advances, Wiley-IEEE Press, 2011.
- [2] T. Nishiura, M. Nakamura, A. Lee, H. Saruwatari and K. Shikano: Talker tracking display on autonomous mobile robot with a moving microphone array, ICAD2002: 8th International Conference on Auditory Display, July 2002.
- [3] Y. Sasaki, M. Kabasawa and S. Thompson: Spherical microphone array for spatial sound localization for a mobile robot, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012.
- [4] R. Schmidt: Multiple emitter location and signal parameter estimation, IEEE Transactions on Antennas and Propagation, Vol. 34, No. 3, pp. 276-280, 1986.
- [5] R. Roy and T. Kailath: 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] B. D. Rao and K. V. S. Hari: Performance analysis of Root-MUSIC, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 37, No. 12, pp. 1939-1949, 1989.
- [7] H. Wang and M. Kaveh: Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 33, No. 4, pp. 823-831, 1985.
- [8] S. Valaee, B. Champagne and P. Kabal: Localization of wideband signals using least-squares and total least-squares approaches, IEEE Transactions on Signal Processing, Vol. 47, No. 5, pp. 1213-1222, 1999.
- [9] C. M. Bishop: Pattern Recognition and Machine Learning, Springer, 2006.

Chapter 4

An Efficient Wideband Direction-of-Arrival Estimation by using Orthogonal Procrustes Analysis

This chapter presents an efficient wideband two-dimensional direction-of-arrival (DOA) estimation for an L-shaped microphone array. We propose a way to construct a wideband sample cross-correlation matrix without any process of DOA preliminary estimation, such as beamforming technique, by exploiting sample cross-correlation matrices of two different frequencies for all frequency bins along with the Orthogonal Procrustes analysis. Subsequently, wideband DOAs can be estimated by using this wideband matrix along with a scheme of estimating DOA in a narrowband subspace method. Therefore, a contribution of our study is providing an alternative framework for recent narrowband subspace methods to estimating the DOA of wideband sources directly. It means that this framework enables cutting-edge techniques in the existing narrowband subspace methods to implement the wideband direction estimation for reducing the computa-

4.1 Robustness Issue in Recent Wideband DOA Estimation

tional complexity and facilitating the estimation algorithm. Theoretical analysis and effectiveness of the proposed method are substantiated through numerical simulations, which are performed in reverberating environments. The results show that performance of the proposed method performs better than others over a range of signal-to-noise ratio with just a few microphones. All these advantages make the proposed method a powerful tool for navigation systems based on acoustic signal processing.

4.1 Robustness Issue in Recent Wideband DOA Estimation

Direction-of-arrival (DOA) estimation in particular is an important part of navigation systems for the exploration of sources [7, 43], and has wide applications including in wireless communications, sonar and radar, as mentioned in Section 5.1 and Chapter 3. In addition, DOA estimation methods have been applied for human computer interaction in acoustic signal processing, as shown in the literature [2, 16, 20, 23, 26, 32]. Several approaches have been proposed as potential solution for DOA estimates. One of the most popular method is time-difference-of-arrival based DOA estimation [1, 10]; this method is generally known as generalized cross-correlation with a weighting function, such as, generalized cross-correlation with phase transform (GCC-PHAT) [10]. DOA is estimated using time shifting related to the location of the largest peak in time-domain representation of a normalized cross-correlation. Although a notable feature of GCC-PHAT is low computational requirement, but a major drawback of GCC-PHAT is its low robustness in noisy and multipath environments. In an alternative approach to narrowband DOA estimates, subspace method has been proposed in an effort to reduce computational complexity, and improve estimation performance in multipath environment. The most prominent methods are multiple signal classification (MUSIC) [24], estimation of

4.1 Robustness Issue in Recent Wideband DOA Estimation

signal parameters via rotational invariance techniques (ESPRIT) [22], and propagator method (PM) [14, 15], which have been used frequently for one-dimensional DOA estimation along with uniform linear array (ULA) of sensors. In case of a two-dimensional (2D) DOA estimation, a new geometrical structure of sensor array is required. The first systematic study on effectiveness of the geometrical structures was conducted since 1991 [9]. According to their research, they demonstrate that an octagonal 2D array, which is close to a circular array in structure, offers the most minimal Cramer Rao bound (CRB) of DOA; CRB indicates a lower bound on the variance of unbiased estimators, and the low CRB is said to be potential high accuracy of DOA estimation. In addition, the authors underline that CRB of DOA for the octagonal 2D array is slightly lower than an L-shaped array; CRB of DOA for the L-shaped array, which consists of two ULAs connected orthogonally at one end of each ULA, can be considerably lower than for other geometrical structure as well as the octagonal 2D array. (For details, see [9].) The authors further suggest that the L-shaped array seems to be innovative and interesting, because it possesses independent properties from distinct coordinate axes; in other words, it allows for simple implementation. Due to these advantages, the L-shaped array is widely applied to 2D DOA estimation method [19, 35]. Further geometric structures based on the L-shaped array are currently studied as shown the literature [42, 44].

Although subspace method for narrowband 2D DOA estimates with L-shaped array may unable to applies wideband signals directly [3, 5, 12, 13, 17, 21, 31, 36]; however, there is one possible solution to overcome this problem by employing the narrowband DOA estimation method intensively each temporal frequency, which can be found in the previous research works [29, 38]. It should be considered that intensive computational costs, encountered in the previous works [29, 38], may restrict to be implemented in practical application.

4.1 Robustness Issue in Recent Wideband DOA Estimation

Several approaches were proposed to deal with wideband sources, for example, the incoherent signal subspace (ISS) is one of the classical wideband DOA estimation method [25]. ISS firstly constructs a noise subspace model each temporal frequency. Then, the wideband DOA result is obtained by interpolating the narrowband DOA results, which are processed by narrowband subspace method individually. In particular, MUSIC algorithm is modified, and the modified MUSIC is employed by ISS as a narrowband subspace method [24]; this integration is known as the incoherent MUSIC (IMUSIC) [25]. Accuracy performance of IMUSIC was demonstrated to be effective method for wideband DOA estimation in high signal-to-noise ratio (SNR) region. However, there is one serious problem at an interpolating process in IMUSIC; a single small distorting narrowband DOA result could affect the wideband DOA result. In order to overcome such undesirable effect, the coherent signal subspace (CSS) method was proposed [37]. CSS constructively focuses multiple cross-correlation matrices of received signals for all frequency bins into a single correlation matrix with a single reference frequency via liner transformation process. Then, the wideband DOA result is estimated by using this matrix. In the transformation process of CSS, in particular, requires an DOA preliminary estimation [33,34,37]; hence, a common shortcoming is clearly recognized as a requirement of DOA preliminary estimation, which means that any inferior initiation can lead to biased estimates.

Various approaches were proposed to overcome a requirement of DOA preliminary estimation in CSS, such as, the test of orthogonality of projected subspace (TOPS) [40]. TOPS estimates the wideband DOA by transforming a signal subspace of a one reference frequency, and then measuring orthogonality between the previous signal subspace and noise subspace each temporal frequency. This method provides much higher accuracy than CSS as stated in [40]; however, the undesirable false peaks in spatial spectrum still remain as well as CSS. Many attempts were made recently to reduce the false

4.1 Robustness Issue in Recent Wideband DOA Estimation

peaks in spatial spectrum [8, 18]. Obviously, the computational complexity increases dramatically compared to the original method.

Therefore the purpose of this research work is to investigate a new way for estimating a wideband 2D DOA in a more efficient way. Firstly, the recent methods for wideband DOA estimation, such as IMUSIC, CSS, TOPS, and the test of orthogonality of frequency subspaces (TOFS) [41], have only investigated in ULA model [8, 18, 25, 33, 37, 40, 41]; therefore, the proposed method presents an alternative solution for estimating the wideband 2D DOA with L-shaped array placed in the X and Z coordinate axes. Secondly, the proposed method is inspired by CSS with further improvement of linear transformation scheme; we propose a way to construct a wideband sample cross-correlation matrix without any process of DOA preliminary estimation. Additionally, the proposed method employ transformation matrices, which are constructed by performing the singular value decomposition (SVD) of a new unique cross-correlation matrix, where elements in the row and column positions are sample cross-correlation matrices of two different frequencies from the same coordinate axis. This new unique cross-correlation matrix seem to play a role as the universal correlation matrix in CSS [33, 34, 37], but the process of initial DOA estimation is not longer required; therefore, a contribution of our proposed method is providing an alternative framework for the recent narrowband subspace methods with L-shaped array, as shown in [3, 5, 12, 13, 17, 21, 31, 36], to estimating the 2D DOA of wideband sources directly. It implies that the new framework enable cutting-edge techniques in the existing narrowband subspace methods to implement the wideband 2D DOA estimation for reducing the computational complexity and facilitating the estimation algorithm. Inspired by useful techniques in recent narrowband subspace methods [3, 17, 28, 31], in this work, the azimuth and elevation are estimated simultaneously using the Eigenvalue decomposition (EVD) of left and right singular vector form the results of SVD of the proposed unique

4.2 Preliminaries

cross-correlation matrix. The theoretical analysis and effectiveness of proposed method are substantiated through numerical simulations. The results show that performance of the proposed method performs better than others over a range of signal-to-noise ratio with just a few microphones. Furthermore, the results show that the performance of proposed method is especially effective for wideband 2D DOA estimation in real reverberating environments.

4.2 Preliminaries

4.2.1 Data Model

An L-shaped array is considered as illustrated in Fig. 2.1, where $\{X, Y, Z\}$ is a set of Cartesian coordinate axes for a three-dimensional space, a reference point is pinned at z_1 , s_k represents a wideband plane wave of source k to the reference point. ψ_k represents the angle between the X and Y axes; it is named as azimuth. ϕ_k is named as x subarray angle, which denotes the angle between the X and the source axes. Similarly, θ_k is named as z subarray angle or zenith, which denotes the angle between the Z and the source axes. By Euler's rotation theorem, we have the following relation;

$$\cos \phi_k = \sin \theta_k \cos \psi_k, \quad (4.1)$$

where $\psi_k = \bar{\phi}_k$ in Fig. 2.1. The method proposed considers an acoustic source as we mentioned in the previous section. Each point on the axis stands for an acoustic sensor, such as a microphone. Unlike other wideband subspace methods that decompose the sensor outputs into multiple narrowband signals using a filter bank and the discrete Fourier transform (DFT), the received signals are transformed into time-frequency representation via short-time Fourier transform (STFT). Additionally, we employ the previous signal model as described in Chapter 2. To make this chapter read easily, we

4.2 Preliminaries

recall some necessary definitions here:

$$\begin{aligned}\mathbf{x}(t, f) &= \mathbf{A}_x(\boldsymbol{\phi}, f) \mathbf{s}(t, f) + \mathbf{w}_x(t, f), \\ \mathbf{z}(t, f) &= \mathbf{A}_z(\boldsymbol{\theta}, f) \mathbf{s}(t, f) + \mathbf{w}_z(t, f),\end{aligned}\tag{4.2}$$

where

$$\begin{aligned}\mathbf{A}_x(\boldsymbol{\phi}, f) &= \begin{bmatrix} \mathbf{a}_x(\phi_1, f) & \mathbf{a}_x(\phi_2, f) & \dots & \mathbf{a}_x(\phi_K, f) \end{bmatrix}, \\ \mathbf{A}_z(\boldsymbol{\theta}, f) &= \begin{bmatrix} \mathbf{a}_z(\theta_1, f) & \mathbf{a}_z(\theta_2, f) & \dots & \mathbf{a}_z(\theta_K, f) \end{bmatrix}, \\ \mathbf{a}_x(\phi_k, f) &= \left[e^{\alpha_x(\phi_k, f)j} \quad e^{2\alpha_x(\phi_k, f)j} \quad \dots \quad e^{M\alpha_x(\phi_k, f)j} \right]^T, \\ \mathbf{a}_z(\theta_k, f) &= \left[1 \quad e^{\alpha_z(\theta_k, f)j} \quad \dots \quad e^{(M-1)\alpha_z(\theta_k, f)j} \right]^T, \\ \alpha_x(\phi_k, f) &= \left(\frac{f}{f_o} \right) \cdot \left(\frac{2\pi d \cos \phi_k}{\lambda} \right), \\ \alpha_z(\theta_k, f) &= \left(\frac{f}{f_o} \right) \cdot \left(\frac{2\pi d \cos \theta_k}{\lambda} \right).\end{aligned}\tag{4.3}$$

From above, $\mathbf{x}(t, f) \in \mathbb{C}^M$ and $\mathbf{z}(t, f) \in \mathbb{C}^M$ are the summation of the received signal vectors for all sources corresponding to the x and z subarrays, $\mathbf{A}_x(\boldsymbol{\phi}, f) \in \mathbb{C}^{M \times K}$ and $\mathbf{A}_z(\boldsymbol{\theta}, f) \in \mathbb{C}^{M \times K}$ are array manifold matrices in the x and z subarrays, $\mathbf{s}(t, f) \in \mathbb{C}^K$ is a signal source vector, and $\mathbf{w}_x(t, f) \in \mathbb{C}^M$ and $\mathbf{w}_z(t, f) \in \mathbb{C}^M$ are additive noise vectors for all microphone elements corresponding to the x and z subarrays. Note that the variable f is a source frequency, f_o is a reference frequency, λ is a wavelength, d is the spacing of the microphone elements any subarray, t is time, M is the number of microphone elements each subarray, K is the number of incident sources, and $\lambda = \frac{c}{f_o}$ where c denotes speed of sound in a current environment.

4.2.2 Basic Assumptions

Based on the recent literature reviews, the following basic assumptions on the data model are required:

Assumption 1: Number of incident sources K is required and can be predicted by the following techniques [30, 39]. Note that $K < M$ [7, 30, 39, 43].

4.2 Preliminaries

Assumption 2: Spacing between the adjacent elements d of each subarray and elements spacing between x_1 and z_1 should be set to $\frac{\lambda}{2}$ for avoid confusion caused by the side lobes in array structure radiation [7, 9, 27, 43].

Assumption 3: Source frequency f has to satisfy the constraint that $f \leq f_o$ for maintain a power beam-width of the grating lobe in array structure radiation [27].

Assumption 4: The source $\mathbf{s}(t, f)$ is assumed to be Gaussian complex random variable with zero mean, which is suggested by the literature [14, 15, 22, 24, 28, 36]. However, we consider the source as human speech; therefore, $\mathbf{s}(t, f)$ can be Super-Gaussian complex random variable, and it is not a stationary signal for the most general case when giving an appropriate period of time capture. According to the acoustic theory of speech production [4], frequency dependence of the source are existed; a cross-covariance between the signal source and itself with distinct frequencies is given as

$$\text{cov}(s_k(t, f), s_k(t, f')) = c_{s_k\{f, f'\}}, \quad (4.4)$$

where $c_{s_k\{f, f'\}} \in \mathbb{C}$. Assume that the sources are uncorrelated each other, which implies that $s_k(t, f)$ and $s_{k'}(t, f')$ are statistically independent; $\text{cov}(s_k(t, f), s_{k'}(t, f')) = 0$ where $k \neq k'$. On the contrary, the source with different frequencies at same element are partially dependent. Therefore, a sample cross-covariance matrix over two distinct frequencies can be expressed as

$$\mathbf{S}_{\{f, f'\}} = \text{diag}(c_{s_1\{f, f'\}}, c_{s_2\{f, f'\}}, \dots, c_{s_k\{f, f'\}}). \quad (4.5)$$

If $f' = f$, then $c_{s_k\{f, f\}} = \sigma_{s_k\{f\}}^2$ where $\sigma_{s_k\{f\}}^2 \in \mathbb{R}_{\geq 0}$ is a variance at frequency f of the source.

Assumption 5: The additive noise vectors $\mathbf{w}_x(t, f)$ and $\mathbf{w}_z(t, f)$ are Gaussian complex random variables as suggested by the previous studies [14, 15, 22, 24, 28, 36].

4.2 Preliminaries

Noise cross-covariance matrices over two distinct frequencies are expressed as follows:

$$\begin{aligned}
\mathbf{W}_{xx\{f,f'\}} &= c_{w_x\{f,f'\}} \mathbf{I}_M, \\
\mathbf{W}_{zz\{f,f'\}} &= c_{w_z\{f,f'\}} \mathbf{I}_M, \\
\mathbf{W}_{xz\{f,f'\}} &= \mathbf{O}_{M \times M}, \\
\mathbf{W}_{zx\{f,f'\}} &= \mathbf{O}_{M \times M}.
\end{aligned} \tag{4.6}$$

where $c_{w_x\{f,f'\}}, c_{w_z\{f,f'\}} \in \mathbb{C}$. Note that $c_{w_x\{f,f\}} = \sigma_{w_x\{f\}}^2$, and $c_{w_z\{f,f\}} = \sigma_{w_z\{f\}}^2$, where $\sigma_{w_x\{f\}}^2, \sigma_{w_z\{f\}}^2 \in \mathbb{R}_{\geq 0}$ are variance at frequency f of the signal noises. If the noise with different frequencies at same element are independent, we have $c_{w_x\{f,f'\}} = 0$, and $c_{w_z\{f,f'\}} = 0$.

4.2.3 Problem Formulation

Under Assumption 1 to 5 and Eq. (6.1), a sample cross-correlation matrices are defined as follows:

$$\begin{aligned}
\mathbf{R}_{xx\{f,f'\}} &= E \{ \mathbf{x}(t, f) \mathbf{x}^H(t, f') \} \\
&= \mathbf{A}_x(\boldsymbol{\phi}, f) E \{ \mathbf{s}(t, f) \mathbf{s}^H(t, f') \} \mathbf{A}_x^H(\boldsymbol{\phi}, f') \\
&\quad + E \{ \mathbf{w}_x(t, f) \mathbf{w}_x^H(t, f') \} \\
&= \mathbf{A}_x(\boldsymbol{\phi}, f) \mathbf{S}_{\{f,f'\}} \mathbf{A}_x^H(\boldsymbol{\phi}, f') + \mathbf{W}_{xx\{f,f'\}}, \\
\mathbf{R}_{zz\{f,f'\}} &= E \{ \mathbf{z}(t, f) \mathbf{z}^H(t, f') \} \\
&= \mathbf{A}_z(\boldsymbol{\theta}, f) E \{ \mathbf{s}(t, f) \mathbf{s}^H(t, f') \} \mathbf{A}_z^H(\boldsymbol{\theta}, f') \\
&\quad + E \{ \mathbf{w}_z(t, f) \mathbf{w}_z^H(t, f') \} \\
&= \mathbf{A}_z(\boldsymbol{\theta}, f) \mathbf{S}_{\{f,f'\}} \mathbf{A}_z^H(\boldsymbol{\theta}, f') + \mathbf{W}_{zz\{f,f'\}},
\end{aligned} \tag{4.7}$$

where $\mathbf{R}_{xx\{f,f'\}}, \mathbf{R}_{zz\{f,f'\}} \in \mathbb{C}^{M \times M}$. Additionally, the sample cross-correlation matrix can be formulated between the distinct subarrays [28]; the effort made those noise matrices vanish to zeros as below

$$\begin{aligned}
\mathbf{R}_{xz\{f,f'\}} &= E \{ \mathbf{x}(t, f) \mathbf{z}^H(t, f') \} \\
&= \mathbf{A}_x(\boldsymbol{\phi}, f) E \{ \mathbf{s}(t, f) \mathbf{s}^H(t, f') \} \mathbf{A}_z^H(\boldsymbol{\theta}, f') \\
&\quad + E \{ \mathbf{w}_x(t, f) \mathbf{w}_z^H(t, f') \} \\
&= \mathbf{A}_x(\boldsymbol{\phi}, f) \mathbf{S}_{\{f,f'\}} \mathbf{A}_z^H(\boldsymbol{\theta}, f'),
\end{aligned} \tag{4.8}$$

4.2 Preliminaries

where $\mathbf{R}_{xz\{f,f'\}} \in \mathbb{C}^{M \times M}$. As mentioned in CSS [33,34,37], transformation matrices are employed in order to transform sample cross-correlation matrices into a single wideband cross-correlation matrix, which is called as a universal cross-correlation matrix. Based on the L-shaped structure, the universal cross-correlation matrix is defined by

$$\begin{aligned} \mathbf{R}_{xz} &= \int_{f_{\min}}^{f_o} \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f,f\}} \mathbf{T}_{z\{f\}}^H df \\ &= \int_{f_{\min}}^{f_o} \mathbf{T}_{x\{f\}} \mathbf{A}_x(\boldsymbol{\phi}, f) \mathbf{S}_{\{f,f\}} \mathbf{A}_z^H(\boldsymbol{\theta}, f) \mathbf{T}_{z\{f\}}^H df \\ &= \mathbf{A}_x(\boldsymbol{\phi}, f_o) \left(\int_{f_{\min}}^{f_o} \mathbf{S}_{\{f,f\}} df \right) \mathbf{A}_z^H(\boldsymbol{\theta}, f_o), \end{aligned} \quad (4.9)$$

where

$$\begin{aligned} \mathbf{A}_x(\boldsymbol{\phi}, f_o) &= \mathbf{T}_{x\{f\}} \mathbf{A}_x(\boldsymbol{\phi}, f), \\ \mathbf{A}_z(\boldsymbol{\theta}, f_o) &= \mathbf{T}_{z\{f\}} \mathbf{A}_z(\boldsymbol{\theta}, f), \end{aligned} \quad (4.10)$$

f_{\min} denotes the minimum frequency, and $\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}} \in \mathbb{C}^{M \times M}$ denote the transformation matrices in the x and z subarrays, respectively. Note that $\mathbf{T}_{x\{f\}}$ and $\mathbf{T}_{z\{f\}}$ are unitary on any given f ;

$$\begin{aligned} \mathbf{T}_{x\{f\}}^H \mathbf{T}_{x\{f\}} &= \mathbf{I}_M, \\ \mathbf{T}_{z\{f\}}^H \mathbf{T}_{z\{f\}} &= \mathbf{I}_M. \end{aligned} \quad (4.11)$$

$\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}}$ are originally designed by minimizing the Frobenius norm for the array manifold errors [37]. Alternatively, the transformation matrices can be designed along with $\mathbf{R}_{xx\{f,f\}} - \mathbf{W}_{xx\{f,f\}}$ and $\mathbf{R}_{zz\{f,f\}} - \mathbf{W}_{zz\{f,f\}}$. (For the details, see [34].) As stated in the introduction, the construction of $\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}}$ requires an DOA preliminary estimation; it means that any inferior initiation can lead to biased estimates. To overcome this difficulty, the new concept and scheme without any process of DOA preliminary estimation are presented in next section.

4.3 Proposed Method

Overview of the proposed method, consisting of the transformation matrix estimation, universal cross-correlation matrix calculation, and angle estimation scheme, is presented in this section.

4.3.1 New Optimization Problems and its Solutions

In this section, a new optimization problem for the transformation matrix estimation and its solution are presented. First we begin by introducing a lemma that will be useful for solving the proposed optimization problems.

Lemma 1. *Given four sets of distinct frequencies in Eq. (6.2) by $\{f, f\}$, $\{f_o, f_o\}$, $\{f, f_o\}$, and $\{f_o, f\}$ respectively where f is less than or equal f_o , since the incident sources have to carry the signal with a minimum frequency of at least f_{\min} and more than or equal the reference frequency f_o , then the rank of sample cross-correlation matrices $\mathbf{R}_{xz\{f,f\}}$, $\mathbf{R}_{xz\{f_o,f_o\}}$, $\mathbf{R}_{xz\{f,f_o\}}$, and $\mathbf{R}_{xz\{f_o,f\}}$ are always K .*

Proof. We start by considering two sets of pair frequencies $\{f, f\}$ and $\{f_o, f_o\}$. From Eq. (4.5), the sample covariance matrices are redefined as follows:

$$\begin{aligned}\mathbf{S}_{\{f,f\}} &= \text{diag}\left(\sigma_{s_1\{f\}}^2, \sigma_{s_2\{f\}}^2, \dots, \sigma_{s_K\{f\}}^2\right), \\ \mathbf{S}_{\{f_o,f_o\}} &= \text{diag}\left(\sigma_{s_1\{f_o\}}^2, \sigma_{s_2\{f_o\}}^2, \dots, \sigma_{s_K\{f_o\}}^2\right).\end{aligned}\tag{4.12}$$

In case without loss of generality, the incident sources have to carry the signal in frequency ranges between f_{\min} and f_o or more. Therefore, it is readily seen that:

$$\begin{aligned}\sigma_{s_k\{f\}}^2 &> 0; \forall f \in \mathbb{R}_{\geq f_{\min}}, \\ \sigma_{s_k\{f_o\}}^2 &> 0; f_o \geq f_{\min}.\end{aligned}\tag{4.13}$$

As a matter of fact that rank of a square matrix is equal to number of non-zero diagonal entries, rank of $\mathbf{S}_{\{f,f\}}$ and $\mathbf{S}_{\{f_o,f_o\}}$ are clearly equal to K . Since the array manifold

4.3 Proposed Method

matrices $\mathbf{A}_x(\phi, f)$, $\mathbf{A}_x(\phi, f_o)$, $\mathbf{A}_z(\theta, f)$, $\mathbf{A}_z(\theta, f_o)$ always have full rank of K from the basic assumptions, thus, a rank of $\mathbf{R}_{xz\{f,f\}}$ and $\mathbf{R}_{xz\{f_o,f_o\}}$ are K .

In case of $\mathbf{S}_{\{f,f_o\}}$, $\mathbf{S}_{\{f_o,f\}}$, and according to the basic assumption that those signal sources with different frequencies on the same element can be partially dependent, therefore:

$$\begin{aligned} E\{s_k(\mathbf{t}, f) s_k^*(\mathbf{t}, f_o)\} &\neq 0, \\ E\{s_k(\mathbf{t}, f_o) s_k^*(\mathbf{t}, f)\} &\neq 0, \end{aligned} \quad (4.14)$$

where $\forall f \in \mathbb{R}_{\geq f_{\min}}$, $f_o \geq f_{\min}$. Substituting the above equations in Eq. (4.5), rank of $\mathbf{R}_{xz\{f,f_o\}}$ and $\mathbf{R}_{xz\{f_o,f\}}$ are K . ■

After Lemma 1 is defined, We continue to introduce a new optimization problem for the transformation matrix estimation, $\mathbf{T}_{x\{f\}}$, $\mathbf{T}_{z\{f\}}$ are solved simultaneously as the solutions to

$$\begin{aligned} &\underset{\mathbf{T}_{x\{f\}}}{\text{minimize}} \quad \left\| \mathbf{R}_{xz\{f_o,f_o\}} - \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f,f_o\}} \right\|_{\text{F}}^2 \\ &\text{subject to} \quad \mathbf{T}_{x\{f\}}^H \mathbf{T}_{x\{f\}} = \mathbf{I}_M \\ &\quad \quad \quad f_{\min} \leq f \leq f_o, \end{aligned} \quad (4.15)$$

$$\begin{aligned} &\underset{\mathbf{T}_{z\{f\}}}{\text{minimize}} \quad \left\| \mathbf{R}_{xz\{f_o,f_o\}} - \mathbf{R}_{xz\{f_o,f\}} \mathbf{T}_{z\{f\}}^H \right\|_{\text{F}}^2 \\ &\text{subject to} \quad \mathbf{T}_{z\{f\}}^H \mathbf{T}_{z\{f\}} = \mathbf{I}_M \\ &\quad \quad \quad f_{\min} \leq f \leq f_o. \end{aligned} \quad (4.16)$$

As the matter of fact that above objective functions is unable to reach a zero-norms, because $\left\| \mathbf{S}_{\{f_o,f_o\}} - \mathbf{S}_{\{f,f_o\}} \right\|_{\text{F}}^2$, and $\left\| \mathbf{S}_{\{f_o,f_o\}} - \mathbf{S}_{\{f_o,f\}} \right\|_{\text{F}}^2$ are not equal to zero-norms. According to the orthogonal Procrustes problem [6], there are several possible solutions to solve these problems without loss of generality. We can obtain the solutions as follows.

Theorem 1. *Given a frequency f in Eqs. (4.15) and (4.16), the dual optimization problems are rewritten as*

$$\begin{aligned} &\underset{\mathbf{T}_{x\{f\}}}{\text{maximize}} \quad \text{tr}(\Re(\mathbf{T}_{x\{f\}} \boldsymbol{\Psi}_{x\{f\}})) \\ &\text{subject to} \quad \mathbf{T}_{x\{f\}}^H \mathbf{T}_{x\{f\}} = \mathbf{I}_M \\ &\quad \quad \quad f_{\min} \leq f \leq f_o, \end{aligned} \quad (4.17)$$

4.3 Proposed Method

$$\begin{aligned}
& \underset{\mathbf{T}_{z\{f\}}}{\text{maximize}} && \text{tr}(\Re(\mathbf{T}_{z\{f\}}\boldsymbol{\Psi}_{z\{f\}})) \\
& \text{subject to} && \mathbf{T}_{z\{f\}}^H \mathbf{T}_{z\{f\}} = \mathbf{I}_M \\
& && f_{\min} \leq f \leq f_o,
\end{aligned} \tag{4.18}$$

where

$$\boldsymbol{\Psi}_{x\{f\}} = \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H, \tag{4.19}$$

and

$$\boldsymbol{\Psi}_{z\{f\}} = \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}}. \tag{4.20}$$

Consider $\boldsymbol{\Psi}_{x\{f\}}$ and $\boldsymbol{\Psi}_{z\{f\}}$ can be factorized into singular value decomposition (SVD) forms; $\boldsymbol{\Psi}_{x\{f\}} = \mathbf{U}_{x_s\{f\}} \boldsymbol{\Sigma}_{x_s\{f\}} \mathbf{V}_{x_s\{f\}}^H + \mathbf{U}_{x_w\{f\}} \boldsymbol{\Sigma}_{x_w\{f\}} \mathbf{V}_{x_w\{f\}}^H$ and $\boldsymbol{\Psi}_{z\{f\}} = \mathbf{U}_{z_s\{f\}} \boldsymbol{\Sigma}_{z_s\{f\}} \mathbf{V}_{z_s\{f\}}^H + \mathbf{U}_{z_w\{f\}} \boldsymbol{\Sigma}_{z_w\{f\}} \mathbf{V}_{z_w\{f\}}^H$. The possible solution in Eqs. (4.17) and (4.18) are given as

$$\mathbf{T}_{x\{f\}} = \mathbf{V}_{x_s\{f\}} \mathbf{U}_{x_s\{f\}}^H, \tag{4.21}$$

and

$$\mathbf{T}_{z\{f\}} = \mathbf{V}_{z_s\{f\}} \mathbf{U}_{z_s\{f\}}^H. \tag{4.22}$$

where

$$\begin{aligned}
\mathbf{U}_{x_s\{f\}} &= \mathbf{A}_x(\phi, f) \mathbf{P}_{x\{f\}}, \\
\mathbf{V}_{x_s\{f\}} &= \mathbf{A}_x(\phi, f_o) \mathbf{Q}_{x\{f\}}, \\
\mathbf{U}_{z_s\{f\}} &= \mathbf{A}_z(\theta, f) \mathbf{P}_{z\{f\}}, \\
\mathbf{V}_{z_s\{f\}} &= \mathbf{A}_z(\theta, f_o) \mathbf{Q}_{z\{f\}},
\end{aligned} \tag{4.23}$$

$\mathbf{P}_{x\{f\}}, \mathbf{P}_{z\{f\}}, \mathbf{Q}_{x\{f\}}, \mathbf{Q}_{z\{f\}} \in \mathbb{C}^{K \times K}$ are invertible matrices. $\mathbf{U}_{x_s\{f\}}, \mathbf{U}_{z_s\{f\}} \in \mathbb{C}^{M \times K}$, $\mathbf{V}_{x_s\{f\}}, \mathbf{V}_{z_s\{f\}} \in \mathbb{C}^{M \times K}$, $\boldsymbol{\Sigma}_{x_s\{f\}}, \boldsymbol{\Sigma}_{z_s\{f\}} \in \mathbb{C}^{K \times K}$ are the matrices of left and right singular vectors and diagonal matrices of singular values in the signal subspace, respectively. Likewise, $\mathbf{U}_{x_w\{f\}}, \mathbf{U}_{z_w\{f\}} \in \mathbb{C}^{M \times M-K}$, $\mathbf{V}_{x_w\{f\}}, \mathbf{V}_{z_w\{f\}} \in \mathbb{C}^{M \times M-K}$, $\boldsymbol{\Sigma}_{x_w\{f\}}, \boldsymbol{\Sigma}_{z_w\{f\}} \in \mathbb{C}^{M-K \times M-K}$ are the singular matrices in the noise subspace.

4.3 Proposed Method

Proof. We begin by introducing *Orthogonal Procrustes problems* from [6]. The objective function in Eq. (4.15) can be rederived by

$$\begin{aligned}
& \left\| \mathbf{R}_{xz\{f_o, f_o\}} - \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \right\|_{\mathbb{F}}^2 \\
&= \text{tr} \left(\left(\mathbf{R}_{xz\{f_o, f_o\}} - \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \right) \left(\mathbf{R}_{xz\{f_o, f_o\}} - \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \right)^H \right) \\
&= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \mathbf{T}_{x\{f\}}^H \right) \\
&\quad - \text{tr} \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H + \mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \mathbf{T}_{x\{f\}}^H \right).
\end{aligned} \tag{4.24}$$

From Eq. (4.11) and basic properties of trace operator, we have

$$\begin{aligned}
& \text{tr} \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \mathbf{T}_{x\{f\}}^H \right) \\
&= \text{tr} \left(\mathbf{R}_{xz\{f, f_o\}} \mathbf{T}_{x\{f\}}^H \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}}^H \right) \\
&= \text{tr} \left(\mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \right),
\end{aligned} \tag{4.25}$$

Since the matrices $\mathbf{R}_{xz\{f_o, f_o\}}$, $\mathbf{R}_{xz\{f, f_o\}}$ consistently possess the Hermitian property, we can have

$$\begin{aligned}
& \text{tr} \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H + \mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \mathbf{T}_{x\{f\}}^H \right) \\
&= \text{tr} \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H + \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right)^H \right) \\
&= \text{tr} \left(2 \cdot \Re \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) \right) \\
&= 2 \cdot \text{tr} \left(\Re \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) \right).
\end{aligned} \tag{4.26}$$

Substituting Eqs. (4.25) and (4.26) into Eq. (4.24), then

$$\begin{aligned}
& \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \mathbf{T}_{x\{f\}}^H \right) \\
&- \text{tr} \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H + \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right)^H \right) \\
&= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \right) \\
&\quad - 2 \cdot \text{tr} \left(\Re \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) \right).
\end{aligned} \tag{4.27}$$

When f and f_o are given, the statements of $\text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right)$ and $\text{tr} \left(\mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \right)$ are constants or coefficients. Therefore, the optimization

4.3 Proposed Method

problem on Eq. (4.15) can be rewritten as

$$\begin{aligned}
& \underset{\mathbf{T}_{x\{f\}}}{\text{minimize}} && -\text{tr} \left(\Re \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) \right) \\
& \text{subject to} && \mathbf{T}_{x\{f\}}^H \mathbf{T}_{x\{f\}} = \mathbf{I}_M \\
& && f_{\min} \leq f \leq f_o.
\end{aligned} \tag{4.28}$$

Obviously, the above minimization problem is equivalent to maximization problem as shown in Eq. (4.17). Although the proof of particular solution of Eq. (4.17) can be based on the CSS theorems whose proof can be found in [33, 34], however, there is no theory supporting that statement of $\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}}$ is in some sense matched to the matrix $\mathbf{R}_{xz\{f_o, f_o\}}$ where $f \neq f_o$; $\mathbf{S}_{\{f_o, f_o\}} \neq \mathbf{S}_{\{f, f_o\}}$. In order to find the feasible solution of Eq. (4.17) along with $\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}}$ and $\mathbf{R}_{xz\{f_o, f_o\}}$, we start by representing the matrices $\mathbf{R}_{xz\{f_o, f_o\}}$ and $\mathbf{R}_{xz\{f, f_o\}}$ into SVD forms only on the signal subspace, as follows:

$$\begin{aligned}
\mathbf{R}_{xz\{f_o, f_o\}} &= \mathbf{U}_{xz\{f_o, f_o\}} \mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f_o\}}^H, \\
\mathbf{R}_{xz\{f, f_o\}} &= \mathbf{U}_{xz\{f, f_o\}} \mathbf{\Sigma}_{xz\{f, f_o\}} \mathbf{V}_{xz\{f, f_o\}}^H,
\end{aligned} \tag{4.29}$$

where

$$\begin{aligned}
\mathbf{U}_{xz\{f_o, f_o\}} &= \mathbf{A}_x(\phi, f_o) \mathbf{F}_{xz\{f_o, f_o\}}^{-1}, \\
\mathbf{\Sigma}_{xz\{f_o, f_o\}} &= \mathbf{F}_{xz\{f_o, f_o\}} \mathbf{S}_{\{f_o, f_o\}} \mathbf{G}_{xz\{f_o, f_o\}}^H, \\
\mathbf{V}_{xz\{f_o, f_o\}} &= \mathbf{A}_z(\theta, f_o) \mathbf{G}_{xz\{f_o, f_o\}}^{-1}, \\
\mathbf{U}_{xz\{f, f_o\}} &= \mathbf{A}_x(\phi, f) \mathbf{F}_{xz\{f, f_o\}}^{-1}, \\
\mathbf{\Sigma}_{xz\{f, f_o\}} &= \mathbf{F}_{xz\{f, f_o\}} \mathbf{S}_{\{f, f_o\}} \mathbf{G}_{xz\{f, f_o\}}^H, \\
\mathbf{V}_{xz\{f, f_o\}} &= \mathbf{A}_z(\theta, f_o) \mathbf{G}_{xz\{f, f_o\}}^{-1}.
\end{aligned} \tag{4.30}$$

$\mathbf{F}_{xz\{f_o, f_o\}}, \mathbf{F}_{xz\{f, f_o\}}, \mathbf{G}_{xz\{f_o, f_o\}}, \mathbf{G}_{xz\{f, f_o\}} \in \mathbb{C}^{K \times K}$ are invertible matrices, $\mathbf{U}_{xz\{f_o, f_o\}}, \mathbf{U}_{xz\{f, f_o\}} \in \mathbb{C}^{M \times K}$, are the matrices of left singular vectors, $\mathbf{V}_{xz\{f_o, f_o\}}, \mathbf{V}_{xz\{f, f_o\}} \in \mathbb{C}^{M \times K}$ are the matrices of right singular vectors, and $\mathbf{\Sigma}_{xz\{f_o, f_o\}}, \mathbf{\Sigma}_{xz\{f, f_o\}} \in \mathbb{R}_{\geq 0}^{K \times K}$ are the diagonal matrices of singular values. From Lemma 1, rank of $\mathbf{\Sigma}_{xz\{f_o, f_o\}}$ and $\mathbf{\Sigma}_{xz\{f, f_o\}}$ can be considered as equal to K . Substituting Eqs. (4.29) and (4.30) into

4.3 Proposed Method

Eq. (4.17), a statement in the previous objective function can be replaced as

$$\begin{aligned}
& \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \\
&= \mathbf{T}_{x\{f\}} \left(\mathbf{U}_{xz\{f, f_o\}} \boldsymbol{\Sigma}_{xz\{f, f_o\}} \mathbf{V}_{xz\{f, f_o\}}^H \right) \\
&\quad \cdot \left(\mathbf{U}_{xz\{f_o, f_o\}} \boldsymbol{\Sigma}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f_o\}}^H \right)^H \\
&= \mathbf{T}_{x\{f\}} \left(\mathbf{U}_{xz\{f, f_o\}} \boldsymbol{\Sigma}_{xz\{f, f_o\}} \mathbf{V}_{xz\{f, f_o\}}^H \right) \\
&\quad \cdot \left(\mathbf{V}_{xz\{f_o, f_o\}} \boldsymbol{\Sigma}_{xz\{f_o, f_o\}}^H \mathbf{U}_{xz\{f_o, f_o\}}^H \right).
\end{aligned} \tag{4.31}$$

Note that $\mathbf{U}_{xz\{f_o, f_o\}}$, $\mathbf{U}_{xz\{f, f_o\}}$, $\mathbf{V}_{xz\{f_o, f_o\}}$, $\mathbf{V}_{xz\{f, f_o\}}$ have column of unitary property. Since the matrices $\mathbf{V}_{xz\{f_o, f_o\}}$ and $\mathbf{V}_{xz\{f, f_o\}}$ possess the property of the array manifold matrix $\mathbf{A}_z(\boldsymbol{\theta}, f_o)$, it is possible to shows $\mathbf{V}_{xz\{f_o, f_o\}}$ and $\mathbf{V}_{xz\{f, f_o\}}$ satisfy a property of $\mathbf{V}_{xz\{f_o, f_o\}}^H \mathbf{V}_{xz\{f, f_o\}} \approx \mathbf{I}_K$; but it strongly depends on an approximation algorithm of SVD. For instance, SVD approximation procedure based on Matlab[®] R2017a appears suitable for exhibiting such that property.

From the earlier property, the objective function in Eq. (4.17) can be given by

$$\begin{aligned}
& \text{tr} \left(\Re \left(\mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) \right) \\
&= \text{tr} \left(\Re \left(\mathbf{T}_{x\{f\}} \left(\mathbf{U}_{xz\{f, f_o\}} \boldsymbol{\Sigma}_{xz\{f, f_o\}} \boldsymbol{\Sigma}_{xz\{f_o, f_o\}}^H \mathbf{U}_{xz\{f_o, f_o\}}^H \right) \right) \right), \\
&= \text{tr} \left(\Re \left(\boldsymbol{\Sigma}_{xz\{f, f_o\}} \boldsymbol{\Sigma}_{xz\{f_o, f_o\}}^H \mathbf{U}_{xz\{f_o, f_o\}}^H \mathbf{T}_{x\{f\}} \mathbf{U}_{xz\{f, f_o\}} \right) \right).
\end{aligned} \tag{4.32}$$

In order to reach the maximum point on Eq. (4.28), a possible solution of $\mathbf{T}_{x\{f\}}$ is $\mathbf{U}_{xz\{f_o, f_o\}} \mathbf{U}_{xz\{f, f_o\}}^H$;

$$\begin{aligned}
& \text{tr} \left(\Re \left(\boldsymbol{\Sigma}_{xz\{f, f_o\}} \boldsymbol{\Sigma}_{xz\{f_o, f_o\}}^H \mathbf{U}_{xz\{f_o, f_o\}}^H \mathbf{T}_{x\{f\}} \mathbf{U}_{xz\{f, f_o\}} \right) \right) \\
&= \text{tr} \left(\boldsymbol{\Sigma}_{xz\{f_o, f_o\}} \boldsymbol{\Sigma}_{xz\{f, f_o\}} \right).
\end{aligned} \tag{4.33}$$

Overall, the error of transformation matrix $\mathbf{T}_{x\{f\}}$ is given by

$$\begin{aligned}
\varepsilon_{x\{f\}} &= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \right) \\
&\quad - 2 \cdot \text{tr} \left(\boldsymbol{\Sigma}_{xz\{f_o, f_o\}} \boldsymbol{\Sigma}_{xz\{f, f_o\}} \right).
\end{aligned} \tag{4.34}$$

It is possible to reduce complexity of two SVD operations by performing a single SVD operation on the matrix $\boldsymbol{\Psi}_{x\{f\}}$ at Eq. (4.19), which is resulted in Eq. (4.21). We can

4.3 Proposed Method

conclude that $\Sigma_{x_s\{f\}} \approx \Sigma_{xz\{f_o, f_o\}} \Sigma_{xz\{f, f_o\}}$ by the earlier property, and $\mathbf{V}_{x_s\{f\}} \mathbf{U}_{x_s\{f\}}^H$ is remarkably similar to $\mathbf{U}_{xz\{f_o, f_o\}} \mathbf{U}_{xz\{f, f_o\}}^H$.

Using such similar technique as in the proof of above, the solution of Eqs. (4.17) and (4.22) can be completed easily. The objective function in Eq. (4.16) is initially redefined as

$$\begin{aligned}
& \left\| \mathbf{R}_{xz\{f_o, f_o\}} - \mathbf{R}_{xz\{f_o, f\}} \mathbf{T}_{z\{f\}} \right\|_{\mathbb{F}}^2 \\
&= \text{tr} \left(\left(\mathbf{R}_{xz\{f_o, f_o\}} - \mathbf{R}_{xz\{f_o, f\}} \mathbf{T}_{z\{f\}} \right) \left(\mathbf{R}_{xz\{f_o, f_o\}} - \mathbf{R}_{xz\{f_o, f\}} \mathbf{T}_{z\{f\}} \right)^H \right) \\
&= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f_o, f\}} \mathbf{R}_{xz\{f_o, f\}}^H \right) \\
&\quad - \text{tr} \left(\mathbf{T}_{z\{f\}} \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}} + \mathbf{R}_{xz\{f_o, f_o\}}^H \mathbf{R}_{xz\{f_o, f\}} \mathbf{T}_{z\{f\}}^H \right) \\
&= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f_o, f\}} \mathbf{R}_{xz\{f_o, f\}}^H \right) \\
&\quad - \text{tr} \left(\mathbf{T}_{z\{f\}} \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}} + \left(\mathbf{T}_{z\{f\}} \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}} \right)^H \right) \\
&= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f_o, f\}} \mathbf{R}_{xz\{f_o, f\}}^H \right) \\
&\quad - 2 \cdot \text{tr} \left(\Re \left(\mathbf{T}_{z\{f\}} \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}} \right) \right).
\end{aligned} \tag{4.35}$$

The optimization problem on Eq. (4.16) can be rewritten as

$$\begin{aligned}
& \underset{\mathbf{T}_{z\{f\}}}{\text{minimize}} && -\text{tr} \left(\Re \left(\mathbf{T}_{z\{f\}} \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}} \right) \right) \\
& \text{subject to} && \mathbf{T}_{z\{f\}}^H \mathbf{T}_{z\{f\}} = \mathbf{I}_M \\
& && f_{\min} \leq f \leq f_o.
\end{aligned} \tag{4.36}$$

Similar to Eq. (4.28), the above minimization problem is equivalent to the maximization problem as shown in Eq. (4.18). To find the solution of Eq. (4.18), we transform $\mathbf{R}_{xz\{f_o, f\}}$ into SVD forms, as follows

$$\mathbf{R}_{xz\{f_o, f\}} = \mathbf{U}_{xz\{f_o, f\}} \Sigma_{xz\{f_o, f\}} \mathbf{V}_{xz\{f_o, f\}}^H, \tag{4.37}$$

where

$$\begin{aligned}
\mathbf{U}_{xz\{f_o, f\}} &= \mathbf{A}_x(\phi, f_o) \mathbf{F}_{xz\{f_o, f\}}^{-1}, \\
\Sigma_{xz\{f_o, f\}} &= \mathbf{F}_{xz\{f_o, f\}} \mathbf{S}_{\{f_o, f\}} \mathbf{G}_{xz\{f_o, f\}}^H, \\
\mathbf{V}_{xz\{f_o, f\}} &= \mathbf{A}_z(\theta, f) \mathbf{G}_{xz\{f_o, f\}}^{-1}.
\end{aligned} \tag{4.38}$$

4.3 Proposed Method

From above, $\mathbf{F}_{xz\{f_o, f\}}, \mathbf{G}_{xz\{f_o, f\}} \in \mathbb{C}^{K \times K}$ are invertible matrices, $\mathbf{U}_{xz\{f_o, f\}} \in \mathbb{C}^{M \times K}$, is the matrices of left singular vectors, $\mathbf{V}_{xz\{f_o, f\}} \in \mathbb{C}^{M \times K}$ is the matrices of right singular vectors, and $\mathbf{\Sigma}_{xz\{f_o, f\}} \in \mathbb{R}_{\geq 0}^{K \times K}$ is the diagonal matrices of singular values. In case of SVD of $\mathbf{R}_{xz\{f_o, f_o\}}$, we borrow the notations from Eqs. (4.29) and (4.30). Note that a rank of $\mathbf{\Sigma}_{xz\{f_o, f\}}$ is equal to K by employing Lemma 1. Substituting Eqs. (4.29), (4.30), (4.37) and (4.38) into Eq. (4.18), the statement in the objective function can be given as

$$\begin{aligned}
& \mathbf{T}_{z\{f\}} \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}} \\
&= \mathbf{T}_{z\{f\}} \left(\mathbf{U}_{xz\{f_o, f\}} \mathbf{\Sigma}_{xz\{f_o, f\}} \mathbf{V}_{xz\{f_o, f\}}^H \right)^H \\
&\quad \cdot \left(\mathbf{U}_{xz\{f_o, f_o\}} \mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f_o\}}^H \right) \\
&= \mathbf{T}_{x\{f\}} \left(\mathbf{V}_{xz\{f_o, f\}} \mathbf{\Sigma}_{xz\{f_o, f\}}^H \mathbf{U}_{xz\{f_o, f\}}^H \right) \\
&\quad \cdot \left(\mathbf{U}_{xz\{f_o, f_o\}} \mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f_o\}}^H \right).
\end{aligned} \tag{4.39}$$

Based on the definition on Eqs. (4.29), (4.30), (4.37) and (4.38), the matrices $\mathbf{U}_{xz\{f_o, f_o\}}$ and $\mathbf{U}_{xz\{f_o, f\}}$ possess the property of array manifold matrix $\mathbf{A}_x(\phi, f_o)$. It suffices to show that $\mathbf{U}_{xz\{f_o, f_o\}}$ and $\mathbf{U}_{xz\{f_o, f\}}$ satisfy a property of $\mathbf{U}_{xz\{f_o, f\}}^H \mathbf{U}_{xz\{f_o, f_o\}} \approx \mathbf{I}_K$. Hence, the objective function in Eq. (4.18) can be expressed as

$$\begin{aligned}
& \text{tr} \left(\Re \left(\mathbf{T}_{z\{f\}} \mathbf{R}_{xz\{f_o, f\}}^H \mathbf{R}_{xz\{f_o, f_o\}} \right) \right) \\
&= \text{tr} \left(\Re \left(\mathbf{T}_{z\{f\}} \left(\mathbf{V}_{xz\{f_o, f\}} \mathbf{\Sigma}_{xz\{f_o, f\}}^H \mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f_o\}}^H \right) \right) \right), \\
&= \text{tr} \left(\Re \left(\mathbf{\Sigma}_{xz\{f_o, f\}}^H \mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f_o\}}^H \mathbf{T}_{z\{f\}} \mathbf{V}_{xz\{f_o, f\}} \right) \right).
\end{aligned} \tag{4.40}$$

A possible solution to reach the maximum point of Eq. (4.36) is $\mathbf{V}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f\}}^H$.

We have

$$\begin{aligned}
& \text{tr} \left(\Re \left(\mathbf{\Sigma}_{xz\{f_o, f\}}^H \mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f_o\}}^H \mathbf{T}_{z\{f\}} \mathbf{V}_{xz\{f_o, f\}} \right) \right) \\
&= \text{tr} \left(\mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{\Sigma}_{xz\{f_o, f\}} \right),
\end{aligned} \tag{4.41}$$

where the error of matrix transformation $\mathbf{T}_{z\{f\}}$ is given by

$$\begin{aligned}
\varepsilon_{x\{f\}} &= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f_o, f\}} \mathbf{R}_{xz\{f_o, f\}}^H \right) \\
&\quad - 2 \cdot \text{tr} \left(\mathbf{\Sigma}_{xz\{f_o, f_o\}} \mathbf{\Sigma}_{xz\{f_o, f\}} \right).
\end{aligned} \tag{4.42}$$

4.3 Proposed Method

It is possible to show that $\Sigma_{z_s\{f\}} \approx \Sigma_{xz\{f_o, f_o\}} \Sigma_{xz\{f_o, f\}}$ by the previously property as applied to Eq. (4.40). Thus, $\mathbf{V}_{z_s\{f\}} \mathbf{U}_{z_s\{f\}}^H$ is notably similar to $\mathbf{V}_{xz\{f_o, f_o\}} \mathbf{V}_{xz\{f_o, f\}}^H$.

All things considered, the remains in Eqs. (4.21) to (4.23) are summarized as follows:

$$\begin{aligned} \Sigma_{x_s\{f\}} &= \mathbf{F}_{xz\{f, f_o\}} \mathbf{S}_{\{f, f_o\}} \mathbf{G}_{xz\{f, f_o\}}^H \\ &\quad \mathbf{G}_{xz\{f_o, f_o\}} \mathbf{S}_{\{f_o, f_o\}}^H \mathbf{F}_{xz\{f_o, f_o\}}^H, \\ \mathbf{P}_{x\{f\}} &= \mathbf{F}_{xz\{f, f_o\}}^{-1}, \\ \mathbf{Q}_{x\{f\}} &= \mathbf{F}_{xz\{f_o, f_o\}}^{-1}, \end{aligned} \tag{4.43}$$

and

$$\begin{aligned} \Sigma_{z_s\{f\}} &= \mathbf{G}_{xz\{f_o, f\}} \mathbf{S}_{\{f_o, f\}}^H \mathbf{F}_{xz\{f_o, f\}}^H \\ &\quad \mathbf{F}_{xz\{f_o, f_o\}} \mathbf{S}_{\{f_o, f_o\}} \mathbf{G}_{xz\{f_o, f_o\}}^H, \\ \mathbf{P}_{z\{f\}} &= \mathbf{G}_{xz\{f_o, f\}}^{-1}, \\ \mathbf{Q}_{z\{f\}} &= \mathbf{G}_{xz\{f_o, f_o\}}^{-1}. \end{aligned} \tag{4.44}$$

The proof is completed. ■

Theorem 1 shows that $\mathbf{T}_{x\{f\}}$, $\mathbf{T}_{z\{f\}}$ are constructed by performing SVD of $\Psi_{x\{f\}}$ and $\Psi_{z\{f\}}$. It can be seen that the process of initial DOA estimation, such as beamforming technique, is not longer required. Although exploiting the singular vectors of $\Psi_{x\{f\}}$, $\Psi_{z\{f\}}$ from distinct subarrays result in a better performance than the identical subarray [33, 34]; this is explained by the existence of $\mathbf{W}_{xx\{f, f'\}}$, $\mathbf{W}_{zz\{f, f'\}}$ as shown in Eqs. (4.7), (5.5) and (6.2). The error of transformations may remain consistent with following equations:

$$\begin{aligned} \varepsilon_{x\{f\}} &= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f, f_o\}} \mathbf{R}_{xz\{f, f_o\}}^H \right) \\ &\quad - 2 \cdot \text{tr} \left(\Sigma_{x_s\{f\}} \right), \\ \varepsilon_{z\{f\}} &= \text{tr} \left(\mathbf{R}_{xz\{f_o, f_o\}} \mathbf{R}_{xz\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{R}_{xz\{f_o, f\}} \mathbf{R}_{xz\{f_o, f\}}^H \right) \\ &\quad - 2 \cdot \text{tr} \left(\Sigma_{z_s\{f\}} \right). \end{aligned} \tag{4.45}$$

We relax the error constraints in the hope of arriving at a reduction in computation [33, 34, 37], but this is sufficient for estimating $\mathbf{T}_{x\{f\}}$, $\mathbf{T}_{z\{f\}}$ without loss of generality.

For the full statement, see 4.3.1.

4.3 Proposed Method

To successfully apply this class of transformation matrices into Eq. (4.9), the propositions of

$$\begin{aligned}\mathbf{Q}_{x\{f\}} \mathbf{P}_{x\{f\}}^H \mathbf{A}_x^H(\phi, f) \mathbf{A}_x(\phi, f) &= \mathbf{I}_K, \\ \mathbf{A}_z^H(\theta, f) \mathbf{A}_z(\theta, f) \mathbf{P}_{z\{f\}} \mathbf{Q}_{z\{f\}}^H &= \mathbf{I}_K,\end{aligned}\tag{4.46}$$

are generally required by Eqs. (4.21) to (4.23) and (6.4). However, it is possible to show that those propositions are not necessary, which the next section will address further.

4.3.2 Extension Properties on SVD of the Correlation Matrix

In this section, we continue to present a special property on SVD of $\mathbf{R}_{xz\{f,f\}}$, $\mathbf{R}_{xz\{f,f_o\}}$, $\mathbf{R}_{xz\{f_o,f\}}$ and $\mathbf{R}_{xz\{f_o,f_o\}}$, in other to show that previous propositions are not necessary to satisfy. Considering SVD to $\mathbf{R}_{xz\{f,f'\}}$, a generalization form of the matrix $\mathbf{R}_{xz\{f,f'\}}$ on signal subspace is given as

$$\mathbf{R}_{xz\{f,f'\}} = \mathbf{U}_{xz\{f,f'\}} \mathbf{\Sigma}_{xz\{f,f'\}} \mathbf{V}_{xz\{f,f'\}}^H,\tag{4.47}$$

where $\mathbf{U}_{xz\{f,f'\}}$, $\mathbf{V}_{xz\{f,f'\}} \in \mathbb{C}^{M \times K}$, and $\mathbf{\Sigma}_{xz\{f,f'\}} \in \mathbb{R}_{\geq 0}^{K \times K}$ are the matrices of left and right singular vectors, and the diagonal matrices of singular values, respectively. According to the past studies [12–15, 17, 28], it is possible to define the matrices $\mathbf{U}_{xz\{f,f'\}}$, $\mathbf{V}_{xz\{f,f'\}}$ and $\mathbf{\Sigma}_{xz\{f,f'\}}$ as following equations:

$$\begin{aligned}\mathbf{\Sigma}_{xz\{f,f'\}} &= \mathbf{F}_{xz\{f,f'\}} \mathbf{S}_{\{f,f'\}} \mathbf{G}_{xz\{f,f'\}}^H, \\ \mathbf{U}_{xz\{f,f'\}} &= \mathbf{A}_x(\phi, f) \mathbf{F}_{xz\{f,f'\}}^{-1}, \\ \mathbf{V}_{xz\{f,f'\}} &= \mathbf{A}_z(\theta, f') \mathbf{G}_{xz\{f,f'\}}^{-1}.\end{aligned}\tag{4.48}$$

From above, $\mathbf{F}_{xz\{f,f'\}}$, $\mathbf{G}_{xz\{f,f'\}} \in \mathbb{C}^{K \times K}$ is invertible and possess the unitary property via SVD theorem. From the above definition, we hypothesize that $\mathbf{F}_{xz\{f,f'\}}$ possess a property of array manifold matrix $\mathbf{A}_x(\phi, f)$, and similarly, $\mathbf{G}_{xz\{f,f'\}}$ possess a property of array manifold matrix $\mathbf{A}_z(\theta, f')$. To support such hypothesis, we start by redefining the left matrix of generalization form of $\mathbf{R}_{xz\{f,f\}}$;

$$\mathbf{U}_{xz\{f,f\}} = \mathbf{A}_x(\phi, f) \mathbf{F}_{xz\{f,f\}}^{-1}.\tag{4.49}$$

4.3 Proposed Method

It is possible to show that

$$\mathbf{U}_{xz\{f,f'\}} \mathbf{U}_{xz\{f,f\}}^H = \mathbf{I}_K, \quad (4.50)$$

$$\mathbf{U}_{xz\{f,f'\}} \mathbf{U}_{xz\{f,f\}}^\dagger = \mathbf{I}_K, \quad (4.51)$$

by employing a provably good approximation algorithm of SVD. For example, SVD approximation procedure based on Matlab[®] R2017a appears suitable for exhibiting the property on Eqs. (4.50) and (4.51). Therefore, we can assert that unitary property is valid by utilizing the suitable approximation algorithm. Additionally, the following statement is revealed that matrix $\mathbf{F}_{xz\{f,f'\}}$ can possess the property of array manifold matrix $\mathbf{A}_x(\boldsymbol{\phi}, f)$ by Eq. (4.50);

$$\mathbf{A}_x^H(\boldsymbol{\phi}, f) \mathbf{A}_x(\boldsymbol{\phi}, f) = \mathbf{F}_{xz\{f,f\}}^H \mathbf{F}_{xz\{f,f'\}}. \quad (4.52)$$

From Eqs. (4.51) and (4.52), we additionally found that

$$\mathbf{F}_{xz\{f,f\}} = \mathbf{F}_{xz\{f,f'\}}. \quad (4.53)$$

Using similar technique from the above on the right matrix of generalization form of $\mathbf{R}_{xz\{f',f'\}}$, we have

$$\mathbf{A}_z^H(\boldsymbol{\theta}, f) \mathbf{A}_z(\boldsymbol{\theta}, f) = \mathbf{G}_{xz\{f',f'\}}^H \mathbf{G}_{xz\{f',f\}}. \quad (4.54)$$

and

$$\mathbf{G}_{xz\{f',f'\}} = \mathbf{G}_{xz\{f,f'\}}. \quad (4.55)$$

Since the relation between $\mathbf{F}_{xz\{f,f\}}$ and $\mathbf{F}_{xz\{f,f'\}}$, and the relation between $\mathbf{G}_{xz\{f',f'\}}$ and $\mathbf{G}_{xz\{f,f'\}}$ have been found, those propositions are no longer necessary.

After Theorem 1 is defined, we can proceed to describe a concept to construct a wideband cross-correlation matrix on Eq. (4.9) with the proposed transformation

4.3 Proposed Method

matrices;

$$\begin{aligned}
& \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f,f\}} \mathbf{T}_{z\{f\}}^H \\
&= \mathbf{T}_{x\{f\}} \mathbf{A}_x(\phi, f) \mathbf{S}_{\{f,f\}} \mathbf{A}_z^H(\theta, f) \mathbf{T}_{z\{f\}}^H \\
&= \mathbf{A}_x(\phi, f_o) \mathbf{F}_{xz\{f_o, f_o\}}^{-1} \mathbf{F}_{xz\{f,f\}} \mathbf{S}_{\{f,f\}} \mathbf{G}_{xz\{f,f\}}^H \\
&\quad \mathbf{G}_{xz\{f_o, f_o\}}^{-H} \mathbf{A}_z^H(\theta, f_o) \\
&= \mathbf{U}_{xz\{f_o, f_o\}} \mathbf{\Sigma}_{xz\{f,f\}} \mathbf{V}_{xz\{f_o, f_o\}}^H.
\end{aligned} \tag{4.56}$$

The proof is obtained easily by performing straightforward calculation. It is apparently seen from Eq. (4.56) that matrices of singular vectors $\mathbf{U}_{xz\{f_o, f_o\}}$ and $\mathbf{V}_{xz\{f_o, f_o\}}$ no longer depend on any frequency f , but only depend on the reference frequency f_o , moreover, the diagonal matrix of singular values $\mathbf{\Sigma}_{xz\{f,f\}}$ completely depends on the temporal frequency f ; this advantage of independent isolation between f_o and f allows $\mathbf{T}_{x\{f\}}$ and $\mathbf{T}_{z\{f\}}$ to transform the sample cross-correlation matrix from any frequency f into the reference frequency f_o without loss of generality, which will be clearly revealed in the next section.

4.3.3 DOA Estimation Scheme

In this section, we give a scheme of wideband DOA estimation. We start by constructing cross-correlation matrices from Eq. (4.56) and performing integration, as follows

$$\begin{aligned}
\mathbf{R}_{xz} &= \int_{f_{\min}}^{f_o} \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f,f\}} \mathbf{T}_{z\{f\}}^H df \\
&= \int_{f_{\min}}^{f_o} \mathbf{U}_{xz\{f_o, f_o\}} \mathbf{\Sigma}_{xz\{f,f\}} \mathbf{V}_{xz\{f_o, f_o\}}^H df \\
&= \mathbf{U}_{xz\{f_o, f_o\}} \mathbf{\Omega} \mathbf{V}_{xz\{f_o, f_o\}}^H,
\end{aligned} \tag{4.57}$$

where

$$\mathbf{\Omega} = \int_{f_{\min}}^{f_o} \mathbf{\Sigma}_{xz\{f,f\}} df. \tag{4.58}$$

From above, $\mathbf{\Omega} \in \mathbb{R}_{\geq 0}^{K \times K}$ represents the summation matrix of singular values for all frequency range between f_{\min} and f_o . To show that matrix $\mathbf{\Omega}$ can be considered as

4.3 Proposed Method

indicator of $\int_{f_{\min}}^{f_o} \mathbf{S}_{\{f,f\}} df$, we introduce the following useful lemma;

Lemma 2. *Given Lemma 1 and Assumption 4 along with frequency range between f_{\min} and f_o , the matrix $\mathbf{\Omega}$ is nonsingular if and only if there exists some nonsingular matrix $\mathbf{S}_{\{f,f\}}$ at some point in frequency f .*

Proof. In this appendix, we provide a mathematical reason to support Lemma 2. We start by recalling the following matrix [37]: $\mathbf{S} = \int_{f_{\min}}^{f_o} \mathbf{S}_{\{f,f\}} df$. It is well known that \mathbf{S} may be singular for all frequency bins within range between f_{\min} and f_o if and only if the sources are correlated somehow each other [37]. Indeed, if any $\mathbf{S}_{\{f,f\}}$ are nonsingular, then \mathbf{S} is directly nonsingular. Therefore, \mathbf{S} can be considered as nonsingular matrix for the most practical situations.

Suppose that the sample matrix $\mathbf{R}_{xz\{f,f\}} \in \mathbb{C}^{M \times M}$ on Eq. (4.47) has a rank of K by employing Lemma 1 and SVD theorem, then the matrix of singular values $\mathbf{\Sigma}_{xz\{f,f\}} \in \mathbb{C}^{K \times K}$ has a positive determinant. As the facts that a square matrix is singular if and only if its determinant is zero; therefore, the singular values $\mathbf{\Sigma}_{xz\{f,f\}}$ have to satisfy the condition of

$$\mathbf{\Sigma}_{xz\{f,f\}} > \mathbf{O}_{K \times K}. \quad (4.59)$$

For making more practical implementation, we can have

$$\mathbf{\Sigma}_{xz\{f,f\}} \geq \eta \mathbf{I}_K > \mathbf{O}_{K \times K}, \quad (4.60)$$

where $\eta \in \mathbb{R}_{>0}$ is a threshold limit value. Note that the threshold value vary and depend on the maximum peak-to-peak amplitude of signal sources; to adjust the threshold value, a pre-processing technique on the received signals may be required such as normalization or pre-amplification. Summing $\mathbf{\Sigma}_{xz\{f,f\}}$ for all f , the inequality in Eq. (4.60) indicates $\mathbf{\Omega}$ can be nonsingular if there exist some nonsingular matrix $\mathbf{\Sigma}_{xz\{f,f\}}$ at some temporal frequency f . ■

4.3 Proposed Method

The matrix \mathbf{R}_{xz} is now able to perform SVD without degrading the left and right matrices singular vectors $\mathbf{U}_{xz\{f_o, f_o\}}$ and $\mathbf{V}_{xz\{f_o, f_o\}}$ by employing Lemma 2. Furthermore, \mathbf{R}_{xz} is compatible with recent narrowband subspace methods [3, 5, 12, 13, 17, 21, 31, 36]. According to useful techniques [3, 17, 28, 31], we form new matrices $\mathbf{\Gamma}_x$ and $\mathbf{\Gamma}_z$ by Eq. (4.57) as follows:

$$\begin{aligned}\mathbf{\Gamma}_x &= \mathbf{U}_{xz_1\{f_o, f_o\}}^\dagger \mathbf{U}_{xz_2\{f_o, f_o\}} \\ &= \mathbf{F}_{xz\{f_o, f_o\}} \mathbf{\Phi}_x \mathbf{F}_{xz\{f_o, f_o\}}^{-1}, \\ \mathbf{\Gamma}_z &= \mathbf{V}_{xz_1\{f_o, f_o\}}^\dagger \mathbf{V}_{xz_2\{f_o, f_o\}} \\ &= \mathbf{G}_{xz\{f_o, f_o\}} \mathbf{\Phi}_z \mathbf{G}_{xz\{f_o, f_o\}}^{-1},\end{aligned}\tag{4.61}$$

where

$$\begin{aligned}\mathbf{\Phi}_x &= \text{diag}(\alpha_x(\phi_1, f_o), \alpha_x(\phi_2, f_o), \dots, \alpha_x(\phi_K, f_o)), \\ \mathbf{\Phi}_z &= \text{diag}(\alpha_z(\theta_1, f_o), \alpha_z(\theta_2, f_o), \dots, \alpha_z(\theta_K, f_o)),\end{aligned}\tag{4.62}$$

$\mathbf{U}_{xz_1\{f_o, f_o\}}, \mathbf{U}_{xz_2\{f_o, f_o\}}, \mathbf{V}_{xz_1\{f_o, f_o\}}, \mathbf{V}_{xz_2\{f_o, f_o\}} \in \mathbb{C}^{M-1 \times K}$ are the first and the last $M - 1$ rows of singular matrices $\mathbf{U}_{xz\{f_o, f_o\}}$ and $\mathbf{V}_{xz\{f_o, f_o\}}$, respectively, Applying Eigenvalue decomposition (EVD), the x and z subarray angles are estimated by the eigenvalues matrices $\mathbf{\Lambda}_x, \mathbf{\Lambda}_z$ of $\mathbf{\Gamma}_x, \mathbf{\Gamma}_z$;

$$\begin{aligned}\phi_k &= \cos^{-1} \left(\text{angle}(\lambda_{x_k}) \frac{\lambda}{2\pi d} \right), \\ \theta_k &= \cos^{-1} \left(\text{angle}(\lambda_{z_k}) \frac{\lambda}{2\pi d} \right),\end{aligned}\tag{4.63}$$

where $\lambda_{x_k}, \lambda_{z_k} \in \mathbb{C}$ are the k^{th} eigenvalue in $\mathbf{\Lambda}_x, \mathbf{\Lambda}_z$.

All thing considered, the following steps summarize the wideband 2D DOA estimation using the proposed method:

1. Construct sample matrix $\mathbf{R}_{xz\{f_o, f_o\}}$ on Eq. (6.2).
2. Given f between f_{\min} and f_o , construct sample matrices $\mathbf{R}_{xz\{f, f_o\}}, \mathbf{R}_{xz\{f_o, f\}}, \mathbf{R}_{xz\{f, f\}}$ on Eq. (6.2).
3. Construct $\mathbf{\Psi}_{x\{f\}}, \mathbf{\Psi}_{z\{f\}}$ in Eqs. (4.19) and (4.20) using previous matrices, perform SVD of $\mathbf{\Psi}_{x\{f\}}, \mathbf{\Psi}_{z\{f\}}$ for formulating $\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}}$ on Eqs. (4.21) and (4.22).

4.4 Computational Complexity

4. Cumulate $\mathbf{T}_{x\{f\}}$, $\mathbf{T}_{z\{f\}}$, $\mathbf{R}_{xz\{f,f\}}$ and construct new cross-correlation matrix \mathbf{R}_{xz} via Eq. (4.57).
5. Estimate $\mathbf{U}_{xz\{f_o, f_o\}}$, $\mathbf{V}_{xz\{f_o, f_o\}}$ by performing SVD of \mathbf{R}_{xz} in Eq. (4.57), and construct $\mathbf{\Gamma}_x$, $\mathbf{\Gamma}_z$ using Eq. (6.7).
6. Perform EVD of $\mathbf{\Gamma}_x$, $\mathbf{\Gamma}_z$, estimate angles by Eq. (6.9).
7. Calculate azimuth angle using Eq. (4.1).

4.4 Computational Complexity

Computational complexity of the proposed method is discussed on a worst-case scenario; the number of computations for an $n \times m$ matrix addition is $\mathcal{O}(nm)$, $n \times m$ and $m \times p$ matrix multiplication is $\mathcal{O}(nmp)$, an $n \times n$ matrix inversion is $\mathcal{O}(n^3)$, EVD of an $n \times n$ matrix is $\mathcal{O}(n^3)$, EVD of an $n \times n$ matrix with eigenvectors is $\mathcal{O}(2n^3)$ where $m \geq n$, SVD of an $m \times n$ matrix is $\mathcal{O}(mn^2)$, SVD of an $m \times n$ matrix with singular vectors is $\mathcal{O}(4m^2n + 22n^3)$, and an $m \times n$ matrix pseudo inversion via SVD is $\mathcal{O}(4m^2n + 22n^3 + m^2n + mn^2 + n)$ where $m \geq n$.

We provided computational complexity of the proposed method in comparison with following methods: IMUSIC [25]; CSS [37]; CSS based on a modification to the total least-square (CSS-MTLS) [33]; TOPS [40]; squared test of TOPS (Squared-TOPS) [18]; weighted squared TOPS (WS-TOPS) [8]; TOFS [41]; and CSS based on the property of Gramian matrix of the array manifold matrix (CSS-PGAM) [28]. Fig. 4.1 show the comparison of the computational complexities which consist of four investigations; (a) changing the number of snapshots where $F = 300, M = 12, K = 3$, (b) frequency bins caused by STFT where $M = 12, N = 48000, K = 3$, (c) microphone elements each subarray where $F = 300, N = 48000, K = 3$, and (d) number of sources where $F = 300, M = 12, N = 48000$. For all selected methods excepting CSS-PGAM and the

4.4 Computational Complexity

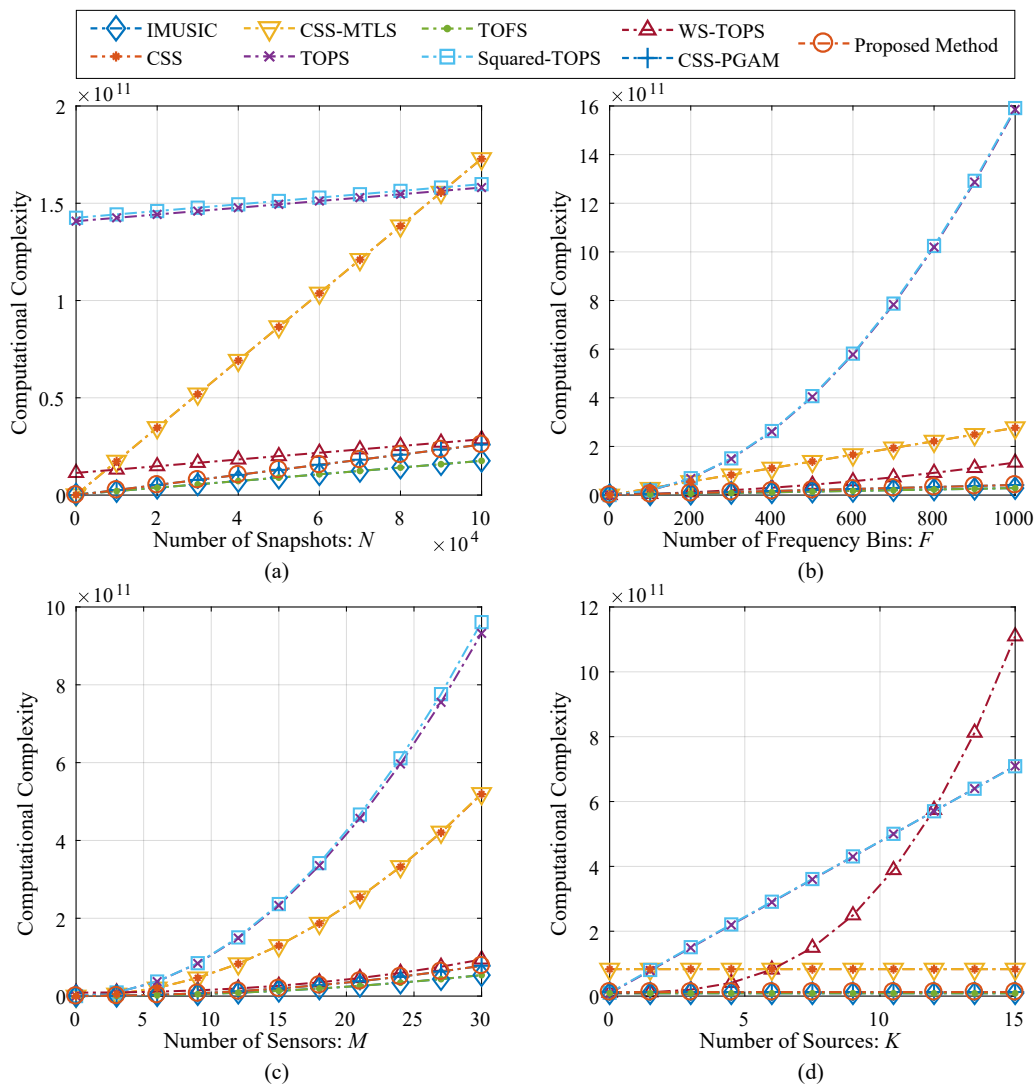


Fig. 4.1 Computational complexities; (a) changing the number of snapshots N , (b) changing the number of frequency bins F , (c) changing the number of microphone elements each subarray M , and (d) changing the number of incident sources K .

proposed method, the number of steering angle is 1800. The number of iterations in CSS and CSS-MTLS are 10. The number of selecting frequency on WS-TOPS is assumed to be 1 for intuitive comparison. Note that a computational burden of peak searching algorithm are not concerned in this study. It is apparently seen in Fig. 4.1 (a) to (d) that the proposed method exhibits remarkably less computational times than other methods, which are presented more high growth rates than the proposed method, such as CSS,

4.5 Numerical Simulations

CSS-MTLS, TOPS, Squared-TOPS, and WS-TOPS for the case (d). To the best of our knowledge, an explanation for the high growth rates is that SVD of large-scale matrices are intensively employed in TOPS, Squared-TOPS, and WS-TOPS [8, 18, 40]. Additionally, the computational complexities of CSS and CSS-MTLS are significantly dominated by the number of iterations, which is directly effecting on the DOA estimation performance [33, 37].

Despite the fact that IMUSIC and TOFS exhibit low computational burden than the proposed method, considering the wideband 2D DOA estimation performance which will be revealed in the next section, we can see that the increase of few computational time can provide significantly improvement of the DOA estimation.

4.5 Numerical Simulations

In this section, the performance of proposed method is demonstrated via computer simulations. The tested methods were as previously described, excepting CSS and CSS-MTLS; this is because unintended biases, causing by the initial DOA estimation step as shown in [33, 37], should be taken into consideration to other candidate methods. Computer simulations were carried out using Matlab[®] R2017a. Simulation parameters were chosen as follows: the number of Monte Carlo trials each scenario was 100 runs, the sampling frequency was 48 kHz, the output of each microphone was captured 1 second or 48000 snapshots number, the speed of sound c was 340 m/s, the spacing of microphone elements d was 5 cm, the minimum and reference frequency f_{\min}, f_o were 0.5 and 3.4 kHz. Note that the sources were voices containing the frequency range from 0.1 to 16 kHz.

The following basic application scenarios are considered in this study: (1) a performance of selected methods and the proposed method with respect to the number of

4.5 Numerical Simulations

Table 4.1 Performance evaluations of Scenario 1 via a computer simulation; $K = 2, 3$

Incident Sources		MAE of DOAs (Degree)							
Number	Position	Angle	IMUSIC	TOPS	TOFS	Squared-TOPS	WS-TOPS	CSS-PGAM	Proposed Method
2	ϕ_1	68.766	0.011	0.372	0.011	0.011	0.011	0.253	0.076
	θ_1	53.000	0.288	0.523	0.288	0.286	0.288	0.960	0.946
	ϕ_2	50.371	0.050	0.571	0.050	0.088	0.050	0.593	0.517
	θ_2	37.000	0.242	1.382	0.242	0.254	0.242	0.586	0.592
Total Average			0.148	0.712	0.148	0.159	0.148	0.598	0.533
3	ϕ_1	75.523	0.108	0.555	0.108	2.893	0.108	0.476	0.232
	θ_1	30.000	1.513	3.165	0.940	2.005	0.837	0.132	0.119
	ϕ_2	60.000	0.412	0.471	0.412	1.071	0.412	0.785	0.408
	θ_2	45.000	0.309	0.424	0.309	0.298	0.298	0.107	0.124
	ϕ_3	41.410	0.416	0.861	0.416	1.864	0.416	0.928	0.529
	θ_3	60.000	0.985	3.019	0.412	2.372	0.985	0.285	0.281
	Total Average			0.624	1.416	0.433	1.751	0.509	0.452

4.5 Numerical Simulations

Table 4.2 Performance evaluations of Scenario 1 via a computer simulation; $K = 4$

Incident Sources		MAE of DOAs (Degree)							
Number	Position	Angle	IMUSIC	TOPS	TOFS	Squared-TOPS	WS-TOPS	CSS-PGAM	Proposed Method
	ϕ_1	85.166	0.204	1.148	0.204	0.285	31.640	0.711	0.366
	θ_1	45.000	0.309	0.331	0.309	2.414	16.844	0.479	0.134
	ϕ_2	64.911	13.918	13.299	8.841	14.032	32.837	5.519	0.818
	θ_2	19.000	19.000	19.000	19.000	18.078	19.000	3.479	3.357
4	ϕ_3	72.612	1.548	0.544	0.993	0.475	27.135	0.590	0.384
	θ_3	30.000	6.818	6.480	7.586	8.283	15.945	2.705	2.435
	ϕ_4	50.727	50.727	50.727	31.552	50.222	50.727	1.079	0.551
	θ_4	80.000	0.214	0.243	0.214	5.497	29.821	1.378	0.417
Total Average			11.592	11.471	8.587	12.411	27.994	1.992	1.058

4.5 Numerical Simulations

Table 4.3 Performance evaluations of Scenario 1 via a computer simulation; $K = 5$

Incident Sources		MAE of DOAs (Degree)									
Number	Position	Angle	IMUSIC	TOPS	TOFS	Squared-TOPS	WS-TOPS	CSS-PGAM	Proposed Method		
5	ϕ_1	74.811	0.900	6.821	31.817	4.996	18.915	22.281	2.523		
	θ_1	44.000	22.228	20.744	22.801	19.197	26.731	0.848	0.900		
	ϕ_2	72.771	8.026	9.390	1.692	8.301	24.419	5.553	0.814		
	θ_2	64.000	0.399	0.508	0.171	0.933	16.312	0.652	0.651		
	ϕ_3	45.694	33.570	38.867	3.610	32.284	36.400	3.370	0.815		
	θ_3	88.000	0.236	0.424	0.236	0.803	22.178	0.214	0.200		
	ϕ_4	12.163	12.163	11.945	1.122	11.943	12.163	5.160	0.437		
	θ_4	50.000	5.309	6.956	5.309	6.810	17.341	0.491	0.489		
	ϕ_5	65.288	20.597	22.064	2.148	19.795	31.627	5.846	0.412		
	θ_5	20.000	20.000	19.864	20.000	20.000	20.000	2.312	2.316		
Total Average			12.343	13.758	8.891	12.506	22.609	4.673	0.955		

4.5 Numerical Simulations

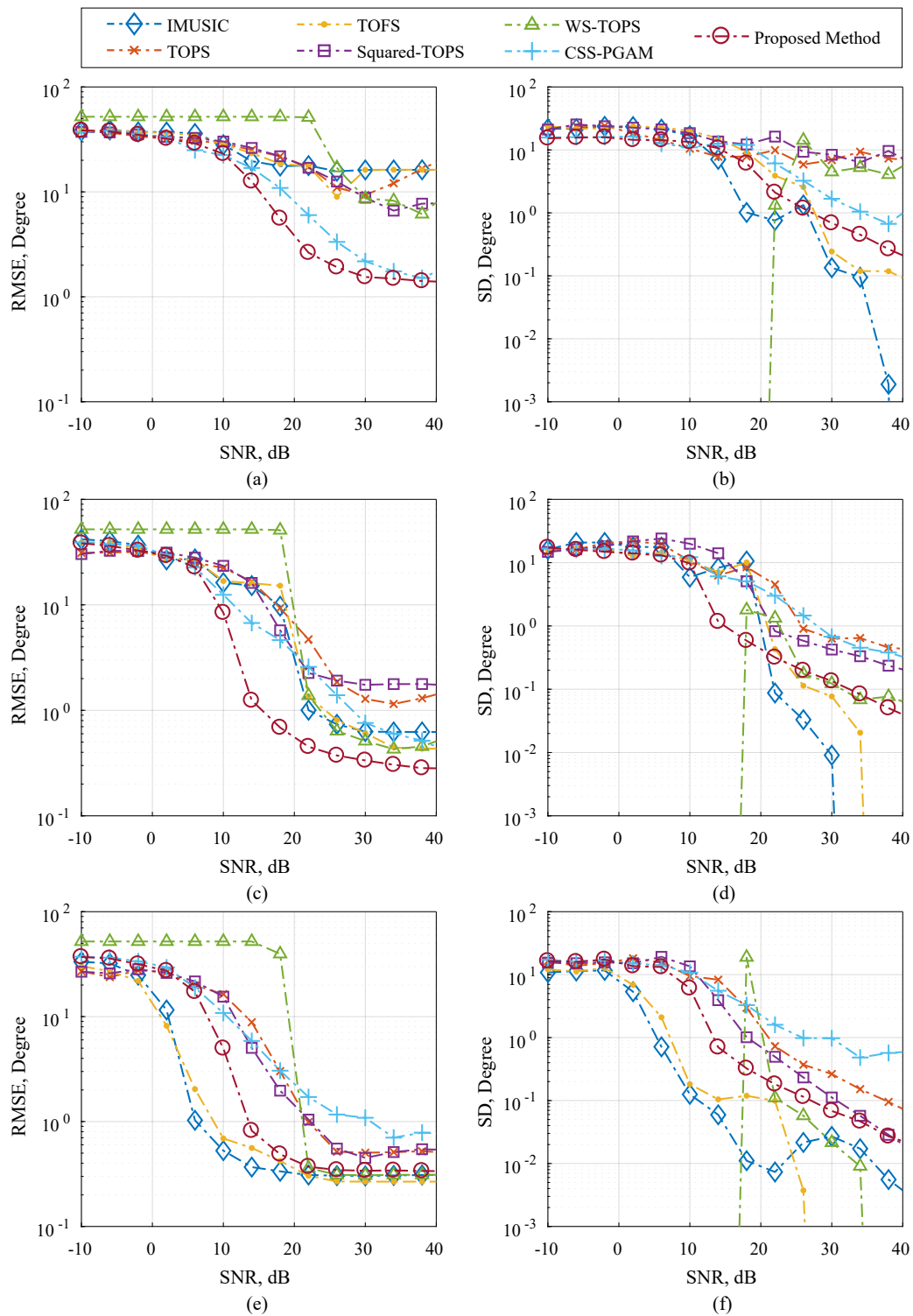


Fig. 4.2 Performance evaluations of Scenario 2 via computer simulation; (a), (c), (e) RMSE versus SNR, (b), (d), (f) SD versus SNR, where the number of microphone elements each subarray on (a), (b) $M = 6$, (c), (d) $M = 10$, and (e), (f) $M = 14$.

4.5 Numerical Simulations

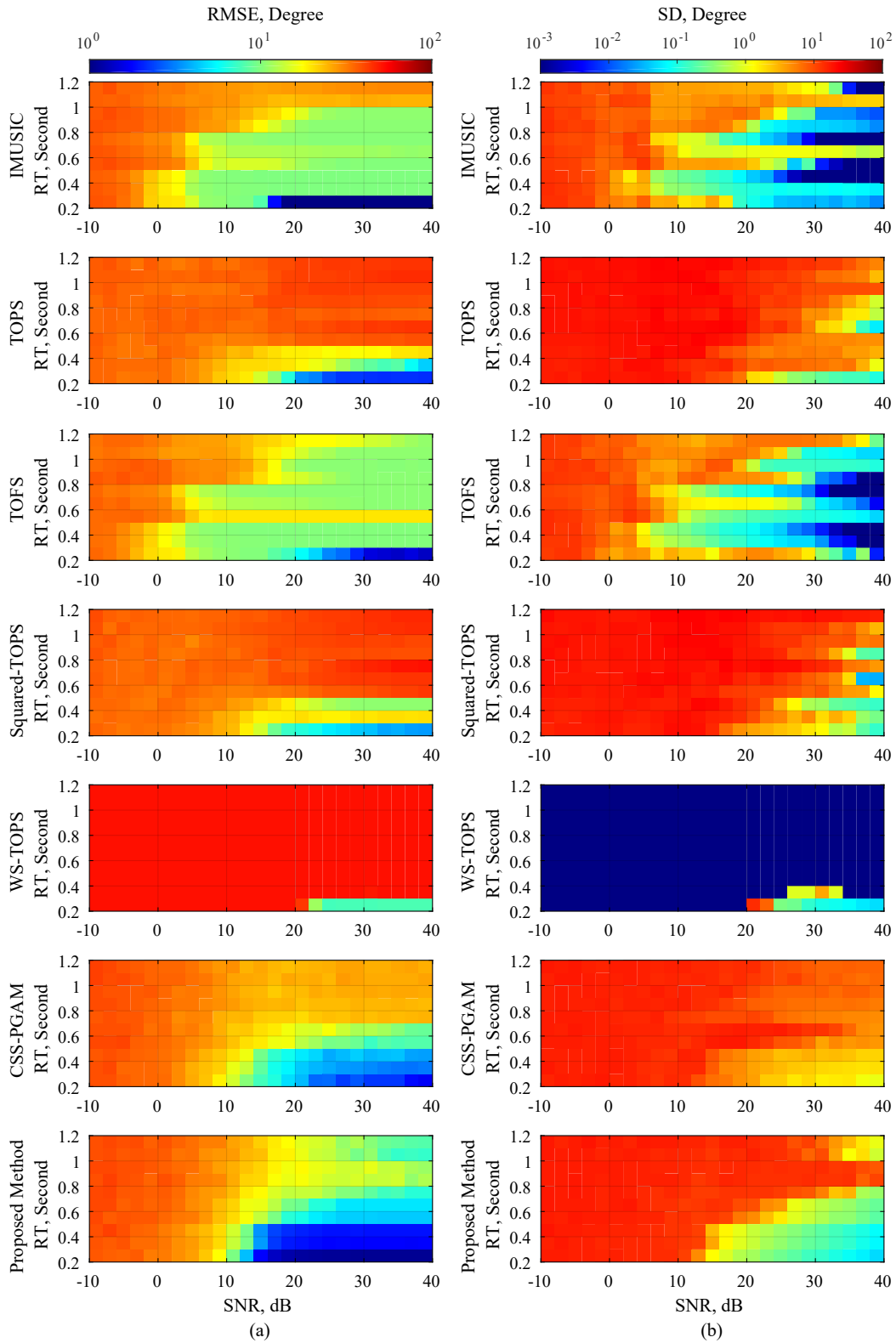


Fig. 4.3 Performance evaluations of Scenario 3 via computer simulation; (a) RMSE versus SNR, and (b) SD versus SNR under the reverberation time environment.

4.5 Numerical Simulations

sources, (2) the performance with respect to the number of microphone elements each subarray, and (3) the performance under reverberation environment.

Tables 4.1 to 4.3 shows performance comparison of the existing methods and the proposed method in term of mean absolute error (MAE) over a range of source number at 40 dB SNR. The boldfaced results highlight the optimal minimum MAE in each problem. Note that pair matching is employed. As highlighted in Tables 4.1 to 4.3, the performance at two sources of IMUSIC, TOFS and WS-TOPS exhibit the lowest MAE. When the three source is performed, the performance of selected methods are slightly dominated, but those overall performances are still satisfactory compared to the previous case. When the incident sources is increasing to four and five, it is interesting to see that the performance of IMUSIC, TOFS, TOPS, Squared-TOPS and WS-TOPS are significantly dominated by the number of sources. In addition, above simulation results are consistent with the other related-researches that the incoherent-based methods, such as IMUSIC and TOFS, require the number of microphone elements to much more higher than the number of sources [8, 18, 25, 28, 33, 37, 40, 41]. Therefore, the simulation results in Tables 4.1 to 4.3 are able to provide evidence that the proposed method performs a better estimation performance than other methods with respect to the number of sources. Although the performance of proposed method is also dominated by the number of sources, but the overall performance of proposed method is still more effective than the other methods.

Fig. 4.2 illustrate performance comparison of the existing methods and the proposed method in terms of a standard deviation (SD) and root-mean-squared error (RMSE) over a range of SNR. Three uncorrelated source angles $(\psi_k^{\text{DOA}}, \phi_k^{\text{DOA}})$ were placed at $(60^\circ, 30^\circ)$, $(45^\circ, 45^\circ)$, and $(30^\circ, 60^\circ)$, where ψ_k^{DOA} and ϕ_k^{DOA} are true DOA azimuth and zenith angles, respectively. Firstly, it is interesting to take a close look at the case of fourteen microphones in Fig. 4.2 (e), (f) that RMSE and SD of the proposed

4.5 Numerical Simulations

method worse than IMUSIC and TOFS in SNR range from 6 to 18 dB; this can be explained by effect of relaxation of the error constraints in Eq. (4.45). It should be mentioned again in the incoherent-based method that criteria of selecting number of microphone elements is based on the number of sources for providing the maximum performance [8, 18, 25, 28, 33, 37, 40, 41]; in other words, the low number of microphones elements should be considered for providing more practical applications. In the case of ten microphones each subarray, the performances of selected methods are dominated by the number of microphone elements as illustrated in Fig. 4.2 (c), (d). Furthermore, the performances of selected methods are dramatically degraded when employing six microphones as illustrated in Fig. 4.2 (a), (b). On the contrary, RMSE performance of proposed method is weakly dominated by the number of microphone elements. The proposed method gives noticeably better RMSE performance than the other methods with respect to the number of microphone elements. This substantial ability is more meaningful for many practical applications.

Fig. 4.3 represent performance comparison of the existing methods and the proposed method in terms of overall SD and RMSE over a range of SNR under reverberation environment. The reverberations were simulated by the following procedure [11]. Dimensions of enclosure room is $13 \times 13 \times 5$ meter. The number of microphone elements each subarray is twelve, and three uncorrelated source angles were placed as same as previously used. Simulation results in Fig. 4.3 indicated that reverberation has stronger effect to RMSE and SD performances in both of the existing methods and the proposed method; RMSE increases more significantly as well as SNR and the reverberation time. For practical applications, RMSE and SD are desired to be minimized simultaneously over the entire SNR range. Consider at the high reverberation time with decreasing SNR, IMUSIC and TOFS exhibit extremely lower RMSE than TOPS and TOPS-based methods. Additionally, IMUSIC and TOFS are able to provide slightly better estimates

4.6 Conclusion

than CSS-PGAM and the proposed method. Since the reverberation time is decreasing, the proposed method and CSS-PGAM begin to demonstrate low RMSE which is much less than other methods. It is apparently seen from Fig. 4.3 (a) that RMSE of the proposed method remarkably outperforms than the other methods when SNR levels ranging from 10 to 40 dB along with reverberation time decreases to less than 0.6 second. All things considered, trade-off between SNR and the reverberation should be pondered deeply in actual applications. For example, applying a noise cancellation technique to the received signal can provide the best SNR, which can be resulted in the higher performance of proposed method than the other methods.

Ultimately, all simulation results show that the performance of proposed method is especially effective for wideband 2D DOA estimation in real reverberating environments. This makes it suitable for many practical applications.

4.6 Conclusion

An efficient 2D DOA estimation with L-shaped microphone array for wideband sources was presented. We proposed a way to construct a wideband sample cross-correlation matrix and addressed a problem of estimating transformation matrices on CSS without any process of DOA preliminary estimation. The proposed transformation matrices were constructed by performing SVD of a new unique cross-correlation matrix, where elements in the row and column positions are sample cross-correlation matrices of two different frequencies. Wideband DOAs can be estimated by using this wideband sample cross-correlation matrix along with a scheme of estimating DOA in a narrowband subspace method. Therefore, a contribution of our study is providing an alternative framework for recent narrowband subspace methods to estimating the DOA of wideband sources directly, which implied that the new framework enable cutting-edge

4.7 References

techniques in the existing narrowband subspace methods to implement the wideband 2D DOA estimation for reducing the computational complexity and facilitating the estimation algorithm. Theoretical analysis and effectiveness of the proposed method are substantiated through numerical simulations, Furthermore, the results show that the proposed method exhibited effective performance than other wideband methods over a range of SNR with just a few microphones in reverberating environments. All these advantages make the proposed method a powerful tool for navigation systems based on acoustic signal processing.

4.7 References

- [1] Soo-Hwan Choi and Doo-Seop Eom. Minimizing false peak errors in generalized cross-correlation time delay estimation using subsample time delay estimation. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E96-A(1):304–311, October 2013.
- [2] Ning Ding and Nozomu Hamada. Doa estimation of multiple speech sources from a stereophonic mixture in underdetermined case. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E95-A(4):735–744, April 2012.
- [3] Y. Y. Dong, C. x. Dong, W. Liu, H. Chen, and G. q. Zhao. 2-d doa estimation for l-shaped array with array aperture and snapshots extension techniques. *IEEE Signal Processing Letters*, 24(4):495–499, April 2017.
- [4] G. Fant. *Acoustic Theory of Speech Production*. D A C S R Series. De Gruyter, 1970.
- [5] A. Ghobadzadeh and R. Adve. Low-complexity 2d root-music pairing for an l-shaped array. In *2017 IEEE Radar Conference (RadarConf)*, pages 0957–0962,

4.7 References

May 2017.

- [6] John C Gower and Garnt B Dijkstrahuis. *Procrustes Problems*. Oxford University Press, 1st edition, 2004.
- [7] Simon Haykin and K.J. Ray Liu. *Handbook on Array Processing and Sensor Networks*. Wiley-IEEE Press, 2010.
- [8] Hayashi Hirotaka and Ohtsuki Tomoaki. Doa estimation for wideband signals based on weighted squared tops. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):243, Oct 2016.
- [9] Y. Hua, T. K. Sarkar, and D. D. Weiner. An l-shaped array for estimating 2-d directions of wave arrival. *IEEE Transactions on Antennas and Propagation*, 39(2):143–146, Feb 1991.
- [10] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, Aug 1976.
- [11] Eric A. Lehmann and Anders M. Johansson. Prediction of energy decay in room impulse responses simulated with an image-source model. *The Journal of the Acoustical Society of America*, 124(1):269–277, Jul 2008.
- [12] J. Li and D. Jiang. Joint elevation and azimuth angles estimation for l-shaped array. *IEEE Antennas and Wireless Propagation Letters*, 16:453–456, 2017.
- [13] X. Luo, P. Wei, and L. Gan. Two-dimensional direction-of-arrival estimation using two transform matrices. In *2014 IEEE 17th International Conference on Computational Science and Engineering*, pages 1051–1054, Dec 2014.
- [14] S. Marcos, A. Marsal, and M. Benidir. Performances analysis of the propagator method for source bearing estimation. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, pages IV/237–IV/240 vol.4, Apr 1994.

4.7 References

- [15] Sylvie Marcos, Alain Marsal, and Messaoud Benidir. The propagator method for source bearing estimation. *Signal Processing*, 42(2):121–138, 1995.
- [16] Naoya Mochiki, Tetsuji Ogawa, and Tetsunori Kobayashi. Ears of the robot: Direction of arrival estimation based on pattern recognition using robot-mounted microphones. *IEICE Transactions on Information and Systems*, E91-D(5):1522–1530, may 2008.
- [17] Xi Nie and Ping Wei. Array aperture extension algorithm for 2-d doa estimation with l-shaped array. *Progress In Electromagnetics Research Letters*, 52:63–69, 2015.
- [18] K. Okane and T. Ohtsuki. Resolution improvement of wideband direction-of-arrival estimation ”squared-tops”. In *2010 IEEE International Conference on Communications*, pages 1–5, May 2010.
- [19] M. Omer, A. A. Quadeer, T. Y. Al-Naffouri, and M. S. Sharawi. An l-shaped microphone array configuration for impulsive acoustic source localization in 2-d using orthogonal clustering based time delay estimation. In *2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, pages 1–6, Feb 2013.
- [20] Keiichi Osako, Yoshimitsu Mori, Yu Takahashi, Hiroshi Saruwatari, and Kiyohiro Shikano. Fast convergence blind source separation using frequency subband interpolation by null beamforming. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E91-A(6):1357–1361, June 2008.
- [21] M. G. Porozantidou and M. T. Chryssomallis. Azimuth and elevation angles estimation using 2-d music algorithm with an l-shape antenna. In *2010 IEEE Antennas and Propagation Society International Symposium*, pages 1–4, July 2010.
- [22] R. Roy and T. Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, Jul 1989.

4.7 References

- [23] Miki Sato, Toru Iwasawa, Akihiko Sugiyama, Toshihiro Nishizawa, and Yosuke Takano. A single-chip speech dialogue module and its evaluation on a personal robot, papero-mini. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E93-A(1):261–271, January 2010.
- [24] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, Mar 1986.
- [25] Guaning Su and M. Morf. The signal subspace approach for multiple wide-band emitter location. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(6):1502–1522, Dec 1983.
- [26] Yuya Sugimoto, Shigeki Miyabe, Takeshi Yamada, Shoji Makino, and Biing-Hwang Juang. An extension of music exploiting higher-order moments via nonlinear mapping. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E99-A(6):1152–1162, June 2016.
- [27] B. Suksiri and M. Fukumoto. Enhanced array manifold matrices for l-shaped microphone array-based 2-d doa estimation. In *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 955–960, Dec 2017.
- [28] B. Suksiri and M. Fukumoto. A computationally efficient wideband direction-of-arrival estimation method for l-shaped microphone arrays. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5, May 2018.
- [29] Bandhit Suksiri and Masahiro Fukumoto. Multiple frequency and source angle estimation by gaussian mixture model with modified microphone array data model. *Journal of Signal Processing*, 21(4):163–166, 2017.
- [30] H. Tao, J. Xin, K. Mei, N. Zheng, and A. Sano. Detection of number of uncorrelated and coherent narrowband signals with l-shaped sensor array. In *2011 1st International Symposium on Access Spaces (ISAS)*, pages 65–70, June 2011.

4.7 References

- [31] N. Tayem, K. Majeed, and A. A. Hussain. Two-dimensional doa estimation using cross-correlation matrix with l-shaped array. *IEEE Antennas and Wireless Propagation Letters*, 15:1077–1080, Dec 2016.
- [32] Masahito Togami and Yasunari Obuchi. Stepwise phase difference restoration method for doa estimation of multiple sources. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E91-A(11):3269–3281, November 2008.
- [33] S. Valaee, B. Champagne, and P. Kabal. Localization of wideband signals using least-squares and total least-squares approaches. *IEEE Transactions on Signal Processing*, 47(5):1213–1222, May 1999.
- [34] S. Valaee and P. Kabal. Wideband array processing using a two-sided correlation transformation. *IEEE Transactions on Signal Processing*, 43(1):160–172, Jan 1995.
- [35] M. Wajid, A. Kumar, and R. Bahl. Direction-of-arrival estimation algorithms using single acoustic vector-sensor. In *2017 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, pages 84–88, Nov 2017.
- [36] G. Wang, J. Xin, N. Zheng, and A. Sano. Computationally efficient subspace-based method for two-dimensional direction estimation with l-shaped array. *IEEE Transactions on Signal Processing*, 59(7):3197–3212, July 2011.
- [37] H. Wang and M. Kaveh. Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(4):823–831, Aug 1985.
- [38] Jin Wang, Yongjun Zhao, and Zhigang Wang. Low complexity subspace fitting method for wideband signal location. In *2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN '08)*, pages 1–4, May 2008.

4.7 References

- [39] J. Xin, N. Zheng, and A. Sano. Simple and efficient nonparametric method for estimating the number of signals without eigendecomposition. *IEEE Transactions on Signal Processing*, 55(4):1405–1420, April 2007.
- [40] Yeo-Sun Yoon, L. M. Kaplan, and J. H. McClellan. Tops: new doa estimator for wideband signals. *IEEE Transactions on Signal Processing*, 54(6):1977–1989, June 2006.
- [41] H. Yu, J. Liu, Z. Huang, Y. Zhou, and X. Xu. A new method for wideband doa estimation. In *2007 International Conference on Wireless Communications, Networking and Mobile Computing*, pages 598–601, Sept 2007.
- [42] Fawad Zaman, Ijaz Mansoor Qureshi, Junaid Ali Khan, and Zafar Ullah Khan. An application of artificial intelligence for the joint estimation of amplitude and two-dimensional direction of arrival of far field sources using 2-l-shape array. *International Journal of Antennas and Propagation*, 2013.
- [43] Reza Zekavat and R. Michael Buehrer. *Handbook of Position Location: Theory, Practice and Advances*. Wiley-IEEE Press, 1st edition, 2011.
- [44] Zhi Zheng, Yuxuan Yang, Wen-Qin Wang, Guangjun Li, Jiao Yang, and Yan Ge. 2-d doa estimation of multiple signals based on sparse l-shaped array. *IEICE Transactions on Communications*, E101-B(2):383–391, February 2018.

Chapter 5

Extension Theory of Orthogonal Procrustes Analysis for Wideband Direction-of-Arrival Estimation and its Experimentation

This chapter presents an efficient framework for estimating direction-of-arrival (DOA) of wideband sound sources. The proposed framework provides an efficient way to construct a wideband cross-correlation matrix from multiple narrowband cross-correlation matrices for all frequency bins via extension theory of Orthogonal Procrustes analysis. In addition, the proposed framework is inspired by the coherent signal subspace technique with further improvement of linear transformation procedure, and the new procedure no longer require any process of DOA preliminary estimation by exploiting unique cross-correlation matrices between received signal and itself on distinct frequencies, along with the higher-order generalized singular value decomposition of the

5.1 Foreword: New Possible DOA Approach via High-Order Signal Subspace

array of this unique matrix. Wideband DOAs are estimated by employing any subspace-based technique for estimating narrowband DOAs, but using the proposed wideband correlation instead of the narrowband correlation matrix. It implies that the proposed framework enables cutting-edge researches in the recent narrowband subspace methods to estimate DOA of the wideband sources directly, which result in the reducing computational complexity and facilitating the estimation algorithm. Practical examples are presented to showcase its applicability and effectiveness, and the results show that the performance of fusion methods perform better than others over a range of signal-to-noise ratio with just a few sensors, which make it suitable for practical use.

5.1 Foreword: New Possible DOA Approach via High-Order Signal Subspace

As mentioned in Section 5.1 and Chapter 3, several approaches have been proposed as a potential way to estimate DOA. For instance, the time-difference-of-arrival based DOA estimation is one of the most frequently used approaches, which is widely known as the generalized cross-correlation with phase transform (GCC-PHAT) [13]. DOA is estimated using time shifting related to location of the largest peak of the generalized cross-correlation between two observed signals. In addition to this approach, low computational requirement makes it attractive for the practical applications; however, the major drawback is its low robustness in noisy and multipath environments. Another relevant approach is adopted from the independent component analysis (ICA) in blind source separation [28, 45]. ICA searches independent components by measuring deviations from Gaussian distributions, such as, maximization of negentropy or kurtosis. DOAs are estimated easily by using the separated components for all frequency bins, but it should be noticed that the estimation accuracy of such method is highly sensitive

5.1 Foreword: New Possible DOA Approach via High-Order Signal Subspace

to the non-Gaussianity measures.

In an alternative approach to estimate narrowband DOAs, the subspace method has been proposed in an effort to improve estimation performance. The most prominent methods observe the signal and noise subspace for achieving more robust results, such as, multiple signal classification (MUSIC) [29], estimation of signal parameters via rotational invariance techniques (ESPRIT) [27], and propagator method [16,17], which have been used frequently for one-dimensional (1D) DOA estimation along with the uniform linear array (ULA) of sensors. In case of a two-dimensional (2D) DOA estimation, a new geometrical structure of a sensor array is required, and it was previously found that the structure of L-shaped array is considerably effective for estimating 2D DOAs [11]. Additionally, the L-shaped array allows for simple implementation, because it consists of two ULAs connected orthogonally at one end of each ULA. For these reasons, the L-shaped array is widely applied to 2D DOA estimation method [3, 4, 10, 15, 21, 26, 34, 40, 42, 43], and its practical applications can be found in the past researches [23, 39]. Although the narrowband subspace method may be unable to directly use for estimating wideband DOAs, one possible way to solve this problem is to employ the narrowband subspace method in each temporal frequency intensively, and then the wideband DOA results can be estimated by interpolating the narrowband DOA results all frequency bins [31, 33]. But it should be noted again that intensive computational costs encountered in the above solution may be limited by practical considerations.

Several approaches were proposed to solve the problem of estimating wideband DOAs, for example, the incoherent MUSIC (IMUSIC) is one of the simplest methods for estimating wideband DOA [30]. There are two steps in IMUSIC: Firstly, noise subspace model each temporal frequency is constructed. Then, wideband DOAs are obtained by minimizing the norm of orthogonal relation between a steering vector and the noise subspace of all frequency bins. Although accuracy performance of IMUSIC was

5.1 Foreword: New Possible DOA Approach via High-Order Signal Subspace

demonstrated to be effective method for estimating DOAs of multiple wideband signals in high signal-to-noise ratio (SNR) region, a single small distortion of the noise subspace at any frequency can affect the whole DOA results. Many attempts were made recently to overcome this problem. For instance, the test of orthogonality of frequency subspaces (TOFS) was proposed to overcome this difficulty [47], but performance degradation caused by the small distortion still remain challenging. Another relevant approach is called as the test of orthogonality of projected subspaces (TOPS) [46]. TOPS estimate DOA by constructing signal subspace of one reference frequency, and then measuring orthogonality of the previous signal subspace and noise subspace for all frequency bins. The simulations showed that TOPS is able to achieve higher accuracy than IMUSIC in mid SNR range, however, the undesirable false peaks still remain. The revised and greatly improved version of TOPSs were proposed recently to reduce this false peaks [7, 22]. Obviously, computational complexities increased dramatically compared to the classical TOPS.

Another notable approach of wideband DOA estimation is the coherent signal subspace method (CSS) [12, 41]. CSS specifically focuses a correlation matrix of received signals each temporal frequency into a single matrix, which is called a universal correlation, associated with one focusing frequency via linear transformation procedure. Wideband DOAs are estimated by applying a single scheme of any narrowband subspace method on the universal correlation matrix. In addition to the transformation procedure of CSS [32, 35, 36], a process of DOA preliminary estimation is required before the wideband DOAs can be estimated. Therefore, a common shortcoming is clearly recognized as a requirement of DOA preliminary estimation, which means that any inferior initiation can lead to biased estimates. Some attempts have been made recently to overcome this difficulty [32, 35, 36]. According to the literature [1, 30, 46, 47], CSS demonstrates deficient performance than others such as TOPS, this is because the solutions

5.1 Foreword: New Possible DOA Approach via High-Order Signal Subspace

of transformation procedure in CSS are solely focused on subspace between a temporal frequency and focusing frequency; to the best knowledge of the authors, it means that a fundamental component of the transformation matrix across all frequency bins may exhibit the different core component, which is clearly apparent when a narrowband DOA result at some frequency is not close enough to the true DOA. A single component distortion can definitely affect the whole DOA results. Therefore, the solutions have to exhibit the exact component even though power present in a received signal at that frequency is very weak; in other words, the solution of transformation matrix have to be focused across all frequency bins instead of the pair of different frequencies.

Therefore the purpose of this research work is to investigate an alternative for estimating wideband 2D DOAs in a more efficient way. We consider wideband sources as sound sources, such as human speeches and musical sounds. In order to estimate the wideband DOAs, we address the issue of transforming multiple narrowband cross-correlation matrices for all frequency bins into a wideband cross-correlation matrix. Additionally, Our study is inspired by a computational model of CSS with further improvement of linear transformation procedure [12,32,35,36,41]. Since the transformation procedure of CSS are only focused on subspace between current and reference frequency as previously mentioned, we propose a new transformation procedure which focus all frequency bins simultaneously and efficiently. The higher-order generalized singular value decomposition (HOGSVD) is firstly used to achieve this important issue [24]. By employing HOGSVD of array of the new unique cross-correlation matrix, where elements in the row and column positions are a sample cross-correlation matrix between received signal and itself on two distinct frequencies, the new transformation procedure no longer require any process of DOA preliminary estimation. Finally, the wideband cross-correlation matrix is constructed via the proposed transformation procedure, and finally, the wideband DOAs can be estimated by employing any subspace-based tech-

5.2 Preliminaries

nique for estimating narrowband DOAs, but using this wideband correlation matrix instead of the narrowband correlation matrix. Therefore, the proposed framework enables cutting-edge researches in the recent narrowband subspace methods to estimate DOA of the wideband sources directly, which result in the reducing computational complexity and facilitating the estimation algorithm. Practical examples, such as 2D-MUSIC and ESPRIT with L-shaped array, are presented to showcase its applicability and effectiveness.

5.2 Preliminaries

5.2.1 Data Model

The proposed method presented in this chapter considers far-field sound sources. Received signals are a composition of the multiple sources, each one consisting of angle in a spherical coordinate system. The received signals are transformed into a time-frequency representation via the short-time Fourier transform (STFT), and are given by

$$\mathbf{r}(t, f) = \mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f) \mathbf{s}(t, f) + \mathbf{w}(t, f), \quad (5.1)$$

where $\mathbf{r}(t, f) \in \mathbb{C}^M$ is the summation of a received signal, $\mathbf{s}(t, f) \in \mathbb{C}^K$ is a source signal, $\mathbf{w}(t, f) \in \mathbb{C}^M$ is an additive noise, the constant M is the number of microphone elements, and K is the number of incident sources. The matrix $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f) \in \mathbb{C}^{M \times K}$ stands for the array manifold where $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ are phase angle components of the source on x and z axes in the spherical coordinate system. Note that the elements in $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f)$ depend on an array geometry.

Consider the L-shaped array structure consisting of two ULAs as illustrated in Fig. 2.1, we recall some necessary definitions here to make this chapter read easily, as

5.2 Preliminaries

follows:

$$\begin{bmatrix} \mathbf{x}(t, f) \\ \mathbf{z}(t, f) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_x(\boldsymbol{\phi}, f) \\ \mathbf{A}_z(\boldsymbol{\theta}, f) \end{bmatrix} \mathbf{s}(t, f) + \begin{bmatrix} \mathbf{w}_x(t, f) \\ \mathbf{w}_z(t, f) \end{bmatrix}, \quad (5.2)$$

where

$$\begin{aligned} \mathbf{A}_x(\boldsymbol{\phi}, f) &= \begin{bmatrix} \mathbf{a}_x(\phi_1, f) & \mathbf{a}_x(\phi_2, f) & \dots & \mathbf{a}_x(\phi_K, f) \end{bmatrix}, \\ \mathbf{A}_z(\boldsymbol{\theta}, f) &= \begin{bmatrix} \mathbf{a}_z(\theta_1, f) & \mathbf{a}_z(\theta_2, f) & \dots & \mathbf{a}_z(\theta_K, f) \end{bmatrix}, \\ \mathbf{a}_x(\phi_k, f) &= \left[e^{\alpha_x(\phi_k, f)j} \quad e^{2\alpha_x(\phi_k, f)j} \quad \dots \quad e^{N\alpha_x(\phi_k, f)j} \right]^T, \\ \mathbf{a}_z(\theta_k, f) &= \left[1 \quad e^{\alpha_z(\theta_k, f)j} \quad \dots \quad e^{(N-1)\alpha_z(\theta_k, f)j} \right]^T, \\ \alpha_x(\phi_k, f) &= \left(\frac{f}{f_o} \right) \cdot \left(\frac{2\pi d \cos \phi_k}{\lambda} \right), \\ \alpha_z(\theta_k, f) &= \left(\frac{f}{f_o} \right) \cdot \left(\frac{2\pi d \cos \theta_k}{\lambda} \right). \end{aligned} \quad (5.3)$$

From the above definitions, $\mathbf{x}(t, f), \mathbf{w}_x(t, f) \in \mathbb{C}^N$, $\mathbf{A}_x(\boldsymbol{\phi}, f) \in \mathbb{C}^{N \times K}$ and a subscript x are belonged to x subarray, and likewise, $\mathbf{z}(t, f), \mathbf{w}_z(t, f) \in \mathbb{C}^N$, $\mathbf{A}_z(\boldsymbol{\theta}, f) \in \mathbb{C}^{N \times K}$ and a subscript z are belonged to z subarray where N is the number of microphone elements each subarray with $M = 2N$. The variable t is time, f is a source frequency, d is the spacing of microphone elements, λ is a wavelength with respect to $\lambda = \frac{c}{f_o}$ where c is the speed of sound in current medium, and f_o is a reference frequency.

5.2.2 Redefinition of Model Assumptions

Unlike the previous chapter, the following assumptions are required on the proposed framework:

Assumption 1: The number of sources is known or predicted in advance [19, 44].

Assumption 2: The spacing between adjacent elements of each subarray and spacing between x_1 and z_1 should be set to $d = \frac{\lambda}{2}$ for avoiding the angle ambiguity in array structure radiation [6, 11, 48].

Assumption 3: The source $\mathbf{s}(t, f)$ is assumed to be Gaussian complex random variable as suggested by the literature [11, 29, 30]. However, we consider wideband

5.2 Preliminaries

sources as sound sources such as human speech; therefore, $\mathbf{s}(t, f)$ can also be Super-Gaussian complex random variable, and it is not stationary signals for the most general case when giving an appropriate period of time.

Assumption 4: According to acoustic theory of speech, frequency dependence of the sound source, especially a human speech, is existed [2]; it means that a cross-covariance between the source and itself with distinct frequencies is not zero; $\text{cov}(s_k(\mathbf{t}, f), s_k(\mathbf{t}, f')) = c_{s_k\{f, f'\}}$, where $c_{s_k\{f, f'\}} \in \mathbb{C}$. Next, suppose that $\mathbf{s}(\mathbf{t}, f)$ are uncorrelated, which implies that $s_k(\mathbf{t}, f)$ and $s_{k'}(\mathbf{t}, f')$ are statistically independent of each other when $k \neq k'$; $\text{cov}(s_k(\mathbf{t}, f), s_{k'}(\mathbf{t}, f')) = 0$. When $k = k'$, the sources can take to be partially dependent by the following literature [2]; therefore, a sample cross-covariance matrix of the incident sources over two different frequencies is given by

$$\begin{aligned} \mathbf{S}_{\{f, f'\}} &= E \{ \mathbf{s}(\mathbf{t}, f) \mathbf{s}^H(\mathbf{t}, f') \} \\ &= \text{diag}(c_{s_1\{f, f'\}}, c_{s_2\{f, f'\}}, \dots, c_{s_k\{f, f'\}}). \end{aligned} \quad (5.4)$$

Remark that $c_{s_k\{f, f'\}}$ is equal to $\sigma_{s_k\{f, f'\}}^2$, and $\sigma_{s_k\{f, f'\}}^2 \in \mathbb{R}_{\geq 0}$ is a variance at frequency f of the source.

Assumption 5: An additive white Gaussian noise is considered in this work, which is modeled as Gaussian random variable as well as the past studies. A noise cross-covariance matrix over two different frequencies is given by

$$\begin{aligned} \mathbf{W}_{\{f, f'\}} &= E \{ \mathbf{w}(\mathbf{t}, f) \mathbf{w}^H(\mathbf{t}, f') \} \\ &= c_w\{f, f'\} \mathbf{I}_M, \end{aligned} \quad (5.5)$$

where $c_w\{f, f'\} \in \mathbb{C}$, and \mathbf{I}_i is a i -by- i identity matrix. Note again that $c_w\{f, f\} = \sigma_w^2\{f\}$ where $\sigma_w^2\{f\} \in \mathbb{R}_{\geq 0}$ is a variance of the noise at frequency f . In case of the L-shaped array structure in Eq. (5.2), we have

$$\begin{bmatrix} \mathbf{W}_{xx\{f, f'\}} & \mathbf{W}_{xz\{f, f'\}} \\ \mathbf{W}_{zx\{f, f'\}} & \mathbf{W}_{zz\{f, f'\}} \end{bmatrix} = \begin{bmatrix} c_w\{f, f'\} \mathbf{I}_N & \mathbf{O}_{N \times N} \\ \mathbf{O}_{N \times N} & c_w\{f, f'\} \mathbf{I}_N \end{bmatrix}, \quad (5.6)$$

where $\mathbf{O}_{i \times j}$ is a i -by- j null matrix.

5.2.3 Transformation Problem

Under the data model and assumptions in Section 5.2.1 and Section 5.2.2, a cross-correlation matrix of the received signals is defined as

$$\begin{aligned}
\mathbf{R}_{\{f,f'\}} &= E \{ \mathbf{r}(\mathbf{t}, f) \mathbf{r}^H(\mathbf{t}, f') \} \\
&= \mathbf{A}(\phi, \boldsymbol{\theta}, f) E \{ \mathbf{s}(\mathbf{t}, f) \mathbf{s}^H(\mathbf{t}, f') \} \mathbf{A}^H(\phi, \boldsymbol{\theta}, f') + E \{ \mathbf{w}(\mathbf{t}, f) \mathbf{w}^H(\mathbf{t}, f') \} \\
&= \mathbf{A}(\phi, \boldsymbol{\theta}, f) \mathbf{S}_{\{f,f'\}} \mathbf{A}^H(\phi, \boldsymbol{\theta}, f') + \mathbf{W}_{\{f,f'\}},
\end{aligned} \tag{5.7}$$

where $\mathbf{R}_{\{f,f'\}} \in \mathbb{C}^{M \times M}$. In order to transform $\mathbf{R}_{\{f,f'\}}$ over the available frequency range into a single smoothed matrix, which is named as a wideband cross-correlation, a transformation procedure is required as mentioned previously [41], which is expressed as

$$\begin{aligned}
\mathbf{R} &= \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}} \mathbf{R}_{\{f_i,f_i\}} \mathbf{T}_{\{f_i\}}^H \\
&= \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}} \mathbf{A}(\phi, \boldsymbol{\theta}, f_i) \mathbf{S}_{\{f_i,f_i\}} \mathbf{A}^H(\phi, \boldsymbol{\theta}, f_i) \mathbf{T}_{\{f_i\}}^H + \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}} \mathbf{W}_{\{f_i,f_i\}} \mathbf{T}_{\{f_i\}}^H \\
&= \mathbf{A}(\phi, \boldsymbol{\theta}, f_o) \left(\frac{1}{P} \sum_{i=1}^P \mathbf{S}_{\{f_i,f_i\}} \right) \mathbf{A}^H(\phi, \boldsymbol{\theta}, f_o) + \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}} \mathbf{W}_{\{f_i,f_i\}} \mathbf{T}_{\{f_i\}}^H,
\end{aligned} \tag{5.8}$$

where

$$\mathbf{A}(\phi, \boldsymbol{\theta}, f_o) = \mathbf{T}_{\{f_i\}} \mathbf{A}(\phi, \boldsymbol{\theta}, f_i), \tag{5.9}$$

$\mathbf{R} \in \mathbb{C}^{M \times M}$ is the wideband cross-correlation matrix, and P is the number of STFT frequency bins. $\mathbf{T}_{\{f_i\}} \in \mathbb{C}^{M \times M}$ is a transformation matrix, which was originally designed by using the ordinary beamforming technique [41], or by minimizing the Frobenius norm of array manifold matrices [12]. The objective of $\mathbf{T}_{\{f\}}$ is to transform any given f of the array manifold $\mathbf{A}(\phi, \boldsymbol{\theta}, f)$ into $\mathbf{A}(\phi, \boldsymbol{\theta}, f_o)$. All previous solutions of $\mathbf{T}_{\{f\}}$ are solely based on subspace between pair of distinct frequencies $\{f, f_o\}$, as emphasized in the introduction [12, 32, 35, 36, 41]. When power of the source at some frequency is weak or

5.3 Proposed Method

less than noise power, the matrix $\mathbf{T}_{\{f\}}$ may not share any common angle of ϕ, θ because its non-zero eigenvalues are not full rank, which is resulted in a performance degradation for estimating both $\mathbf{T}_{\{f\}}$ and wideband DOAs. If the transformation matrix can be focused by all frequency bins instead of the pair of frequencies, a good estimate of DOAs on Eq. (5.8) might be expected. Based on this hypothesis, a new concept and scheme are presented in next section.

5.3 Proposed Method

This section introduces a new procedure for estimating a transformation matrix, its alternative solution by using the higher-order generalized singular value decomposition (HOGSVD), and practical examples of wideband DOA estimation scheme.

5.3.1 Problem for Estimating the Transformation Matrix and its Solution

We start by introducing the following lemma that will be useful for obtaining a solution of transformation matrix.

Lemma 3. *Given a set of two distinct frequencies by $\{f, f_o\}$ into Eq. (5.7), and given a transformation matrix $\mathbf{T}_{\{f\}}$ which satisfy the property in Eq. (6.4). Assume that $K < M$, the cross-correlation $\mathbf{R}_{\{f, f_o\}}$ can be factorized into the singular value decomposition (SVD) form;*

$$\mathbf{R}_{\{f, f_o\}} = \mathbf{U}_{\{f, f_o\}_s} \mathbf{\Sigma}_{\{f, f_o\}_s} \mathbf{V}_{\{f, f_o\}_s}^H + \mathbf{U}_{\{f, f_o\}_n} \mathbf{\Sigma}_{\{f, f_o\}_n} \mathbf{V}_{\{f, f_o\}_n}^H, \quad (5.10)$$

where $\mathbf{U}_{\{f, f_o\}_s}, \mathbf{V}_{\{f, f_o\}_s} \in \mathbb{C}^{M \times K}$, $\mathbf{\Sigma}_{\{f, f_o\}_s} \in \mathbb{R}^{K \times K}$ are the matrix of left and right singular vectors and diagonal matrix of singular values in signal subspace, and likewise, $\mathbf{U}_{\{f, f_o\}_n}, \mathbf{V}_{\{f, f_o\}_n} \in \mathbb{C}^{M \times M-K}$, $\mathbf{\Sigma}_{\{f, f_o\}_n} \in \mathbb{C}^{M-K \times M-K}$ are with noise subspace. If the K largest singular values of $\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}}$ and $\mathbf{R}_{\{f, f_o\}}$ are equal, then $\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s}$

5.3 Proposed Method

is a matrix with orthonormal columns.

Proof. Since the transformation procedure of $\mathbf{R}_{\{f, f_o\}}$ is expressed by $\mathbf{T}_{\{f\}}\mathbf{R}_{\{f, f_o\}}$ and the array manifold $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f)$ and $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f_o)$ are full rank matrices [41], Lemma 3 is valid if and only if the K largest singular values of $\mathbf{T}_{\{f\}}\mathbf{R}_{\{f, f_o\}}$ and $\mathbf{R}_{\{f, f_o\}}$ are equal; therefore, $\mathbf{U}_{\{f, f_o\}_s}^H \mathbf{T}_{\{f\}}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} = \mathbf{I}_K$. Considering the $M - K$ smallest singular values of $\mathbf{R}_{\{f, f_o\}}$ are close to zeros by assuming a noise-free signal and using solely the signal subspace $\mathbf{U}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s} \mathbf{V}_{\{f, f_o\}_s}^H$, we have

$$\begin{aligned} & (\mathbf{T}_{\{f\}} (\mathbf{R}_{\{f, f_o\}} - \mathbf{W}_{\{f, f_o\}}))^H (\mathbf{T}_{\{f\}} (\mathbf{R}_{\{f, f_o\}} - \mathbf{W}_{\{f, f_o\}})) \\ &= \mathbf{V}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s}^H \left(\mathbf{U}_{\{f, f_o\}_s}^H \mathbf{T}_{\{f\}}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right) \boldsymbol{\Sigma}_{\{f, f_o\}_s} \mathbf{V}_{\{f, f_o\}_s}^H \\ &= \mathbf{V}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s} \mathbf{V}_{\{f, f_o\}_s}^H. \end{aligned} \quad (5.11)$$

Since $\mathbf{U}_{\{f, f_o\}_s}$ have orthonormal columns [37], we have

$$\begin{aligned} & (\mathbf{R}_{\{f, f_o\}} - \mathbf{W}_{\{f, f_o\}})^H (\mathbf{R}_{\{f, f_o\}} - \mathbf{W}_{\{f, f_o\}}) \\ &= \mathbf{V}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s} \mathbf{V}_{\{f, f_o\}_s}^H. \end{aligned} \quad (5.12)$$

Performing the Eigenvalue decomposition (EVD) to Eqs. (5.11) and (5.12), the square roots of the non-zero eigenvalues of both matrices are identical [8, 37]. This completes the proof of the lemma. ■

Lemma 3 shows that $\mathbf{R}_{\{f, f_o\}}$ and $\mathbf{T}_{\{f\}}\mathbf{R}_{\{f, f_o\}}$ share the common components on the singular values and right singular vectors, whereas the both left singular vectors may be different. Since $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f)$ and $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f_o)$ are full rank, its remaining components are given by [20]:

$$\begin{aligned} \mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f) &= \mathbf{U}_{\{f, f_o\}_s} \mathbf{F}_{\{f, f_o\}}, \\ \mathbf{T}_{\{f\}} \mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f) &= (\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s}) \mathbf{F}_{\{f, f_o\}}, \\ \mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f_o) &= \mathbf{V}_{\{f, f_o\}_s} \mathbf{G}_{\{f, f_o\}}, \end{aligned} \quad (5.13)$$

where

$$\boldsymbol{\Sigma}_{\{f, f_o\}_s} = \mathbf{F}_{\{f, f_o\}} \mathbf{S}_{\{f, f_o\}} \mathbf{G}_{\{f, f_o\}}^H, \quad (5.14)$$

5.3 Proposed Method

$\mathbf{F}_{\{f,f_o\}}, \mathbf{G}_{\{f,f_o\}} \in \mathbb{C}^{K \times K}$ are full rank matrices and have invertibility. From Eqs. (5.13) and (6.4), we have

$$\mathbf{T}_{\{f\}} \mathbf{U}_{\{f,f_o\}_s} = \mathbf{V}_{\{f,f_o\}_s} \mathbf{G}_{\{f,f_o\}} \mathbf{F}_{\{f,f_o\}}^{-1}, \quad (5.15)$$

which is assured that the right singular vectors of $\mathbf{R}_{\{f,f_o\}}$ and the left singular vectors of $\mathbf{T}_{\{f\}} \mathbf{R}_{\{f,f_o\}}$ share the common subspace when $\mathbf{G}_{\{f,f_o\}} \mathbf{F}_{\{f,f_o\}}^{-1}$ has unitary property.

Since the left singular vectors of $\mathbf{T}_{\{f\}} \mathbf{R}_{\{f,f_o\}}$ is existed, we continue to introduce a new transformation procedure. The matrix $\mathbf{T}_{\{f\}}$ can be found as solution to

$$\begin{aligned} & \underset{\mathbf{T}_{\{f\}}}{\text{minimize}} \quad \left\| \mathbf{R}_{\{f_o,f_o\}} - \mathbf{T}_{\{f\}} \mathbf{R}_{\{f,f_o\}} \right\|_{\text{F}}^2 \\ & \text{subject to} \quad \sum_{k=1}^K \sigma_k^2(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f,f_o\}}) = \sum_{k=1}^K \sigma_k^2(\mathbf{R}_{\{f,f_o\}}), \end{aligned} \quad (5.16)$$

where $\|\cdot\|_{\text{F}}$ is the Frobenius norm, and $\sum_{k=1}^K \sigma_k^2(\mathbf{A})$ is the sum-of-squares K largest singular values of \mathbf{A} . If the constraint on Eq. (5.16) is not imposed, then one of the possible choices is obtained by the least squares problem [9,18]; the solution is derived by observing the point where derivative of cost function with respect to $\mathbf{T}_{\{f\}}$ is zero, then we can have $\mathbf{T}_{\{f\}}_{\text{LS}} = \mathbf{R}_{\{f_o,f_o\}} \mathbf{R}_{\{f,f_o\}}^H (\mathbf{R}_{\{f,f_o\}} \mathbf{R}_{\{f,f_o\}}^H)^{-1}$, and $\Sigma_{\{f_o,f_o\}_s} \Sigma_{\{f,f_o\}_s}^{-1} = \mathbf{I}_K$, which is not possible in practice. To solve the problem much more practicable, an alternative solution is introduced, which is based on the constraint on Eq. (5.16) and Lemma 3:

Theorem 2. *Let $\mathbf{U}_{\psi\{f\}}, \mathbf{V}_{\psi\{f\}} \in \mathbb{C}^{M \times K}$ are the matrices in signal subspace containing the left and right singular vectors of $\mathbf{R}_{\{f,f_o\}} \mathbf{R}_{\{f_o,f_o\}}^H$. Imposing the constraint on Eq. (5.16) and Lemma 3, along with the modification of orthogonal Procrustes problem (MOP), an alternative solution to Eq. (5.16) is given by*

$$\mathbf{T}_{\{f\}}_{\text{MOP}} = \mathbf{V}_{\psi\{f\}} \mathbf{U}_{\psi\{f\}}^\dagger. \quad (5.17)$$

Defining the square matrix $\mathbf{\Omega}_{\{f\}} \in \mathbb{C}^{K \times K}$ as the matrix containing error corrections,

5.3 Proposed Method

the error of transformation remains consistent with the following equation;

$$\begin{aligned} & \varepsilon_{\{f\}_{MOP}} \\ &= \left| 2 \cdot \Re \left(\text{tr} \left(\mathbf{\Sigma}_{\psi\{f\}} \left(\mathbf{\Omega}_{\{f\}} - \mathbf{I}_K \right) \right) \right) + \text{tr} \left(\mathbf{\Sigma}_{\{f, f_o\}_n}^2 \left(\mathbf{U}_{\psi\{f\}}^\dagger \mathbf{U}_{\varepsilon\{f\}} \right)^H \mathbf{U}_{\psi\{f\}}^\dagger \mathbf{U}_{\varepsilon\{f\}} \right) \right| \end{aligned} \quad (5.18)$$

where $\mathbf{\Sigma}_{\psi\{f\}} \in \mathbb{R}^{K \times K}$ and $\mathbf{U}_{\varepsilon\{f\}} \in \mathbb{C}^{M \times M-K}$ are the diagonal matrix of the K largest singular values, and the noise subspace left singular vectors of $\mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H$, respectively.

Proof. This appendix provides a detailed derivation of Theorem 2. We begin by considering the cross-correlation matrices in Eq. (5.7). $\mathbf{R}_{\{f_o, f_o\}}$ can be constructed into the EVD form, which is given by

$$\mathbf{R}_{\{f_o, f_o\}} = \mathbf{Q}_{\{f_o\}_s} \mathbf{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H + \mathbf{Q}_{\{f_o\}_n} \mathbf{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H, \quad (5.19)$$

where $\mathbf{Q}_{\{f_o\}_s} \in \mathbb{C}^{M \times K}$, $\mathbf{\Lambda}_{\{f_o\}_s} \in \mathbb{R}^{K \times K}$ are the matrix of eigenvectors and diagonal matrix of eigenvalues in signal subspace, and likewise, $\mathbf{Q}_{\{f_o\}_n} \in \mathbb{C}^{M \times M-K}$, $\mathbf{\Lambda}_{\{f_o\}_n} \in \mathbb{R}^{M-K \times M-K}$ are with noise subspace. In case of $\mathbf{R}_{\{f, f_o\}}$, it can be derived by performing SVD, which directly follows from Eq. (5.10). Since $\mathbf{A}(\phi, \theta, f)$ and $\mathbf{A}(\phi, \theta, f_o)$ are full rank matrices [41], its remaining components are expressed as follows [20]:

$$\begin{aligned} \mathbf{U}_{\{f, f_o\}_s} &= \mathbf{A}(\phi, \theta, f) \mathbf{F}_{\{f, f_o\}}^{-1}, \\ \mathbf{V}_{\{f, f_o\}_s} &= \mathbf{A}(\phi, \theta, f_o) \mathbf{G}_{\{f, f_o\}}^{-1}, \\ \mathbf{Q}_{\{f_o\}_s} &= \mathbf{A}(\phi, \theta, f_o) \mathbf{H}_{\{f_o\}}^{-1}, \end{aligned} \quad (5.20)$$

where

$$\mathbf{\Sigma}_{\{f, f_o\}_s} = \mathbf{F}_{\{f, f_o\}} \mathbf{S}_{\{f, f_o\}} \mathbf{G}_{\{f, f_o\}}^H, \quad \mathbf{\Lambda}_{\{f_o\}_s} = \mathbf{H}_{\{f_o\}} \mathbf{S}_{\{f_o, f_o\}} \mathbf{H}_{\{f_o\}}^H, \quad (5.21)$$

$\mathbf{F}_{\{f, f_o\}}, \mathbf{G}_{\{f, f_o\}}, \mathbf{H}_{\{f_o\}} \in \mathbb{C}^{K \times K}$ are also full rank and invertible. Note again that $\mathbf{U}_{\{f, f_o\}_s}, \mathbf{V}_{\{f, f_o\}_s}, \mathbf{Q}_{\{f_o\}_s}$ have orthonormal columns [37], hence, it is obvious to see that

$$\mathbf{Q}_{\{f_o\}_s} = \mathbf{V}_{\{f, f_o\}_s} \mathbf{G}_{\{f, f_o\}} \mathbf{H}_{\{f_o\}}^{-1}, \quad (5.22)$$

5.3 Proposed Method

$$\begin{aligned}\mathbf{H}_{\{f_o\}}^H \mathbf{H}_{\{f_o\}} &= \mathbf{A}^H(\phi, \boldsymbol{\theta}, f_o) \mathbf{A}(\phi, \boldsymbol{\theta}, f_o) \\ &= \mathbf{G}_{\{f, f_o\}}^H \mathbf{G}_{\{f, f_o\}}.\end{aligned}\quad (5.23)$$

From Eq. (5.22), we may expect that $\mathbf{G}_{\{f, f_o\}}$, $\mathbf{H}_{\{f_o\}}$ have unitary property, but it is incorrect when considering Eq. (5.23). Therefore, a proposition of $\mathbf{G}_{\{f, f_o\}} \mathbf{H}_{\{f_o\}}^{-1}$ have to be identity;

$$\begin{aligned}\mathbf{G}_{\{f, f_o\}} \mathbf{H}_{\{f_o\}}^{-1} &= \mathbf{V}_{\{f, f_o\}_s}^H \mathbf{Q}_{\{f_o\}_s} \\ &= \mathbf{I}_K.\end{aligned}\quad (5.24)$$

When considering only the signal subspace, it can be seen from Eqs. (5.22) to (5.24) that the right singular vectors of $\mathbf{R}_{\{f, f_o\}}$ and the eigenvectors of $\mathbf{R}_{\{f_o, f_o\}}$ are identical.

Next, we continue to generalize the objective function in Eq. (5.16) by utilizing *Orthogonal Procrustes* (OP) [5], but with some modification (MOP). The objective function in Eq. (5.16) is rederived by

$$\begin{aligned}& \left\| \mathbf{R}_{\{f_o, f_o\}} - \mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \right\|_{\text{F}}^2 \\ &= \text{tr} \left(\mathbf{R}_{\{f_o, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f, f_o\}}^H \mathbf{T}_{\{f\}}^H \right) \\ &\quad - \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H + \mathbf{R}_{\{f_o, f_o\}} \mathbf{R}_{\{f, f_o\}}^H \mathbf{T}_{\{f\}}^H \right) \\ &= \text{tr} \left(\mathbf{R}_{\{f_o, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f, f_o\}}^H \mathbf{T}_{\{f\}}^H \right) \\ &\quad - \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H + \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \right)^H \right) \\ &= \text{tr} \left(\mathbf{R}_{\{f_o, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \right) + \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f, f_o\}}^H \mathbf{T}_{\{f\}}^H \right) \\ &\quad - 2 \cdot \Re \left(\text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \right) \right),\end{aligned}\quad (5.25)$$

where $\Re(a)$ returns the real part of the variable a , $\text{tr}(\mathbf{A})$ trace of the square matrix \mathbf{A} .

Considering each expression in Eq. (5.25), the trace of a product of two square matrices is independent of the orders;

$$\begin{aligned}\text{tr} \left(\mathbf{R}_{\{f_o, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \right) &= \text{tr} \left(\mathbf{Q}_{\{f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \right) + \text{tr} \left(\mathbf{Q}_{\{f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \right) \\ &= \text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{Q}_{\{f_o\}_s} \right) + \text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \mathbf{Q}_{\{f_o\}_n} \right) \\ &= \text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \right) + \text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \right).\end{aligned}\quad (5.26)$$

5.3 Proposed Method

Employing Lemma 3, next, we have

$$\begin{aligned}
& \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f, f_o\}}^H \mathbf{T}_{\{f\}}^H \right) \\
&= \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s}^H \left(\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right)^H \right) \\
&\quad + \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n}^H \mathbf{U}_{\{f, f_o\}_n}^H \mathbf{T}_{\{f\}}^H \right) \\
&= \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s}^H \left(\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right)^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right) \\
&\quad + \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n}^H \mathbf{U}_{\{f, f_o\}_n}^H \mathbf{T}_{\{f\}}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \\
&= \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s}^H \right) + \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n}^H \mathbf{U}_{\{f, f_o\}_n}^H \mathbf{T}_{\{f\}}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right), \quad (5.27)
\end{aligned}$$

From Eq. (5.24), we finally have

$$\begin{aligned}
& \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \right) \\
&= \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \right) + \text{tr} \left(\mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \right). \\
&= \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right) + \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \quad (5.28)
\end{aligned}$$

Substituting Eqs. (5.26) to (5.28) into Eq. (5.25), the objective function is simplified as

$$\begin{aligned}
& \left\| \mathbf{R}_{\{f_o, f_o\}} - \mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \right\|_F^2 \\
&= \text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \right) + \text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \right) + \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s}^H \right) \\
&\quad + \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n}^H \mathbf{U}_{\{f, f_o\}_n}^H \mathbf{T}_{\{f\}}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \quad (5.29) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right) \right) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \right).
\end{aligned}$$

Three expressions of $\text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \right)$, $\text{tr} \left(\boldsymbol{\Lambda}_{\{f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \right)$, $\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Sigma}_{\{f, f_o\}_s}^H \right)$ are completely isolated from $\mathbf{T}_{\{f\}}$. Therefore, the optimization problem is redefined as

$$\begin{aligned}
& \text{minimize}_{\mathbf{T}_{\{f\}}} \quad \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n}^H \mathbf{U}_{\{f, f_o\}_n}^H \mathbf{T}_{\{f\}}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right) \right) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \right) \quad (5.30) \\
& \text{subject to} \quad \sum_{k=1}^K \sigma_k^2 \left(\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}} \right) = \sum_{k=1}^K \sigma_k^2 \left(\mathbf{R}_{\{f, f_o\}} \right).
\end{aligned}$$

5.3 Proposed Method

Now there are two possible cases which we need to consider. The first case is when the $M - K$ smallest singular values of $\mathbf{R}_{\{f, f_o\}}$ are close to zeros; the other is when some of the $M - K$ smallest singular values of $\mathbf{R}_{\{f, f_o\}}$ are more than zeros.

Case 1: Assume that all the $M - K$ smallest singular values of $\mathbf{R}_{\{f, f_o\}}$ are close to zeros, we have

$$\begin{aligned} & \underset{\mathbf{T}_{\{f\}}}{\text{maximize}} && 2 \cdot \Re \left(\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right) \right) \\ & \text{subject to} && \sum_{k=1}^K \sigma_k^2 (\mathbf{T}_{\{f\}} \mathbf{R}_{\{f, f_o\}}) = \sum_{k=1}^K \sigma_k^2 (\mathbf{R}_{\{f, f_o\}}), \quad \sum_{m=K+1}^M \sigma_m^2 (\mathbf{R}_{\{f, f_o\}}) = 0. \end{aligned} \quad (5.31)$$

Using the proposition of Eq. (5.24) and employing Lemma 3, two possible solutions to reach the maximum point of Eq. (5.31) can be found. The first solution is given by

$$\mathbf{T}_{\{f\}_{\text{MOP}}} = \mathbf{Q}_{\{f_o\}_s} \boldsymbol{\Omega}_{\{f\}_{\text{MOP}}} \mathbf{U}_{\{f, f_o\}_s}^\dagger, \quad (5.32)$$

where orthonormal columns has not yet been defined on $\mathbf{U}_{\{f, f_o\}_s}$, and the second solution is given by

$$\mathbf{T}_{\{f\}_{\text{OP}}} = \mathbf{Q}_{\{f_o\}_s} \boldsymbol{\Omega}_{\{f\}_{\text{OP}}} \mathbf{U}_{\{f, f_o\}_s}^H, \quad (5.33)$$

where $\mathbf{U}_{\{f, f_o\}_s}$ has orthonormal columns. Note that the subscript \dagger denotes the pseudo-inverse. When the constraints on Eq. (5.31) are imposed into Eqs. (5.32) and (5.33), we can have that $\boldsymbol{\Omega}_{\{f\}_{\text{MOP}}} = \mathbf{I}_K$, $\boldsymbol{\Omega}_{\{f\}_{\text{OP}}} = \mathbf{I}_K$, and the maximum is achieved;

$$\bar{q}_{\{f\}_{\text{case 1}}} = 2 \cdot \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \right). \quad (5.34)$$

Case 2: Assume that some of the $M - K$ smallest singular values of $\mathbf{R}_{\{f, f_o\}}$ are more than zeros, the best solution of Eq. (5.30) can be given the same as Eq. (5.33), and its minimum is equal to Eq. (5.34);

$$\begin{aligned} \underline{q}_{\{f\}_{\text{case 2:OP}}} &= \text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Sigma}_{\{f, f_o\}_n} \mathbf{U}_{\{f, f_o\}_n}^H \mathbf{T}_{\{f\}}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \\ &\quad - 2 \cdot \Re \left(\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_s} \boldsymbol{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_s} \right) \right) \\ &\quad - 2 \cdot \Re \left(\text{tr} \left(\boldsymbol{\Sigma}_{\{f, f_o\}_n} \boldsymbol{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \mathbf{T}_{\{f\}} \mathbf{U}_{\{f, f_o\}_n} \right) \right) \end{aligned}$$

5.3 Proposed Method

$$\begin{aligned}
&= \text{tr} \left(\Sigma_{\{f, f_o\}_n} \Sigma_{\{f, f_o\}_n} \mathbf{U}_{\{f, f_o\}_n}^H \mathbf{U}_{\{f, f_o\}_s} \Omega_{\{f\}_{\text{OP}}}^H \Omega_{\{f\}_{\text{OP}}} \mathbf{U}_{\{f, f_o\}_s}^H \mathbf{U}_{\{f, f_o\}_n} \right) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\Sigma_{\{f, f_o\}_s} \Lambda_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{Q}_{\{f_o\}_s} \Omega_{\{f\}_{\text{OP}}} \mathbf{U}_{\{f, f_o\}_s}^H \mathbf{U}_{\{f, f_o\}_s} \right) \right) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\Sigma_{\{f, f_o\}_n} \Lambda_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \mathbf{Q}_{\{f_o\}_s} \Omega_{\{f\}_{\text{OP}}} \mathbf{U}_{\{f, f_o\}_s}^H \mathbf{U}_{\{f, f_o\}_n} \right) \right) \\
&= -2 \cdot \text{tr} \left(\Sigma_{\{f, f_o\}_s} \Lambda_{\{f_o\}_s} \right). \tag{5.35}
\end{aligned}$$

On the contrary, when using Eq. (5.32) on Eq. (5.30), the minimum of cost function is remained by

$$\begin{aligned}
&\stackrel{\underline{\varrho}_{\{f\}_{\text{case 2:MOP}}}}{=} \text{tr} \left(\Sigma_{\{f, f_o\}_n} \Sigma_{\{f, f_o\}_n} \mathbf{U}_{\{f, f_o\}_n}^H \left(\mathbf{U}_{\{f, f_o\}_s}^\dagger \right)^H \Omega_{\{f\}_{\text{MOP}}}^H \Omega_{\{f\}_{\text{MOP}}} \mathbf{U}_{\{f, f_o\}_s}^\dagger \mathbf{U}_{\{f, f_o\}_n} \right) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\Sigma_{\{f, f_o\}_s} \Lambda_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \mathbf{Q}_{\{f_o\}_s} \Omega_{\{f\}_{\text{MOP}}} \mathbf{U}_{\{f, f_o\}_s}^\dagger \mathbf{U}_{\{f, f_o\}_s} \right) \right) \\
&\quad - 2 \cdot \Re \left(\text{tr} \left(\Sigma_{\{f, f_o\}_n} \Lambda_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \mathbf{Q}_{\{f_o\}_s} \Omega_{\{f\}_{\text{MOP}}} \mathbf{U}_{\{f, f_o\}_s}^\dagger \mathbf{U}_{\{f, f_o\}_n} \right) \right) \\
&= \text{tr} \left(\Sigma_{\{f, f_o\}_n} \Sigma_{\{f, f_o\}_n} \left(\mathbf{U}_{\{f, f_o\}_s}^\dagger \mathbf{U}_{\{f, f_o\}_n} \right)^H \mathbf{U}_{\{f, f_o\}_s}^\dagger \mathbf{U}_{\{f, f_o\}_n} \right) \\
&\quad - 2 \cdot \text{tr} \left(\Sigma_{\{f, f_o\}_s} \Lambda_{\{f_o\}_s} \right). \tag{5.36}
\end{aligned}$$

Using the solution of Eq. (5.32) rather than Eq. (5.33) allows us to relax the error constraint in the hope of arriving at a reduction in the computation of HOGSVD (For details, see Section 5.3.2), but this is still sufficient for estimating $\mathbf{T}_{\{f\}}$ without loss of generality; the squares of $M - K$ smallest singular values of $\mathbf{R}_{\{f, f_o\}}$ are very close to zeros, so we can assume that $\Sigma_{\{f, f_o\}_n}^2 \approx \mathbf{O}_{M-K \times M-K}$. Remark that error of the transformation remains consistent with the following equation;

$$\begin{aligned}
\varepsilon_{\{f\}_{\text{MOP}}} &= \left| 2 \cdot \Re \left(\text{tr} \left(\Sigma_{\{f, f_o\}_s} \Lambda_{\{f_o\}_s} \left(\Omega_{\{f\}} - \mathbf{I}_K \right) \right) \right) \right. \\
&\quad \left. + \text{tr} \left(\Sigma_{\{f, f_o\}_n}^2 \left(\mathbf{U}_{\{f, f_o\}_s}^\dagger \mathbf{U}_{\{f, f_o\}_n} \right)^H \mathbf{U}_{\{f, f_o\}_s}^\dagger \mathbf{U}_{\{f, f_o\}_n} \right) \right|. \tag{5.37}
\end{aligned}$$

To further reduce a computational burden caused by performing SVD of $\mathbf{R}_{\{f, f_o\}}$ and EVD of $\mathbf{R}_{\{f_o, f_o\}}$, we reinitialize the cross-correlation matrix as

$$\begin{aligned}
&\mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \\
&= \left(\mathbf{U}_{\{f, f_o\}_s} \Sigma_{\{f, f_o\}_s} \mathbf{V}_{\{f, f_o\}_s}^H + \mathbf{U}_{\{f, f_o\}_n} \Sigma_{\{f, f_o\}_n} \mathbf{V}_{\{f, f_o\}_n}^H \right)
\end{aligned}$$

5.3 Proposed Method

$$\begin{aligned}
& \left(\mathbf{Q}_{\{f_o\}_s} \mathbf{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H + \mathbf{Q}_{\{f_o\}_n} \mathbf{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \right)^H \\
&= \left(\mathbf{U}_{\{f, f_o\}_s} \mathbf{\Sigma}_{\{f, f_o\}_s} \mathbf{V}_{\{f, f_o\}_s}^H \mathbf{Q}_{\{f_o\}_s} \mathbf{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H \right) \\
&+ \left(\mathbf{U}_{\{f, f_o\}_n} \mathbf{\Sigma}_{\{f, f_o\}_n} \mathbf{V}_{\{f, f_o\}_n}^H \mathbf{Q}_{\{f_o\}_n} \mathbf{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H \right) \\
&= \mathbf{U}_{\{f, f_o\}_s} \mathbf{\Sigma}_{\{f, f_o\}_s} \mathbf{\Lambda}_{\{f_o\}_s} \mathbf{Q}_{\{f_o\}_s}^H + \mathbf{U}_{\{f, f_o\}_n} \mathbf{\Sigma}_{\{f, f_o\}_n} \mathbf{\Lambda}_{\{f_o\}_n} \mathbf{Q}_{\{f_o\}_n}^H, \tag{5.38}
\end{aligned}$$

which is possible to reduce the computation by performing single SVD operation on $\mathbf{R}_{\{f, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H$. The proof is completed. \blacksquare

Theorem 2 provides an efficient way to construct $\mathbf{T}_{\{f\}}$ without any process of DOA preliminary estimation, but the solution are still solely based on subspace between pair of distinct frequencies. In order to observe the solution across all frequency bins, we will present an alternative for constructing $\mathbf{T}_{\{f\}}$ by using HOGSVD along with Theorem 2, which the next section will address further.

5.3.2 Estimation of the Transformation Matrices by HOGSVD

Suppose we have a set of P complex matrices $\mathbf{E}_{\{f_i\}} \in \mathbb{C}^{M \times M}$ and all of them have a full rank;

$$\begin{aligned}
\mathbf{E}_{\{f_1\}} &= \mathbf{R}_{\{f_1, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H, \\
\mathbf{E}_{\{f_2\}} &= \mathbf{R}_{\{f_2, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H, \\
&\vdots \\
\mathbf{E}_{\{f_P\}} &= \mathbf{R}_{\{f_P, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H,
\end{aligned} \tag{5.39}$$

where $\{f_1, f_2, \dots, f_P\}$ is a set of frequency intervals, and the cross-correlation matrices $\mathbf{R}_{\{f_i, f_o\}}$ and $\mathbf{R}_{\{f_o, f_o\}}$ are obtained form Eq. (5.7). The definition of HOGSVD of these P matrices are given by the generalized singular value decomposition (GSVD) of $P \geq 2$

5.3 Proposed Method

datasets and its right singular vectors are identical in all decomposition [25], as follows:

$$\begin{bmatrix} \mathbf{E}_{\{f_1\}} \\ \mathbf{E}_{\{f_2\}} \\ \vdots \\ \mathbf{E}_{\{f_P\}} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{e_{\{f_1\}}_s} \boldsymbol{\Sigma}_{e_{\{f_1\}}_s} \\ \mathbf{U}_{e_{\{f_2\}}_s} \boldsymbol{\Sigma}_{e_{\{f_2\}}_s} \\ \vdots \\ \mathbf{U}_{e_{\{f_P\}}_s} \boldsymbol{\Sigma}_{e_{\{f_P\}}_s} \end{bmatrix} \mathbf{V}_{e_s}^H + \begin{bmatrix} \mathbf{U}_{e_{\{f_1\}}_n} \boldsymbol{\Sigma}_{e_{\{f_1\}}_n} \\ \mathbf{U}_{e_{\{f_2\}}_n} \boldsymbol{\Sigma}_{e_{\{f_2\}}_n} \\ \vdots \\ \mathbf{U}_{e_{\{f_P\}}_n} \boldsymbol{\Sigma}_{e_{\{f_P\}}_n} \end{bmatrix} \mathbf{V}_{e_n}^H, \quad (5.40)$$

where $\mathbf{U}_{e_{\{f_i\}}_s} \in \mathbb{C}^{M \times K}$, $\mathbf{U}_{e_{\{f_i\}}_n} \in \mathbb{C}^{M \times M-K}$ are the matrix of left singular vectors, $\mathbf{V}_{e_s} \in \mathbb{C}^{M \times K}$, $\mathbf{V}_{e_n} \in \mathbb{C}^{M \times M-K}$ are the matrix of right singular vectors, and $\boldsymbol{\Sigma}_{e_{\{f_i\}}_s} \in \mathbb{R}^{K \times K}$, $\boldsymbol{\Sigma}_{e_{\{f_i\}}_n} \in \mathbb{R}^{M-K \times M-K}$ are the diagonal matrix of singular values. Remark that subscripts s and n denote subspace of signal and noise, respectively. Unlike the left singular vectors $\mathbf{U}_{\{f, f_o\}_s}$ and $\mathbf{U}_{\{f, f_o\}_n}$ that have orthonormal columns by performing SVD, $\mathbf{U}_{e_{\{f_i\}}_s}$ and $\mathbf{U}_{e_{\{f_i\}}_n}$ now have unit 2-norm columns instead.

To show that \mathbf{V}_{e_s} is equal to $\mathbf{V}_{\psi_{\{f\}}_s}$ for all frequency bins, let us start from brief description of HOGSVD benchmark. The matrix \mathbf{V}_{e_s} is obtained by performing EVD the following matrix;

$$\begin{aligned} \mathbf{S} &= \frac{1}{P(P-1)} \sum_{i=1}^P \sum_{j=i+1}^P \left(\left(\mathbf{E}_{\{f_i\}}^H \mathbf{E}_{\{f_i\}} \right) \left(\mathbf{E}_{\{f_j\}}^H \mathbf{E}_{\{f_j\}} \right)^{-1} \right. \\ &\quad \left. + \left(\mathbf{E}_{\{f_j\}}^H \mathbf{E}_{\{f_j\}} \right) \left(\mathbf{E}_{\{f_i\}}^H \mathbf{E}_{\{f_i\}} \right)^{-1} \right). \end{aligned} \quad (5.41)$$

Let us redefine

$$\begin{aligned} \mathbf{E}_{\{f_i\}} &= \mathbf{R}_{\{f_i, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H \\ &= \mathbf{U}_{o_{\{f_i\}}} \boldsymbol{\Sigma}_{o_{\{f_i\}}} \mathbf{V}_{o_{\{f_i\}}}^H, \end{aligned} \quad (5.42)$$

where

$$\begin{aligned} \mathbf{U}_{o_{\{f_i\}}} &= \begin{bmatrix} \mathbf{U}_{\psi_{\{f_i\}}} & \mathbf{U}_{\varepsilon_{\{f_i\}}} \end{bmatrix}, \\ \mathbf{V}_{o_{\{f_i\}}} &= \begin{bmatrix} \mathbf{V}_{\psi_{\{f_i\}}} & \mathbf{V}_{\varepsilon_{\{f_i\}}} \end{bmatrix}, \\ \mathbf{V}_{o_{\{f_i\}}} &= \begin{bmatrix} \mathbf{Q}_{\{f_o\}_s} & \mathbf{Q}_{\{f_o\}_n} \end{bmatrix}, \\ \boldsymbol{\Sigma}_{o_{\{f_i\}}} &= \begin{bmatrix} \boldsymbol{\Sigma}_{\psi_{\{f_i\}}} & \mathbf{O}_{K \times M-K} \\ \mathbf{O}_{M-K \times K} & \boldsymbol{\Sigma}_{\varepsilon_{\{f_i\}}} \end{bmatrix}, \end{aligned} \quad (5.43)$$

5.3 Proposed Method

$\Sigma_{\varepsilon\{f_i\}} \in \mathbb{R}^{M-K \times M-K}$ is the matrix of the $M - K$ smallest singular values of $\mathbf{R}_{\{f_i, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H$, and $\mathbf{V}_{\psi\{f_i\}} = \mathbf{Q}_{\{f_o\}_s}$, $\mathbf{V}_{\varepsilon\{f_i\}} = \mathbf{Q}_{\{f_o\}_n}$ by employing Theorem 2 (For details, see Section 5.3.1). Substituting Eqs. (5.42) and (5.43) into Eq. (5.41), we have:

$$\begin{aligned} \left(\mathbf{E}_{\{f_i\}}^H \mathbf{E}_{\{f_i\}} \right) &= \left(\mathbf{U}_{o\{f_i\}} \Sigma_{o\{f_i\}} \mathbf{V}_{o\{f_i\}} \right)^H \mathbf{U}_{o\{f_i\}} \Sigma_{o\{f_i\}} \mathbf{V}_{o\{f_i\}}^H \\ &= \mathbf{V}_{o\{f_i\}} \Sigma_{o\{f_i\}} \mathbf{U}_{o\{f_i\}}^H \mathbf{U}_{o\{f_i\}} \Sigma_{o\{f_i\}} \mathbf{V}_{o\{f_i\}}^H \\ &= \mathbf{V}_{o\{f_i\}} \Sigma_{o\{f_i\}} \Sigma_{o\{f_i\}} \mathbf{V}_{o\{f_i\}}^H, \end{aligned} \quad (5.44)$$

$$\begin{aligned} \left(\mathbf{E}_{\{f_j\}}^H \mathbf{E}_{\{f_j\}} \right)^{-1} &= \left(\mathbf{V}_{o\{f_j\}} \Sigma_{o\{f_j\}} \Sigma_{o\{f_j\}} \mathbf{V}_{o\{f_j\}}^H \right)^{-1} \\ &= \left(\mathbf{V}_{o\{f_j\}}^H \right)^{-1} \Sigma_{o\{f_j\}}^{-1} \Sigma_{o\{f_j\}}^{-1} \mathbf{V}_{o\{f_j\}}^{-1} \\ &= \mathbf{V}_{o\{f_j\}}^{-H} \Sigma_{o\{f_j\}}^{-1} \Sigma_{o\{f_j\}}^{-1} \mathbf{V}_{o\{f_j\}}^{-1}, \end{aligned} \quad (5.45)$$

and then,

$$\begin{aligned} \left(\mathbf{E}_{\{f_i\}}^H \mathbf{E}_{\{f_i\}} \right) \left(\mathbf{E}_{\{f_j\}}^H \mathbf{E}_{\{f_j\}} \right)^{-1} &= \mathbf{V}_{o\{f_i\}} \Sigma_{o\{f_i\}} \Sigma_{o\{f_i\}} \mathbf{V}_{o\{f_i\}}^H \mathbf{V}_{o\{f_j\}}^{-H} \Sigma_{o\{f_j\}}^{-1} \Sigma_{o\{f_j\}}^{-1} \mathbf{V}_{o\{f_j\}}^{-1} \\ &= \mathbf{V}_{o\{f_i\}} \Sigma_{o\{f_i\}} \Sigma_{o\{f_i\}} \Sigma_{o\{f_j\}}^{-1} \Sigma_{o\{f_j\}}^{-1} \mathbf{V}_{o\{f_j\}}^{-1}, \\ \left(\mathbf{E}_{\{f_j\}}^H \mathbf{E}_{\{f_j\}} \right) \left(\mathbf{E}_{\{f_i\}}^H \mathbf{E}_{\{f_i\}} \right)^{-1} &= \mathbf{V}_{o\{f_j\}} \Sigma_{o\{f_j\}} \Sigma_{o\{f_j\}} \Sigma_{o\{f_i\}}^{-1} \Sigma_{o\{f_i\}}^{-1} \mathbf{V}_{o\{f_i\}}^{-1}. \end{aligned} \quad (5.46)$$

Since $\mathbf{V}_{\psi\{f_i\}} = \mathbf{Q}_{\{f_o\}_s} = \mathbf{V}_{\psi\{f_j\}}$, $\mathbf{V}_{\varepsilon\{f_i\}} = \mathbf{Q}_{\{f_o\}_n} = \mathbf{V}_{\varepsilon\{f_j\}}$ for all frequency bins, therefore

$$\begin{aligned} &\mathbf{V}_e^{-1} \mathbf{S} \mathbf{V}_e \\ &= \frac{1}{P(P-1)} \sum_{i=1}^P \sum_{j=i+1}^P \left(\Sigma_{o\{f_i\}} \Sigma_{o\{f_i\}} \Sigma_{o\{f_j\}}^{-1} \Sigma_{o\{f_j\}}^{-1} + \Sigma_{o\{f_j\}} \Sigma_{o\{f_j\}} \Sigma_{o\{f_i\}}^{-1} \Sigma_{o\{f_i\}}^{-1} \right). \end{aligned} \quad (5.47)$$

where

$$\mathbf{V}_e = \begin{bmatrix} \mathbf{Q}_{\{f_o\}_s} & \mathbf{Q}_{\{f_o\}_n} \end{bmatrix}. \quad (5.48)$$

Performing EVD on Eq. (5.47), we can obtain \mathbf{V}_{e_s} , which reveal that \mathbf{V}_{e_s} is equal to $\mathbf{V}_{\psi\{f\}}$ for all frequency bins. In addition, it can be seen that the matrix \mathbf{V}_{e_s} or $\mathbf{V}_{\psi\{f\}}$ is estimated by focusing all frequency bins simultaneously; when power of the source at some frequency is weak or less than noise power, the matrices $\mathbf{V}_{\psi\{f\}}$ still share common angle of ϕ, θ across all frequency bands effectively and identically.

5.3 Proposed Method

After obtaining the right singulars vectors of $\mathbf{E}_{\{f_i\}}$, we then moved forward to find its left singulars vectors. We start by considering the following equations based on Eqs. (5.40) and (5.48);

$$\begin{aligned}
\begin{bmatrix} \mathbf{E}_{\{f_1\}} \\ \mathbf{E}_{\{f_2\}} \\ \vdots \\ \mathbf{E}_{\{f_P\}} \end{bmatrix} \mathbf{V}_e &= \begin{bmatrix} \mathbf{U}_{e\{f_1\}_s} \boldsymbol{\Sigma}_{e\{f_1\}_s} \\ \mathbf{U}_{e\{f_2\}_s} \boldsymbol{\Sigma}_{e\{f_2\}_s} \\ \vdots \\ \mathbf{U}_{e\{f_P\}_s} \boldsymbol{\Sigma}_{e\{f_P\}_s} \end{bmatrix} \mathbf{V}_{e_s}^H \mathbf{V}_e + \begin{bmatrix} \mathbf{U}_{e\{f_1\}_n} \boldsymbol{\Sigma}_{e\{f_1\}_n} \\ \mathbf{U}_{e\{f_2\}_n} \boldsymbol{\Sigma}_{e\{f_2\}_n} \\ \vdots \\ \mathbf{U}_{e\{f_P\}_n} \boldsymbol{\Sigma}_{e\{f_P\}_n} \end{bmatrix} \mathbf{V}_{e_n}^H \mathbf{V}_e \\
&= \begin{bmatrix} \mathbf{U}_{e\{f_1\}_s} \boldsymbol{\Sigma}_{e\{f_1\}_s} \\ \mathbf{U}_{e\{f_2\}_s} \boldsymbol{\Sigma}_{e\{f_2\}_s} \\ \vdots \\ \mathbf{U}_{e\{f_P\}_s} \boldsymbol{\Sigma}_{e\{f_P\}_s} \end{bmatrix} \begin{bmatrix} \mathbf{I}_K & \mathbf{O}_{K \times M-K} \end{bmatrix} \\
&\quad + \begin{bmatrix} \mathbf{U}_{e\{f_1\}_n} \boldsymbol{\Sigma}_{e\{f_1\}_n} \\ \mathbf{U}_{e\{f_2\}_n} \boldsymbol{\Sigma}_{e\{f_2\}_n} \\ \vdots \\ \mathbf{U}_{e\{f_P\}_n} \boldsymbol{\Sigma}_{e\{f_P\}_n} \end{bmatrix} \begin{bmatrix} \mathbf{O}_{M-K \times K} & \mathbf{I}_{M-K} \end{bmatrix}.
\end{aligned} \tag{5.49}$$

We remark again that $\mathbf{U}_{e\{f_i\}_s}$, $\mathbf{U}_{e\{f_i\}_n}$ have unit 2-norm columns instead of orthonormal columns [25];

$$\begin{bmatrix} \mathbf{U}_{e\{f_i\}_s}^H \\ \mathbf{U}_{e\{f_i\}_n}^H \end{bmatrix} \begin{bmatrix} \mathbf{U}_{e\{f_i\}_s} & \mathbf{U}_{e\{f_i\}_n} \end{bmatrix} = \begin{bmatrix} 1 & \xi_{12} & \cdots & \xi_{1M} \\ \xi_{21} & 1 & \cdots & \xi_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{M1} & \xi_{M2} & \cdots & 1 \end{bmatrix}, \tag{5.50}$$

where $\xi_{jk} \in \mathbb{C}$, $\forall j \in M, \forall k \in M : j \neq k$. Then, the singular values are obtained as follows:

$$\begin{aligned}
\boldsymbol{\Sigma}_{e\{f_i\}_s} &= \text{diag}(\|\mathbf{e}_1\|_2, \|\mathbf{e}_2\|_2, \dots, \|\mathbf{e}_K\|_2), \\
\boldsymbol{\Sigma}_{e\{f_i\}_n} &= \text{diag}(\|\mathbf{e}_{K+1}\|_2, \|\mathbf{e}_{K+2}\|_2, \dots, \|\mathbf{e}_M\|_2),
\end{aligned} \tag{5.51}$$

where $\|\cdot\|_2$ is the Euclidean norm, and $\mathbf{e}_j \in \mathbb{C}^M$ is a j^{th} column of $\mathbf{E}_{\{f_i\}} \mathbf{V}_e$. Finally, the matrices $\mathbf{U}_{e\{f_i\}_s}$, $\mathbf{U}_{e\{f_i\}_n}$ are obtained by solving Eq. (5.49) with Eq. (5.51), which also satisfy the condition in Eq. (5.50).

5.3 Proposed Method

After performing HOGSVD of Eq. (5.39) to obtain the left and right singular vectors of $\mathbf{R}_{\{f_i, f_o\}} \mathbf{R}_{\{f_o, f_o\}}^H$, the transformation matrices $\mathbf{T}_{\{f_i\}_{\text{MOP}}}$ can be assembled as follows:

$$\begin{bmatrix} \mathbf{T}_{\{f_1\}_{\text{MOP}}} \\ \mathbf{T}_{\{f_2\}_{\text{MOP}}} \\ \vdots \\ \mathbf{T}_{\{f_P\}_{\text{MOP}}} \end{bmatrix} = \mathbf{V}_{e_s} \begin{bmatrix} \mathbf{U}_{e\{f_1\}_s}^\dagger \\ \mathbf{U}_{e\{f_2\}_s}^\dagger \\ \vdots \\ \mathbf{U}_{e\{f_P\}_s}^\dagger \end{bmatrix}. \quad (5.52)$$

Note that since orthonormal columns has not yet been assumed on the matrix $\mathbf{U}_{\psi\{f\}}$ in Theorem 2, hence, the transformation procedure via HOGSVD is still compatible with Theorem 2 without requiring any modifications (For details, see Eqs. (5.31) and (5.32) in Section 5.3.1).

We now consider a computational complexity of HOGSVD. It is not surprised that HOGSVD has a heavy computational burden; that is because matrix inversions are intensively used on Eq. (5.41). To avoid the computational burden caused by the matrix inversions, Eq. (5.41) is reformulated by the following technique [38]. It begins by performing the economy-sized QR decomposition of Eq. (5.40);

$$\begin{bmatrix} \mathbf{E}_{\{f_1\}} \\ \mathbf{E}_{\{f_2\}} \\ \vdots \\ \mathbf{E}_{\{f_P\}} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{\zeta_1} \\ \mathbf{Q}_{\zeta_2} \\ \vdots \\ \mathbf{Q}_{\zeta_P} \end{bmatrix} \mathbf{R}_\zeta, \quad (5.53)$$

where $\mathbf{R}_{\zeta_i} \in \mathbb{C}^{M \times M}$ is the upper triangular matrix, and $\mathbf{Q}_{\zeta_i} \in \mathbb{C}^{M \times M}$ is a one portion of the $(M \times P)$ -by- M matrix resulting from the QR decomposition of Eq. (5.40). Next, \mathbf{S} is simplified as

$$\mathbf{S}_\zeta = \frac{1}{P(P-1)} (\mathbf{D}_\zeta - P\mathbf{I}_M), \quad (5.54)$$

where

$$\mathbf{D}_\zeta = \sum_{i=1}^P (\mathbf{Q}_{\zeta_i}^H \mathbf{Q}_{\zeta_i})^{-1}. \quad (5.55)$$

Performing EVD of Eq. (5.55), then we have $\mathbf{D}_\zeta = \mathbf{Z}_\zeta \mathbf{\Lambda}_\zeta \mathbf{Z}_\zeta^H$, where $\mathbf{Z}_\zeta \in \mathbb{C}^{M \times M}$ and $\mathbf{\Lambda}_\zeta \in \mathbb{R}^{M \times M}$ are the matrix of eigenvectors and matrix of eigenvalues, respectively.

5.3 Proposed Method

Finally, the alternative computation of \mathbf{V}_e is expressed as $\mathbf{R}_\zeta^H \mathbf{Z}_\zeta$, where the K smallest eigenvalues of \mathbf{D}_ζ are belonged to signal subspace. Comparing Eq. (5.41) and Eq. (5.55), it is clearly seen that the technique on Eqs. (5.54) to (5.55) simplifies the mathematical model, reduces the matrix operations and improves the speed of \mathbf{V}_e computation.

5.3.3 DOA Estimation Scheme

After the transformation matrices are formed by using HOGSVD, we now proceed to describe a framework for estimating the wideband DOAs. We start by simplifying the wideband cross-correlation matrix on Eq. (5.8) with EVD form and substituting with $\mathbf{T}_{\{f_i\}_{\text{MOP}}}$, as follows:

$$\begin{aligned}
& \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}_{\text{MOP}}} \mathbf{R}_{\{f_i, f_i\}} \mathbf{T}_{\{f_i\}_{\text{MOP}}}^H \\
&= \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}_{\text{MOP}}} \left(\mathbf{Q}_{\{f_i\}_s} \mathbf{\Lambda}_{\{f_i\}_s} \mathbf{Q}_{\{f_i\}_s}^H + \mathbf{Q}_{\{f_i\}_n} \mathbf{\Lambda}_{\{f_i\}_n} \mathbf{Q}_{\{f_i\}_n}^H \right) \mathbf{T}_{\{f_i\}_{\text{MOP}}}^H \\
&= \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}_{\text{MOP}}} \mathbf{Q}_{\{f_i\}_s} \mathbf{\Lambda}_{\{f_i\}_s} \mathbf{Q}_{\{f_i\}_s}^H \mathbf{T}_{\{f_i\}_{\text{MOP}}}^H \\
&\quad + \frac{1}{P} \sum_{i=1}^P \mathbf{T}_{\{f_i\}_{\text{MOP}}} \mathbf{Q}_{\{f_i\}_n} \mathbf{\Lambda}_{\{f_i\}_n} \mathbf{Q}_{\{f_i\}_n}^H \mathbf{T}_{\{f_i\}_{\text{MOP}}}^H \\
&= \frac{1}{P} \sum_{i=1}^P \mathbf{V}_{e_s} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_s} \right) \mathbf{\Lambda}_{\{f_i\}_s} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_s} \right)^H \mathbf{V}_{e_s}^H \\
&\quad + \frac{1}{P} \sum_{i=1}^P \mathbf{V}_{e_s} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_n} \right) \mathbf{\Lambda}_{\{f_i\}_n} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_n} \right)^H \mathbf{V}_{e_s}^H \\
&= \mathbf{V}_{e_s} \left(\frac{1}{P} \sum_{i=1}^P \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_s} \right) \mathbf{\Lambda}_{\{f_i\}_s} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_s} \right)^H \right) \mathbf{V}_{e_s}^H \\
&\quad + \mathbf{V}_{e_s} \left(\frac{1}{P} \sum_{i=1}^P \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_n} \right) \mathbf{\Lambda}_{\{f_i\}_n} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_n} \right)^H \right) \mathbf{V}_{e_s}^H \\
&= \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H + \mathbf{\Pi},
\end{aligned} \tag{5.56}$$

5.3 Proposed Method

where

$$\begin{aligned}
\mathbf{\Lambda} &= \mathbf{L}^H \left(\frac{1}{P} \sum_{i=1}^P \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_s} \right) \mathbf{\Lambda}_{\{f_i\}_s} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_s} \right)^H \right) \mathbf{L}, \\
\mathbf{\Pi} &= \mathbf{V}_{e_s} \left(\frac{1}{P} \sum_{i=1}^P \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_n} \right) \mathbf{\Lambda}_{\{f_i\}_n} \left(\mathbf{U}_{e\{f_i\}_s}^\dagger \mathbf{Q}_{\{f_i\}_n} \right)^H \right) \mathbf{V}_{e_s}^H, \\
\mathbf{Q} &= \mathbf{V}_{e_s} \mathbf{L}.
\end{aligned} \tag{5.57}$$

Here, $\mathbf{\Lambda} \in \mathbb{C}^{K \times K}$ and $\mathbf{Q} \in \mathbb{C}^{M \times K}$ are the diagonal matrix of eigenvalues and matrix of eigenvectors of Eq. (5.56) in signal subspace, and $\mathbf{L} \in \mathbb{C}^{K \times K}$ possess unitary property by the fact that $\mathbf{Q}, \mathbf{V}_{e_s}$ are the matrices with orthonormal columns [8, 37]. Remark that $\mathbf{R}_{\{f_i, f_i\}}$ is also derived by performing EVD; the matrices $\mathbf{Q}_{\{f_i\}_s} \in \mathbb{C}^{M \times K}$, $\mathbf{\Lambda}_{\{f_i\}_s} \in \mathbb{R}^{K \times K}$ are the eigenvectors and diagonal matrix of eigenvalues in signal subspace, and likewise, $\mathbf{Q}_{\{f_i\}_n} \in \mathbb{C}^{M \times M-K}$, $\mathbf{\Lambda}_{\{f_i\}_n} \in \mathbb{R}^{M-K \times M-K}$ are with noise subspace. Furthermore, considering only the signal subspace by focusing on the K largest singular values $\mathbf{\Lambda}$, we can expect that Eq. (5.56) is equivalent to Eq. (5.8);

$$\begin{aligned}
\mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H &\equiv (\mathbf{V}_{\{f, f_o\}_s} \mathbf{L}) \mathbf{L}^H \left(\frac{1}{P} \sum_{i=1}^P \left(\mathbf{G}_{\{f, f_o\}} \mathbf{H}_{\{f_i\}}^{-1} \right) \mathbf{\Lambda}_{\{f_i\}_s} \left(\mathbf{G}_{\{f, f_o\}} \mathbf{H}_{\{f_i\}}^{-1} \right)^H \right) \mathbf{L} \\
&\quad (\mathbf{V}_{\{f, f_o\}_s} \mathbf{L})^H \\
&\equiv \mathbf{A}(\phi, \boldsymbol{\theta}, f_o) \left(\frac{1}{P} \sum_{i=1}^P \mathbf{S}_{\{f_i, f_i\}} \right) \mathbf{A}^H(\phi, \boldsymbol{\theta}, f_o),
\end{aligned} \tag{5.58}$$

which can be proofed by employing Lemma 3, Eqs. (5.13) to (5.15), and Eqs. (5.22) to (5.24) on Section 5.3.1 (We omit the proof since the result is easily obtained by performing straightforward substitution). In this state, $\mathbf{T}_{\{f\}_{\text{MOP}}}$ provides an efficient way to transform any given f into f_o by observing the solution across frequency bands without loss of generality; it means that the transformation is no longer biased by the pair of distinct frequencies $\{f, f_o\}$. Furthermore, it is clearly seen that the wideband cross-correlation matrix on Eq. (5.56) is the combination of narrowband sample cross-correlation matrices across all frequency bins, but its array manifold are focused on

5.3 Proposed Method

the single reference frequency by using $\mathbf{T}_{\{f\}_{\text{MOP}}}$, which is now feasible to estimate the wideband DOAs by employing any recent subspace-based technique for estimating narrowband DOAs [3, 4, 10, 15, 34, 40, 42, 43], but using this wideband correlation matrix instead of the narrowband correlation matrix. Practical examples, such as MUSIC and ESPRIT, will be presented to showcase its applicability and effectiveness in the next section.

In case of the L-shaped array structure in Eq. (5.2), we can repeat the proposed transformation procedure to find the solution for x subarray in Eqs. (5.2) and (5.3); starting from Eq. (5.7) by replacing $\mathbf{r}(\mathbf{t}, f)$ with $\mathbf{x}(\mathbf{t}, f)$, the solution for the x subarray can be given by:

$$\mathbf{T}_{x\{f_i\}_{\text{MOP}}} = \mathbf{V}_{x,e_s} \mathbf{U}_{x,e\{f_i\}_s}^\dagger, \quad (5.59)$$

$$\frac{1}{P} \sum_{i=1}^P \mathbf{T}_{x\{f_i\}_{\text{MOP}}} \mathbf{R}_{x\{f_i,f_i\}} \mathbf{T}_{x\{f_i\}_{\text{MOP}}}^H = \mathbf{Q}_x \mathbf{\Lambda}_x \mathbf{Q}_x^H + \mathbf{\Pi}_x, \quad (5.60)$$

$$\mathbf{Q}_x \mathbf{\Lambda}_x \mathbf{Q}_x^H \equiv \mathbf{A}_x(\phi, f_o) \left(\frac{1}{P} \sum_{i=1}^P \mathbf{S}_{\{f_i,f_i\}} \right) \mathbf{A}_x^H(\phi, f_o). \quad (5.61)$$

By performing the same procedure, the solution for z subarray is likewise given by replacing $\mathbf{x}(\mathbf{t}, f)$, $\mathbf{A}_x(\phi, f_o)$ with $\mathbf{z}(\mathbf{t}, f)$, $\mathbf{A}_z(\theta, f_o)$ and the subscript x with z on Eqs. (5.59) to (5.61);

$$\mathbf{T}_{z\{f_i\}_{\text{MOP}}} = \mathbf{V}_{z,e_s} \mathbf{U}_{z,e\{f_i\}_s}^\dagger, \quad (5.62)$$

$$\frac{1}{P} \sum_{i=1}^P \mathbf{T}_{z\{f_i\}_{\text{MOP}}} \mathbf{R}_{z\{f_i,f_i\}} \mathbf{T}_{z\{f_i\}_{\text{MOP}}}^H = \mathbf{Q}_z \mathbf{\Lambda}_z \mathbf{Q}_z^H + \mathbf{\Pi}_z, \quad (5.63)$$

$$\mathbf{Q}_z \mathbf{\Lambda}_z \mathbf{Q}_z^H \equiv \mathbf{A}_z(\theta, f_o) \left(\frac{1}{P} \sum_{i=1}^P \mathbf{S}_{\{f_i,f_i\}} \right) \mathbf{A}_z^H(\theta, f_o). \quad (5.64)$$

DOA Estimation Scheme via MUSIC

MUSIC estimates DOA of the sources by locating the peaks of MUSIC spectrum along with exploiting the orthogonality of the signal and noise subspaces [20, 29]. Let

5.3 Proposed Method

we define the complementary orthogonal space $(\mathbf{I}_M - \mathbf{Q}\mathbf{Q}^H)$ which is orthogonal to $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f_o)$;

$$\mathbf{a}^H(\phi_k, \theta_k, f_o) (\mathbf{I}_M - \mathbf{Q}\mathbf{Q}^H) \mathbf{a}(\phi_k, \theta_k, f_o) = 0, \quad (5.65)$$

for all $k \in \{1, 2, \dots, K\}$, where $\mathbf{a}(\phi_k, \theta_k, f_o) \in \mathbb{C}^M$ is a k^{th} column of $\mathbf{A}(\boldsymbol{\phi}, \boldsymbol{\theta}, f_o)$ as shown in Eq. (5.3). Additionally, the following complementary orthogonal space is also valid;

$$\mathbf{a}^H(\phi_k, \theta_k, f_o) (\mathbf{I}_M - \mathbf{V}_{e_s} \mathbf{V}_{e_s}^H) \mathbf{a}(\phi_k, \theta_k, f_o) = 0, \quad (5.66)$$

by the fact that $\mathbf{Q}\mathbf{Q}^H = \mathbf{V}_{e_s} (\mathbf{L}\mathbf{L}^H) \mathbf{V}_{e_s}^H = \mathbf{V}_{e_s} \mathbf{V}_{e_s}^H$, which implies that it is possible to reduce a computational complexity of Eq. (5.56) by using only \mathbf{V}_{e_s} instead of calculating \mathbf{Q} . The computationally efficient two-dimensional MUSIC (2D-MUSIC) spectrum is expressed as

$$p_{\text{2D-MUSIC}}(\phi, \theta) = \frac{1}{\mathbf{a}^H(\phi, \theta, f_o) (\mathbf{I}_M - \mathbf{V}_{e_s} \mathbf{V}_{e_s}^H) \mathbf{a}(\phi, \theta, f_o)}. \quad (5.67)$$

When the denominator in Eq. (5.67) closes to zero for the true angles of the signals, 2D-MUSIC spectrum will have peak spikes indicating this angles. In case of the L-shaped array structure, the x and z subarray angles are estimated separately by locating the spectral peaks of the following equations:

$$\begin{aligned} p_{x\text{MUSIC}}(\phi) &= \frac{1}{\mathbf{a}_x^H(\phi, f_o) (\mathbf{I}_N - \mathbf{V}_{x,e_s} \mathbf{V}_{x,e_s}^H) \mathbf{a}_x(\phi, f_o)}, \\ p_{z\text{MUSIC}}(\theta) &= \frac{1}{\mathbf{a}_z^H(\theta, f_o) (\mathbf{I}_N - \mathbf{V}_{z,e_s} \mathbf{V}_{z,e_s}^H) \mathbf{a}_z(\theta, f_o)}, \end{aligned} \quad (5.68)$$

where $\mathbf{a}_x(\phi, f_o), \mathbf{a}_z(\theta, f_o) \in \mathbb{C}^N$ are i^{th} column of $\mathbf{A}_x(\boldsymbol{\phi}, f_o), \mathbf{A}_z(\boldsymbol{\theta}, f_o)$, respectively.

5.3 Proposed Method

DOA Estimation Scheme via ESPRIT

We start by recalling the array manifold $\mathbf{A}_x(\boldsymbol{\phi}, f_o)$ and $\mathbf{A}_z(\boldsymbol{\theta}, f_o)$ on Eq. (5.3). ESPRIT takes advantage of the rotational invariance property of ULA [27], as follows:

$$\begin{aligned}\mathbf{A}_{x_2}(\boldsymbol{\phi}, f_o) &= \mathbf{A}_{x_1}(\boldsymbol{\phi}, f_o) \boldsymbol{\Phi}_x, \\ \mathbf{A}_{z_2}(\boldsymbol{\theta}, f_o) &= \mathbf{A}_{z_1}(\boldsymbol{\theta}, f_o) \boldsymbol{\Theta}_z,\end{aligned}\tag{5.69}$$

where

$$\begin{aligned}\boldsymbol{\Phi}_x &= \text{diag}\left(e^{\alpha_x(\phi_1, f_o)j}, e^{\alpha_x(\phi_2, f_o)j}, \dots, e^{\alpha_x(\phi_K, f_o)j}\right), \\ \boldsymbol{\Theta}_z &= \text{diag}\left(e^{\alpha_z(\theta_1, f_o)j}, e^{\alpha_z(\theta_2, f_o)j}, \dots, e^{\alpha_z(\theta_K, f_o)j}\right),\end{aligned}\tag{5.70}$$

$\mathbf{A}_{x_1}(\boldsymbol{\phi}, f_o), \mathbf{A}_{z_1}(\boldsymbol{\theta}, f_o) \in \mathbb{C}^{N-1 \times K}$ and $\mathbf{A}_{x_2}(\boldsymbol{\phi}, f_o), \mathbf{A}_{z_2}(\boldsymbol{\theta}, f_o) \in \mathbb{C}^{N-1 \times K}$ stand for the first and last $(N-1)$ rows of $\mathbf{A}_x(\boldsymbol{\phi}, f_o), \mathbf{A}_z(\boldsymbol{\theta}, f_o)$, respectively. Similar to [4, 34, 42], the matrices $\mathbf{Q}_x, \mathbf{Q}_z$ can be simplified with Eqs. (5.3), (5.59) to (5.61) and (5.69), as follows:

$$\begin{aligned}\mathbf{Q}_{x_1} &= \mathbf{A}_{x_1}(\boldsymbol{\phi}, f_o) \mathbf{C}_x^{-1}, \\ \mathbf{Q}_{x_2} &= \mathbf{A}_{x_2}(\boldsymbol{\phi}, f_o) \mathbf{C}_x^{-1}, \\ \mathbf{Q}_{z_1} &= \mathbf{A}_{z_1}(\boldsymbol{\theta}, f_o) \mathbf{C}_z^{-1}, \\ \mathbf{Q}_{z_2} &= \mathbf{A}_{z_2}(\boldsymbol{\theta}, f_o) \mathbf{C}_z^{-1},\end{aligned}\tag{5.71}$$

where $\mathbf{C}_x, \mathbf{C}_z \in \mathbb{C}^{K \times K}$ are invertible matrices, $\mathbf{Q}_{x_1}, \mathbf{Q}_{z_1} \in \mathbb{C}^{N-1 \times K}$ and $\mathbf{Q}_{x_2}, \mathbf{Q}_{z_2} \in \mathbb{C}^{N-1 \times K}$ stand for the first and last $(N-1)$ rows of $\mathbf{Q}_x, \mathbf{Q}_z$, respectively. Considering Eq. (5.71), we can construct new matrices $\boldsymbol{\Gamma}_x, \boldsymbol{\Gamma}_z$ as follows:

$$\begin{aligned}\boldsymbol{\Gamma}_x &= \mathbf{Q}_{x_1}^\dagger \mathbf{Q}_{x_2} \\ &= \mathbf{C}_x \left(\mathbf{A}_{x_1}^\dagger(\boldsymbol{\phi}, f_o) \mathbf{A}_{x_2}(\boldsymbol{\phi}, f_o) \right) \mathbf{C}_x^{-1} \\ &= \mathbf{C}_x \boldsymbol{\Phi}_x \mathbf{C}_x^{-1}, \\ \boldsymbol{\Gamma}_z &= \mathbf{Q}_{z_1}^\dagger \mathbf{Q}_{z_2} \\ &= \mathbf{C}_z \left(\mathbf{A}_{z_1}^\dagger(\boldsymbol{\theta}, f_o) \mathbf{A}_{z_2}(\boldsymbol{\theta}, f_o) \right) \mathbf{C}_z^{-1} \\ &= \mathbf{C}_z \boldsymbol{\Theta}_z \mathbf{C}_z^{-1}.\end{aligned}\tag{5.72}$$

5.4 Numerical Simulations

The angles ϕ_k, θ_k can thus be estimated by the eigenvalues of $\mathbf{\Gamma}_x, \mathbf{\Gamma}_z$, as follows:

$$\begin{aligned}\phi_k &= \cos^{-1} \left(\text{angle}(\lambda_{x_k}) \frac{\lambda}{2\pi d} \right), \\ \theta_k &= \cos^{-1} \left(\text{angle}(\lambda_{z_k}) \frac{\lambda}{2\pi d} \right),\end{aligned}\tag{5.73}$$

where $\lambda_{x_k}, \lambda_{z_k} \in \mathbb{C}$ is the k^{th} eigenvalue of $\mathbf{\Gamma}_x, \mathbf{\Gamma}_z$, respectively. Furthermore, it is possible to reduce the computational complexity by using only \mathbf{V}_{e_s} as well as MUSIC;

$$\begin{aligned}\mathbf{V}_{x_1, e_s}^\dagger \mathbf{V}_{x_2, e_s} &= \mathbf{L}_x \mathbf{\Gamma}_x \mathbf{L}_x^{-1} \\ &= (\mathbf{L}_x \mathbf{C}_x) \mathbf{\Phi}_x (\mathbf{L}_x \mathbf{C}_x)^{-1}, \\ \mathbf{V}_{z_1, e_s}^\dagger \mathbf{V}_{z_2, e_s} &= \mathbf{L}_z \mathbf{\Gamma}_z \mathbf{L}_z^{-1} \\ &= (\mathbf{L}_z \mathbf{C}_z) \mathbf{\Theta}_z (\mathbf{L}_z \mathbf{C}_z)^{-1},\end{aligned}\tag{5.74}$$

where

$$\begin{aligned}\mathbf{V}_{x_1, e_s} &= \mathbf{A}_{x_1}(\phi, f_o) (\mathbf{L}_x \mathbf{C}_x)^{-1}, \\ \mathbf{V}_{x_2, e_s} &= \mathbf{A}_{x_2}(\phi, f_o) (\mathbf{L}_x \mathbf{C}_x)^{-1}, \\ \mathbf{V}_{z_1, e_s} &= \mathbf{A}_{z_1}(\theta, f_o) (\mathbf{L}_z \mathbf{C}_z)^{-1}, \\ \mathbf{V}_{z_2, e_s} &= \mathbf{A}_{z_2}(\theta, f_o) (\mathbf{L}_z \mathbf{C}_z)^{-1},\end{aligned}\tag{5.75}$$

$\mathbf{V}_{x_1, e_s}, \mathbf{V}_{z_1, e_s} \in \mathbb{C}^{N-1 \times K}$ and $\mathbf{V}_{x_2, e_s}, \mathbf{V}_{z_2, e_s} \in \mathbb{C}^{N-1 \times K}$ stand for the first and last $(N-1)$ rows of $\mathbf{V}_{x, e_s}, \mathbf{V}_{z, e_s}$, respectively.

5.4 Numerical Simulations

In this section, performances of fusion methods by using the proposed framework are demonstrated in four types of the following scenarios: (1) a performance of selected methods and the proposed methods with respect to source types, (2) the performance with respect to the number of microphone elements, (3) the performance with considering automatic pairing of the x and z subarray angles, and (4) the performance under reverberation environment. Scenario 1, 2 and 4 have to find DOA of x and z subarray angles separately by using the data model on Eq. (5.2). Whereas Scenario 3 has to find

5.4 Numerical Simulations

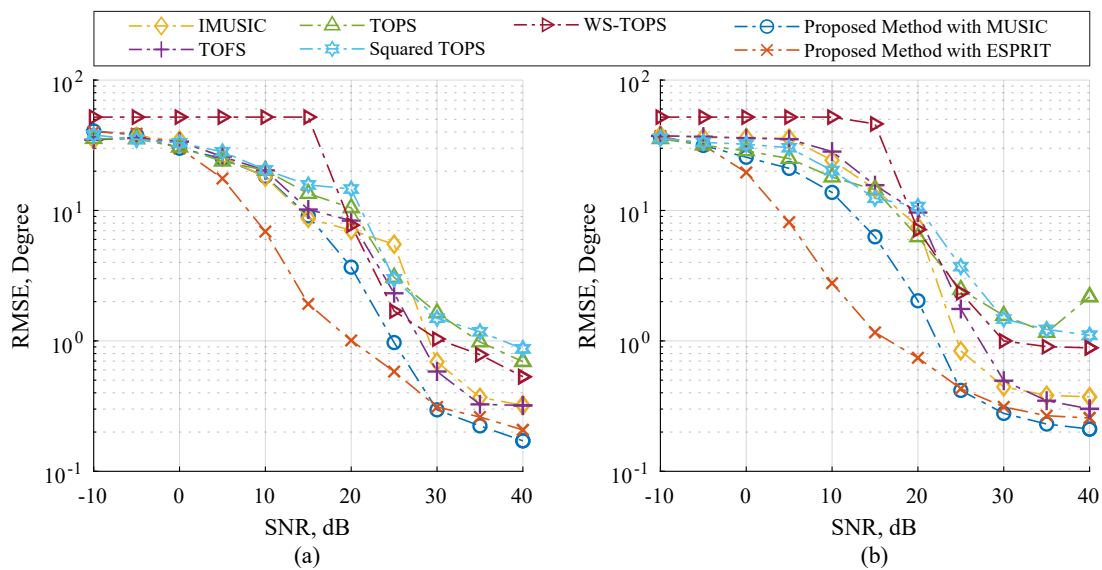


Fig. 5.1 RMSE estimation performance versus SNR on Scenario 1; (a) three different human speeches, and (b) three uncorrelated musical sounds where six microphones is employed each subarray.

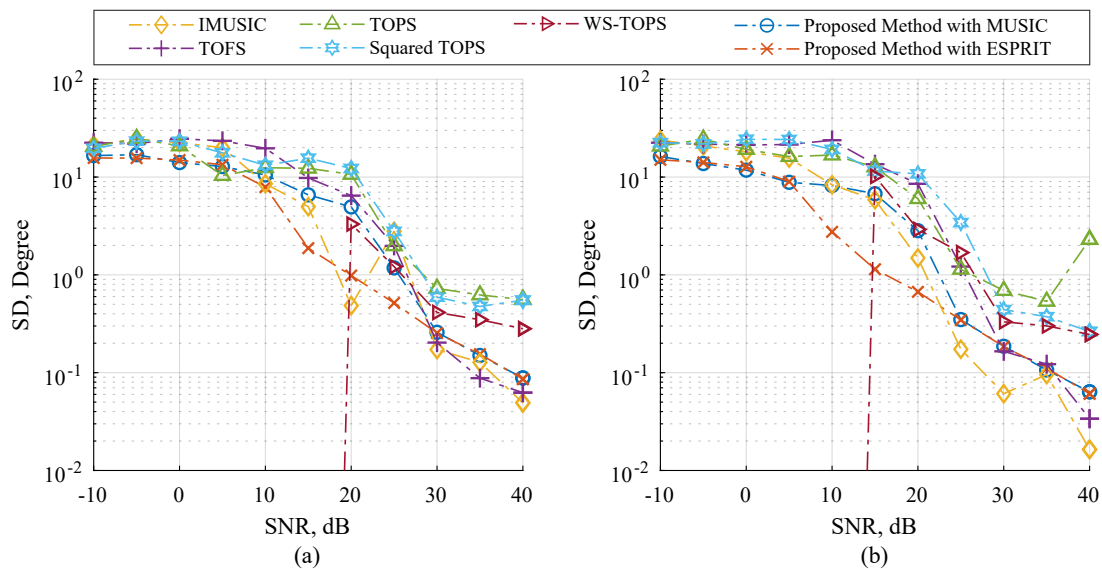


Fig. 5.2 SD estimation performance versus SNR on Scenario 1; (a) three different human speeches, and (b) three uncorrelated musical sounds where six microphones is employed each subarray.

5.4 Numerical Simulations

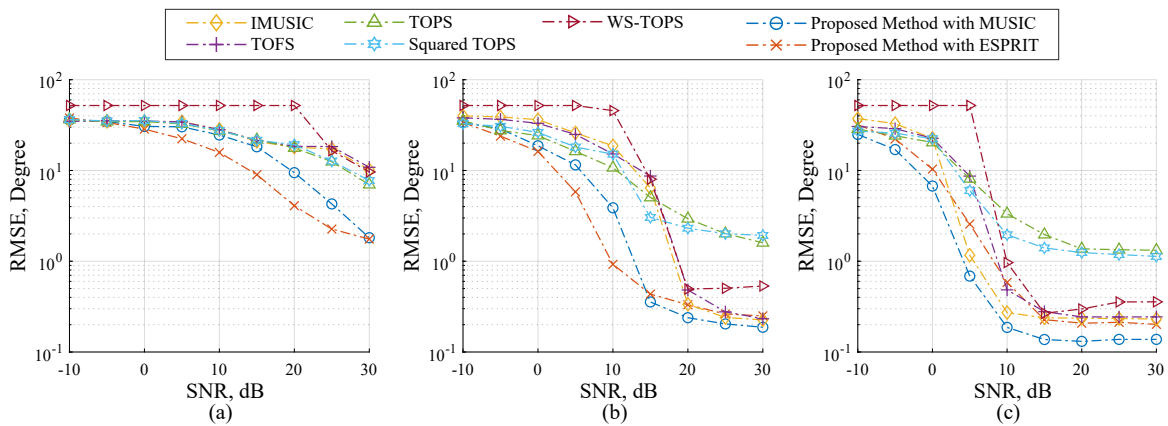


Fig. 5.3 RMSE estimation performance versus SNR on Scenario 2; three human speeches are employed and the number of microphone elements each subarray on (a) $N = 4$, (b) $N = 8$, and (c) $N = 12$.

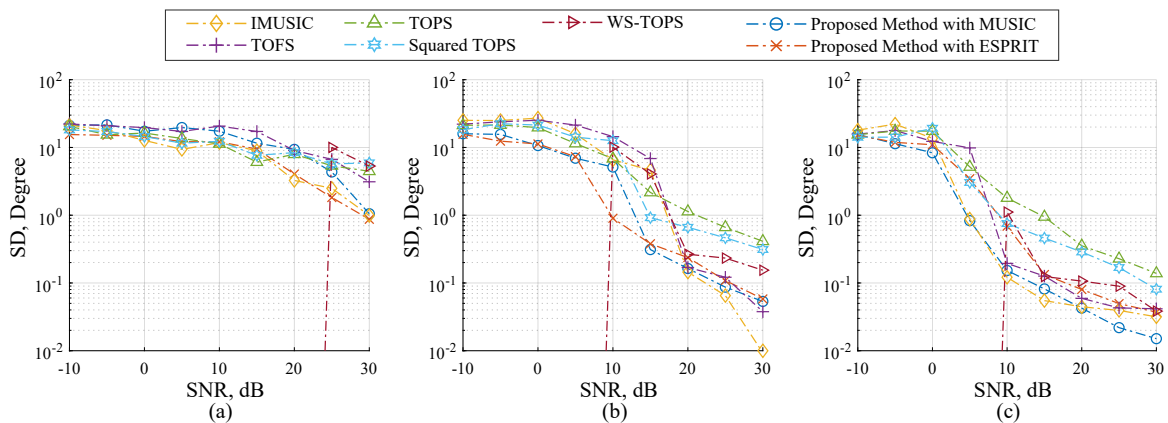


Fig. 5.4 SD estimation performance versus SNR on Scenario 2; three human speeches are employed and the number of microphone elements each subarray on (a) $N = 4$, (b) $N = 8$, and (c) $N = 12$.

DOA of x and z subarray angles simultaneously with considering automatic pairing, by using the data model on Eq. (6.1). We provided the simulation tests of the proposed methods in comparison with following methods: IMUSIC [30], TOFS [47], TONS [46], Squared-TONS [22], WS-TONS [7]. Remark that the CSS-based methods are excluded in these tests; this is because unintended biases, causing by a process of DOA preliminary estimation, should be taken into consideration to other candidate methods as

5.4 Numerical Simulations

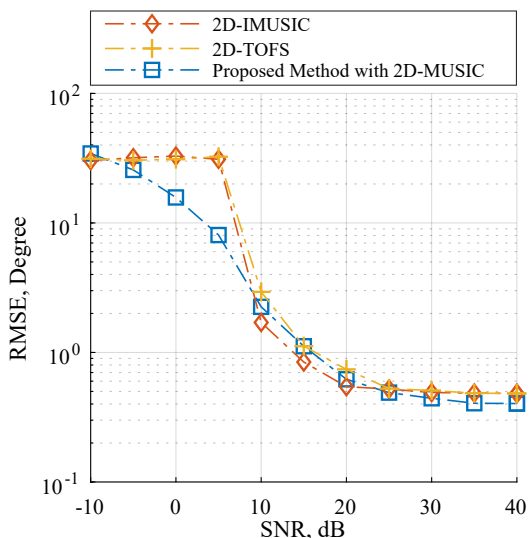


Fig. 5.5 RMSE estimation performance versus SNR on Scenario 3 where $M = 8$.

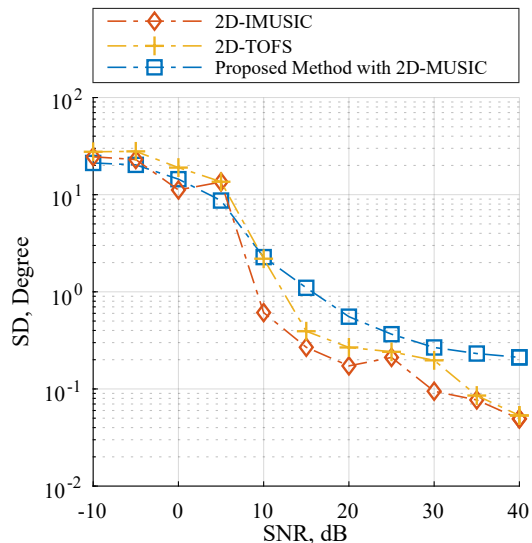


Fig. 5.6 SD estimation performance versus SNR on Scenario 3 where $M = 8$.

discussed in the literature [1, 30, 46, 47].

To measure the overall performance of estimating the x and z subarray angles for each scenario, root-mean-square-error (RMSE) and standard deviation (SD) are defined as the following equations;

$$\text{RMSE} = \sqrt{\frac{1}{2JK} \sum_{j=1}^J \sum_{k=1}^K \left(\left(\hat{\phi}_k^{(j)} - \phi_k \right)^2 + \left(\hat{\theta}_k^{(j)} - \theta_k \right)^2 \right)}, \quad (5.76)$$

$$\text{SD} = \sqrt{\frac{1}{2JK} \sum_{j=1}^J \sum_{k=1}^K \left(\left(\hat{\phi}_k - \bar{\phi}_k^{(j)} \right)^2 + \left(\hat{\theta}_k^{(j)} - \bar{\theta}_k \right)^2 \right)}, \quad (5.77)$$

where K is the source number, J is the number of trials, $\hat{\phi}_k^{(j)}$, $\hat{\theta}_k^{(j)}$ represent the estimated x and z subarray angles each trial, $\bar{\phi}_k$, $\bar{\theta}_k$ represent an average of the estimated x and z subarray angles, and ϕ_k , θ_k represent true x and z subarray angles.

Computer simulations were carried out in Matlab[®] R2017a, using PC with Debian GNU/Linux 9.4 x86_64, Intel[®] Core[™] i5-4590 CPU 3.30 GHz, 16G RAM, Intel[®] Math Kernel Library 11.3.1 on BLAS and LAPACK 3.5.0. Each scenario is repeated 100 times, and simulation parameters are chosen as follows: sampling frequency is 48

5.4 Numerical Simulations

kHz, an output of each microphone is captured 1 second, speed of sound c is 343 m/s, the spacing of microphone elements d is 5 cm, STFT focusing frequency range is from 0.1 to 16 kHz, the reference frequency f_o is 3.43 kHz. Note that we used perturbations of the true angles by adding Gaussian random noise.

5.4.1 Scenario 1: Performance with Respect to Source Types

Figs. 5.1 and 5.2 showed performance comparisons of the selected methods and the proposed methods in term of RMSE and SD over a range of SNR. The proposed methods are the modified MUSIC on Eq. (5.68) and ESPRIT on Eqs. (5.73) to (5.75). The number of microphone elements each subarray is six, and the three uncorrelated source angles (ϕ_k, θ_k) are placed at $(41.41^\circ, 60^\circ)$, $(60^\circ, 45^\circ)$ and $(75.52^\circ, 30^\circ)$. In Fig. 5.1 (a) and Fig. 5.2 (a), sources are human speeches. Whereas sources in Fig. 5.1 (b) and Fig. 5.2 (b) are musical sounds or instrument tones containing frequency range up to 48 kHz. Remark that all sources are not stationary signals. The results in Figs. 5.1 and 5.2 showed that the proposed method with ESPRIT can efficiently handle the both source types than other candidate methods with acceptable SNR range. Subsequently, it is interesting to take a close look at 40 dB SNR in Figs. 5.1 and 5.2 that IMUSIC, TOFS, the proposed method with MUSIC and ESPRIT showed very low RMSE, which could attest to good DOA estimation. When decreasing the SNR to 25 dB, IMUSIC and TOFS begin to demonstrate worse RMSE quality which is much higher than the proposed methods, and it is clearly seen when decreasing the SNR to 10 dB that all tested methods are significantly dominated, but the proposed method with ESPRIT is still associated with more satisfactory results than using other methods. It should be mentioned furthermore that IMUSIC and TOFS require the number of sensor elements to much more higher than the number of sources to achieve fairly good results [1, 30, 46, 47]. Hence, the simulation results in Figs. 5.1 and 5.2 are able to provide evidence that

5.4 Numerical Simulations

the proposed methods perform the better estimation performance than other candidate methods when the incident sources are wideband and non-stationary signals. Although the performances of proposed method with MUSIC is also dominated by the noises, but the overall performances is still more effective than other methods.

5.4.2 Scenario 2: Performance with Respect to the Number of Microphone Elements

Figs. 5.3 and 5.4 illustrated performance comparisons of the selected methods and the proposed methods in term of RMSE and SD over a range of SNR. The three uncorrelated source angles are human speeches, and are placed as same as previously used. Firstly, let us start by looking at the case of twelve microphones in Fig. 5.3 (c) and Fig. 5.4 (c). IMUSIC, TOFS and WS-TOPS exhibited remarkably low level of RMSE in SNR range from 15 to 30 dB; this is because there performances dramatically depend on the number of sensor elements to higher than the number of sources [1, 30, 46, 47]. Likewise, the proposed method with MUSIC and ESPRIT also demonstrated very low RMSE, which may imply that the performance of proposed methods, IMUSIC, TOFS and WS-TOPS are especially effective for a wideband DOA estimation. However, the low number of microphones elements should be considered for providing more practical applications. In the case of eight microphones each subarray, the performances of the selected methods are dominated by the number of microphone elements as illustrated in Fig. 5.3 (b) and Fig. 5.4 (b). Furthermore, the performances of selected methods are dramatically degraded when employing four microphones as illustrated in Fig. 5.3 (a) and Fig. 5.4 (a). The relevant reason is that an undesirable false peaks in the spatial spectrum of the selected methods are occurred, caused by the perturbation of noise subspace; when power of the noise at some frequency is high or grater than source power, the orthogonality between the noise subspace and search space at that frequency may

5.4 Numerical Simulations

be not sufficient to prevent the false-alarm peaks [1]. On the contrary, RMSE performance of proposed methods are also dominated, but less than the other methods, by exhibiting the subspace for all frequency bins simultaneously as shown in Section 5.3. Therefore, the proposed methods provide substantially better RMSE performance than the other methods, which implies that dependency between the number of microphone elements and sources can be relaxed. This substantial ability is more meaningful for many practical applications.

5.4.3 Scenario 3: Performance with Considering Automatic Pairing

This scenario estimated the DOA of x and z subarray angles simultaneously with considering automatic pairing and following the data model on Eq. (6.1). As the L-shaped array structure consisting of two ULAs as illustrated in Fig. 2.1, some research works estimate the DOA of x and z subarray angles separately by implementing 1D DOA estimation for each ULA [3, 4, 10, 15, 21, 26, 34, 40, 42, 43]. When utilizing more than one source, these algorithms require an additional angle pair matching procedure to mapping relationship between the two independent subarray angles, which may results in a performance degradation caused by pair matching error. In order to achieve the automatic pairing without the pair matching procedure, we selected the modified 2D-MUSIC on Eq. (5.67) as the proposed method in this scenario. Furthermore, TOPS, Squared-TOPS, WS-TOPS are excluded in these tests by the fact that there methods have only supported for ULA model. Note that 2D peak finding algorithm was employed on 2D-IMUSIC, 2D-TOFS and the proposed method. Figs. 5.3 and 5.4 showed performance comparisons of 2D-IMUSIC, 2D-TOFS and the proposed method in term of RMSE and SD over a range of SNR, where the number of microphone elements including all subarray is eight, the three uncorrelated source angles are human speeches,

5.4 Numerical Simulations

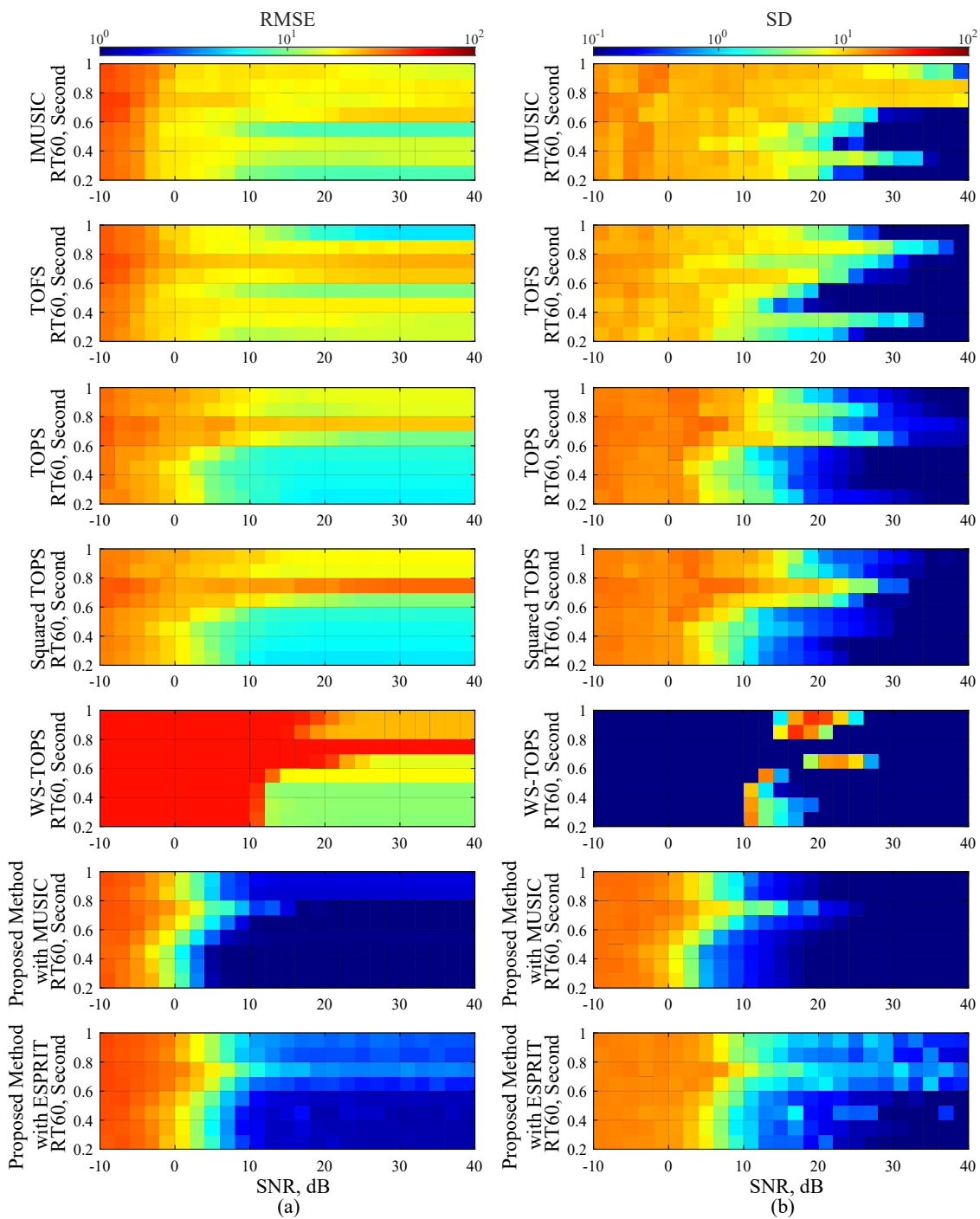


Fig. 5.7 Performance evaluations of Scenario 4; (a) RMSE estimation performance versus SNR, and (b) SD estimation performance versus SNR, where three uncorrelated human speeches are employed along with a reverberant environment, where dimensions of enclosure room is $15 \times 15 \times 5$ meter, a measurement protocol of reverberation time is RT60, and wall absorption coefficients are followed on Table 5.1.

5.4 Numerical Simulations

Table 5.1 Wall absorption coefficients at various reverberation time in Scenario 4

Reverberation Time based on RT60 (Millisecond)	Axial Wall Plane					
	Positive Direction			Negative Direction		
	$x - z$	$x - z$	$x - y$	$x - z$	$x - z$	$x - y$
200	0.7236	0.2021	0.6844	0.0792	0.2436	0.5586
300	0.7142	0.1687	0.7666	0.2650	0.2387	0.7043
400	0.7306	0.0555	0.7731	0.4091	0.8493	0.8587
500	0.5064	0.4974	0.8248	0.4189	0.8069	0.7572
600	0.6074	0.6299	0.8028	0.7599	0.6373	0.8209
700	0.7442	0.7624	0.8734	0.6922	0.6480	0.7893
800	0.6779	0.6827	0.7865	0.8045	0.8386	0.8430
900	0.6992	0.7111	0.7741	0.8752	0.8233	0.9081
1000	0.7622	0.7707	0.9394	0.8248	0.8192	0.8398

and are placed as same as previously used. Fig. 5.3 indicated that the proposed method with 2D-MUSIC exhibits extremely similar overall performances to 2D-IMUSIC and 2D-TOFS when the SNR increases to more than 10 dB; however, computational burden of the proposed method can be significantly lower than those of the other methods, which Section 5.4.5 will reveal further insight.

5.4.4 Scenario 4: Performance under Reverberation Environment

In this scenario, we compared RMSE and SD performances of proposed methods to other methods with respect to reverberation time. This scenario estimated DOA of x and z subarray angles separately by using the data model on Eq. (5.2) without

5.4 Numerical Simulations

considering automatic pairing. The proposed methods in this scenario are the modified MUSIC on Eq. (5.68) and ESPRIT on Eqs. (5.73) to (5.75). The reverberations were simulated by the following procedure [14], and its simulated wall absorption coefficients are shown in Table 5.1, where the dimensions of enclosure room is $15 \times 15 \times 5$ meter, a measurement protocol of reverberation time is RT60, and the reverberation time is from 200 to 1000 ms. The three uncorrelated source angles are employed in the same way as previously used, and the number of microphone elements each subarray is twelve. Fig. 5.7 illustrated performance comparisons of the selected methods and the proposed methods, where a color of the graph on Fig. 5.7 (a) denotes RMSE, whereas a color of the graph on Fig. 5.7 (b) denotes SD estimation performance. The vertical axis is represented as the reverberation time and horizontal axis is represented as a range of SNR. Simulation results in Fig. 5.7 indicated that reverberation has strong effect to RMSE and SD performances in both of the selected methods and the proposed methods, and the performances decreased more significantly at the high noise levels and the long reverberation times. Since the reverberation time is decreasing, all selected methods begin to demonstrate low RMSE. It means that trade-off between the robustness of reverberation and SNR should be considered deeply in actual applications, for instance, applying a reverberation cancellation technique or a noise cancellation technique to provide much more reliable estimation performances of both RMSE and SD. The proposed methods, however, largely outperform the other methods with respect to the reverberation time index and SNR levels range between 10 and 40 dB without considering the trade-off. This can support that the performance of proposed methods can be especially effective for a wideband DOA estimation under reverberant environment.

5.4 Numerical Simulations

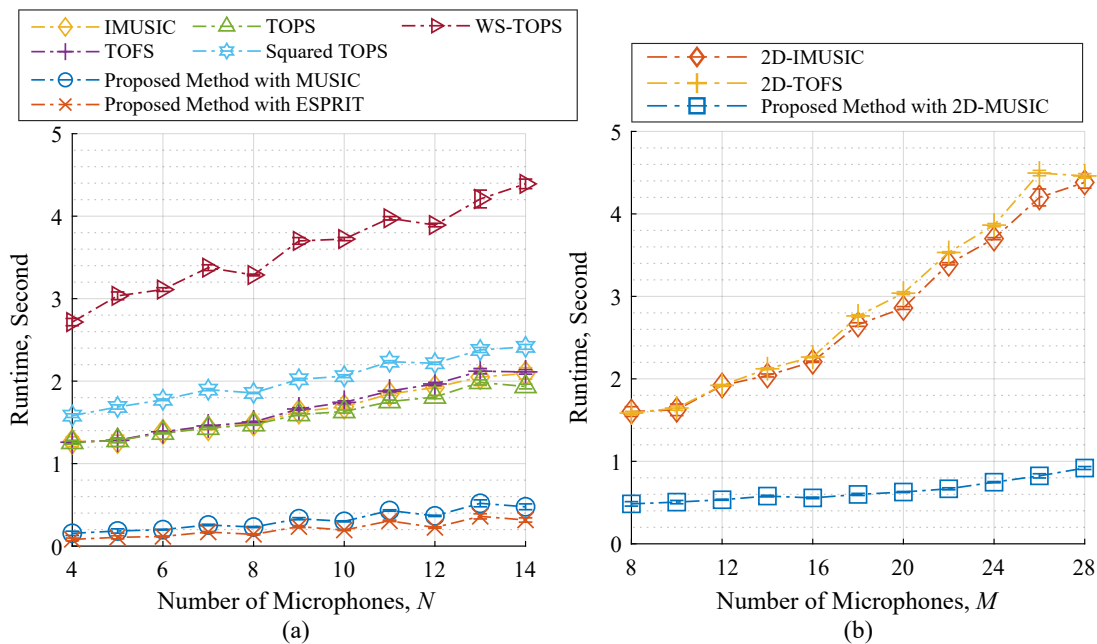


Fig. 5.8 Computational complexities; (a) changing the number of microphone elements each subarray N , and (b) the number of microphone elements including all subarray M where the number of incident sources $K = 3$.

5.4.5 Computational Complexity

A computational complexity of the proposed methods were evaluated using execution time measurement under a stable environment. We provided a computational complexity in comparison with following cases: (1) calculating DOAs of x and z subarray angles separately as shown in Fig. 5.8 (a), and (2) calculating the DOAs of both subarray angles simultaneously as shown in Fig. 5.8 (b). Note that computational burden of a peak searching algorithm are concerned in this study, where the number of searching angle each subarray is 180. It is apparently seen in Fig. 5.8 that computation time of the other methods presented more high growth rates than the proposed methods. This is because the peak searching algorithm execution time is potentially high, and almost all selected methods require intensive computations by testing the orthogonality of subspace and search space of narrowband sample cross-correlation matrices

5.5 Experimental Results

for all frequency bins, which results in high computation costs. On the contrary, the proposed methods transform all narrowband sample cross-correlation matrices across all frequency bins into the single matrix as shown in Eqs. (5.56) to (5.58), and this matrix contains useful information of source cross-correlation matrices across all frequency bins as $\frac{1}{P} \sum_{i=1}^P \mathbf{S}_{\{f_i, f_i\}}$; in other words, the orthogonality testing of subspace and search space can be done by using the wideband cross-correlation matrix on Eqs. (5.56) to (5.58) instead of narrowband sample cross-correlation matrices for all frequency bins. Therefore, the computational complexity of proposed methods remarkably less than the other methods, which is confirmed by the test results on Fig. 5.8.

5.5 Experimental Results

In this section, experiments were carried to examine the performance of proposed methods. Experiment parameters were chosen as the previously simulations, except as follows: We used human speakers as sources of the original speech with random sentences. Their speeches were recorded 20 runs continuously, and each record signal, approximating 1 minute long, was cut into 3 second epochs. Structure of microphone was followed by Fig. 2.1, and the specifications of microphone and its recording device were followed on Table 5.2. The experiment was performed in an indoor meetings room, and its dimensions are shown in Fig. 5.10, where sound pressure level in the meeting room on a normal situation is 46.6 dBA, and the estimated reverberation time based on RT60 is 219 millisecond.

Two scenarios are considered: (1) estimating DOA of x and z subarray angles separately, and (2) estimating DOA of x and z subarray angles simultaneously with considering automatic pairing. In case of Experiment 1, the proposed methods are the modified MUSIC on Eq. (5.68) and ESPRIT on Eqs. (5.73) to (5.75), comparing with

5.5 Experimental Results

Table 5.2 System specification

Hardware Type / Parameter	Specification / Value
Audio Interface	Roland [®] Octa-capture (UA-1010)
Sampling Frequency	48,000 Hz
Microphone Name	Behringer [®] C-2 studio condenser microphone
Number of Microphones	8
Pickup Patterns	Cardioid (8.9 mV/Pa; 20 - 20,000 Hz)
Diaphragm Diameter	16 mm
Equivalent Noise Level	19.0 dBA (IEC 651)
SNR Ratio	75 dB
Microphone Structure	L-shaped Array
Spacing of Microphone	9 cm

the following methods: IMUSIC [30], TOFS [47], TOPS [46], Squared-TOPS [22], WS-TOPS [7]. In case of Experiment 2, the proposed method is the modified 2D-MUSIC on Eq. (5.67), comparing with 2D-IMUSIC [30], and 2D-TOFS [47].

Tables 5.3 to 5.6 showed performance comparisons of the selected methods and the proposed method in term of RMSE over the range of source number, where Tables 5.3 and 5.4 is for Experiment 1, and Tables 5.5 and 5.6 is for Experiment 2. The bold-faced results highlight the optimal minimum RMSE in each problem. As highlighted in Tables 5.3 and 5.4, the performance of IMUSIC exhibited the lowest RMSE when a single source was used, but the performance of other methods including the proposed methods also exhibited similarly low RMSE in acceptable error range. When the two sources are performed, the performance of TOPS, Squared-TOPS and WS-TOPS are directly dominated, whereas IMUSIC, TOFS and the proposed methods are slightly dominated, but it still maintained sufficiently good performance. When the incident

5.5 Experimental Results

Table 5.3 Performance evaluation on Experiment 1; $K = 1, 2$

Incident Sources		RMSE of DOAs (Degree)									
Number	Position	Angle	IMUSIC	TOFS	TOPS	Squared TOPS	WS-TOPS	Proposed Method with MUSIC	Proposed Method with ESPRIT		
1	ϕ_1	96	0.3050	0.2050	1.0950	1.3350	0.5600	0.7750	0.7074		
	θ_1	86	0.5400	1.2600	1.2750	2.0150	0.6850	0.5700	0.6915		
	Total		0.4225	0.7325	1.1850	1.6750	0.6225	0.6725	0.6995		
2	ϕ_1	65	1.1857	1.7286	20.0143	28.5857	37.8714	1.5000	2.0284		
	θ_1	150	9.6000	6.6857	26.3571	39.7857	88.2000	8.8143	8.6800		
	ϕ_2	55	1.0714	1.6857	22.2571	19.4000	32.2429	2.9714	3.8695		
	θ_2	100	8.3714	8.3857	5.0143	6.7857	60.2286	6.6714	3.1630		
Total		5.0571	4.6214	18.4107	23.6393	54.6357	4.9893	4.4353			

5.5 Experimental Results

Table 5.4 Performance evaluation on Experiment 1; $K = 3$

Incident Sources		RMSE of DOAs (Degree)									
Number	Position	Angle	IMUSIC	TOFS	TOPS	Squared TOPS	WS-TOPS	Proposed Method with MUSIC	Proposed Method with ESPRIT		
	ϕ_1	58	2.1400	2.3900	46.5500	52.8100	40.9500	3.6600	4.0334		
	θ_1	55	55.0000	55.0000	55.0000	55.0000	55.0000	9.4300	4.1057		
	ϕ_2	100	1.8400	2.0000	41.5700	62.4000	70.9100	1.8700	2.4554		
	θ_2	95	95.0000	83.4200	52.4500	71.4800	95.0000	9.7700	5.8638		
	ϕ_3	130	10.9300	11.8900	28.8300	32.2800	95.2400	8.2500	6.9071		
	θ_3	120	26.9800	25.8400	16.1200	18.0100	91.2800	5.9400	7.3165		
	Total		31.9817	30.0900	40.0867	48.6633	74.7300	6.4867	5.1137		

5.5 Experimental Results

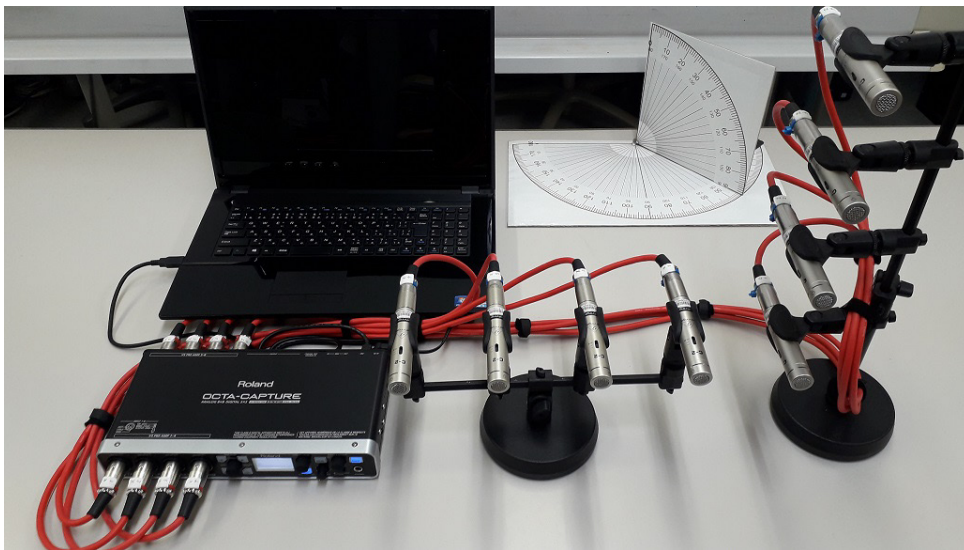


Fig. 5.9 Photograph of the microphone array system.

sources are increasing to three, we clearly see that the performance of IMUSIC, TOFS, TOPS, Squared-TOPS and WS-TOPS are significantly dominated by the number of incident sources, because those methods requires the number of sensor elements to much more higher than the number of sources to achieve reasonably good results, which can be verified by referring to the simulation results on Section 5.4 and Figs. 5.3 and 5.4. The proposed methods, however, are able estimate the DOA of three sources effectively and better than the selected methods. The reason is that the proposed methods focus on the subspace across all frequency bins simultaneously instead of focusing each frequency band individually, which is stated in Section 5.3.2. In case of Experiment 2 in Tables 5.5 and 5.6, the experiment results indicates that the proposed method with 2D-MUSIC exhibits extremely similar overall performances to 2D-IMUSIC and 2D-TOFS. As already stated in Section 5.4.5, the computational complexity of the proposed method is definitely lower than 2D-IMUSIC and 2D-TOFS by the fact that those methods check the orthogonality of subspace and search space of narrowband sample cross-correlation matrices for all frequency bins, resulting in very high computation requirement. Whereas the proposed method test the orthogonality of subspace and search

5.5 Experimental Results

Table 5.5 Performance evaluation on Experiment 2; $K = 1, 2, 3$

Incident Sources			RMSE of DOAs (Degree)		
Number	Position	Angle (Degree)	2D-IMUSIC	2D-TOFS	Proposed Method with 2D-MUSIC
1	ϕ_1	96	0.9000	0.9000	0.9000
	θ_1	86	0.4000	1.0500	0.7500
	Total		0.6500	0.9750	0.8250
2	ϕ_1	57	0.9500	1.1500	1.1000
	θ_1	91	1.0500	1.8000	1.7000
	ϕ_2	139	4.9500	5.2000	5.4500
	θ_2	96	3.1500	3.3000	2.0500
	Total		2.5250	2.8625	2.5750
3	ϕ_1	48	0.9500	1.5500	1.9500
	θ_1	86	1.4500	0.8000	2.4500
	ϕ_2	98	0.9000	1.8000	1.1500
	θ_2	95	1.4500	2.1500	2.6000
	ϕ_3	152	2.7000	2.4000	5.9000
	θ_3	95	4.5000	3.9000	1.4500
Total		1.9917	2.1000	2.5833	

space by using the wideband sample cross-correlation matrix on Eq. (5.56) instead of using the subspace of narrowband sample cross-correlation matrices for all frequency bins, but it is sufficient to exhibit significant effects as well as using the subspace of narrowband sample cross-correlation matrices for all frequency bins. In the end, the experiment results from Tables 5.3 to 5.6 are able to provide the evidence that proposed

5.6 Conclusions

Table 5.6 Performance evaluation on Experiment 2; $K = 4$

Incident Sources			RMSE of DOAs (Degree)		
Number	Position	Angle (Degree)	2D-IMUSIC	2D-TOFS	Proposed Method with 2D-MUSIC
4	ϕ_1	100	5.8095	6.5238	3.2857
	θ_1	94	2.4286	2.6190	1.6667
	ϕ_2	51	1.2381	1.0952	2.5714
	θ_2	95	0.5714	0.6667	1.3333
	ϕ_3	134	1.9524	1.8571	3.9524
	θ_3	103	10.0952	10.2857	9.2857
	ϕ_4	153	7.4762	7.8095	7.8571
	θ_4	89	4.7143	4.7143	5.3810
Total			4.2857	4.4464	4.4167

methods have better estimating performance than other methods with respect to the number of incident sources.

Since the sound source direction are static in Tables 5.3 to 5.6, it is necessary to consider moving sound sources for more practical use. In future work, we will extend the proposed method for moving sound sources, and further develop the prototype to support more realistic tasks.

5.6 Conclusions

An efficient framework for estimating DOA of wideband sound sources was presented. The issue of transforming multiple narrowband cross-correlation matrices for all frequency bins into a wideband cross-correlation matrix has been addressed suc-

5.6 Conclusions

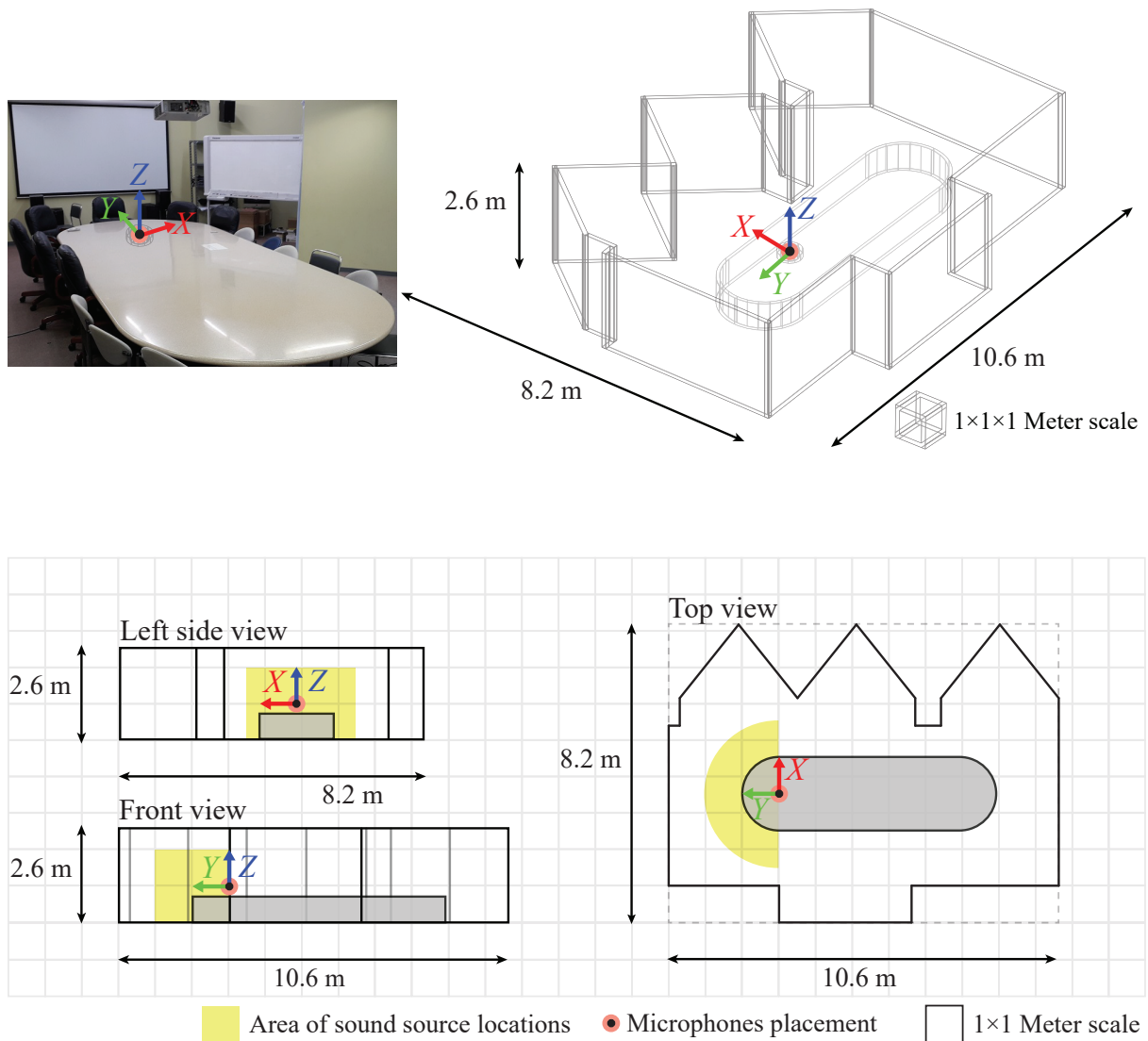


Fig. 5.10 Photograph of the experimental environment, floor plan and the room dimensions.

cessfully by focusing on signal subspace for all frequency bins simultaneously instead of pairing of temporal and reference frequency as done by the CSS-based methods. A new solution to this problem has been given by employing the extension theory of Orthogonal Procrustes analysis along with performing HOGSVD of array of the novel cross-correlation matrices, where elements in the row and column positions are a sample cross-correlation matrix between received signal and itself on two distinct frequencies.

5.7 References

It was shown in the theoretical analysis that the proposed transformation procedure provided the best solution under appropriate constraints, and no longer require any process of DOA preliminary estimation. Subsequently, we provided an alternative to construct the wideband cross-correlation matrix via the proposed transformation procedure, and wideband DOAs were estimated easily using this wideband matrix along with a single scheme of estimating DOAs in any narrowband subspace methods. A major contribution of this research work is that the proposed framework enables cutting-edge researches in the recent narrowband subspace methods to estimate DOA of the wideband sources directly, which result in the reducing computational complexity and facilitating the estimation algorithm. We also have performed several examples of using the proposed framework, such as, 2D-MUSIC, MUSIC, and ESPRIT method integration with the L-shaped microphone arrays. Furthermore, the simulation and experimental results showed that the fusion methods by using the proposed framework exhibited especially effective performance than other wideband DOA estimation methods over a range of SNR with much fewer sensors, high noise and reverberation conditions. We believe that the proposed method represents an efficient way of a wideband DOA estimation and would be able to improve wideband DOA estimates not only for acoustic signal processing but also other possible related fields.

5.7 References

- [1] Amr Abdelbari. *Direction of Arrival Estimation of Wideband RF Sources*. PhD thesis, Near East University, 12 2018.
- [2] Randy Diehl. Acoustic and auditory phonetics: The adaptive design of speech sound systems. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363:965–78, 04 2008.

5.7 References

- [3] Yang-Yang Dong and Xin Chang. Computationally efficient 2d doa estimation for l-shaped array with unknown mutual coupling. *Mathematical Problems in Engineering*, 2018:1–9, 01 2018.
- [4] Xiaofeng Gao, Xinhong Hao, Ping Li, and Guolin Li. An improved two-dimensional direction-of-arrival estimation algorithm for l-shaped nested arrays with small sample sizes. *Sensors*, 19(9), 2019.
- [5] John C Gower and Garnt B Dijkstrahuis. *Procrustes Problems*. Oxford University Press, 1st edition, 2004.
- [6] Simon Haykin and K.J. Ray Liu. *Handbook on Array Processing and Sensor Networks*. Wiley-IEEE Press, 2010.
- [7] Hayashi Hirotsuka and Ohtsuki Tomoaki. Doa estimation for wideband signals based on weighted squared tops. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):243, Oct 2016.
- [8] L. Hogben. *Handbook of Linear Algebra*. Discrete Mathematics and Its Applications. CRC Press, 2006.
- [9] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 2 edition, 2012.
- [10] Kai-Chieh Hsu and Jean-Fu Kiang. Joint estimation of doa and frequency of multiple sources with orthogonal coprime arrays. *Sensors*, 19(2), 2019.
- [11] Y. Hua, T. K. Sarkar, and D. D. Weiner. An l-shaped array for estimating 2-d directions of wave arrival. *IEEE Transactions on Antennas and Propagation*, 39(2):143–146, Feb 1991.
- [12] H. Hung and M. Kaveh. Focussing matrices for coherent signal-subspace processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(8):1272–1281, Aug 1988.
- [13] C. Knapp and G. Carter. The generalized correlation method for estimation of time

5.7 References

- delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, Aug 1976.
- [14] Eric A. Lehmann and Anders M. Johansson. Prediction of energy decay in room impulse responses simulated with an image-source model. *The Journal of the Acoustical Society of America*, 124(1):269–277, Jul 2008.
- [15] J. Li and D. Jiang. Joint elevation and azimuth angles estimation for l-shaped array. *IEEE Antennas and Wireless Propagation Letters*, 16:453–456, 2017.
- [16] S. Marcos, A. Marsal, and M. Benidir. Performances analysis of the propagator method for source bearing estimation. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, pages IV/237–IV/240 vol.4, Apr 1994.
- [17] Sylvie Marcos, Alain Marsal, and Messaoud Benidir. The propagator method for source bearing estimation. *Signal Processing*, 42(2):121–138, 1995.
- [18] Carl D. Meyer, editor. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.
- [19] B. Nadler. Nonparametric detection of signals by information theoretic criteria: Performance analysis and an improved estimator. *IEEE Transactions on Signal Processing*, 58(5):2746–2756, May 2010.
- [20] P.S. Naidu. *Sensor Array Signal Processing*. Taylor & Francis, 2000.
- [21] Xi Nie and Ping Wei. Array aperture extension algorithm for 2-d doa estimation with l-shaped array. *Progress In Electromagnetics Research Letters*, 52:63–69, 2015.
- [22] K. Okane and T. Ohtsuki. Resolution improvement of wideband direction-of-arrival estimation ”squared-tops”. In *2010 IEEE International Conference on Communications*, pages 1–5, May 2010.
- [23] M. Omer, A. A. Quadeer, T. Y. Al-Naffouri, and M. S. Sharawi. An l-shaped microphone array configuration for impulsive acoustic source localization in 2-d using

5.7 References

- orthogonal clustering based time delay estimation. In *2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, pages 1–6, Feb 2013.
- [24] Sri Priya Ponnappalli, Michael A. Saunders, Charles F. Van Loan, and Orly Alter. A higher-order generalized singular value decomposition for comparison of global mrna expression from multiple organisms. *PLOS ONE*, 6(12):1–11, 12 2011.
- [25] Sri Priya Ponnappalli, Michael A. Saunders, Charles F. Van Loan, and Orly Alter. A higher-order generalized singular value decomposition for comparison of global mrna expression from multiple organisms. *PLOS ONE*, 6(12):1–11, 12 2011.
- [26] M. G. Porozantidou and M. T. Chryssomallis. Azimuth and elevation angles estimation using 2-d music algorithm with an l-shape antenna. In *2010 IEEE Antennas and Propagation Society International Symposium*, pages 1–4, July 2010.
- [27] R. Roy and T. Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, Jul 1989.
- [28] H. Sawada, R. Mukai, S. Araki, and S. Makino. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *IEEE Transactions on Speech and Audio Processing*, 12(5):530–538, Sep. 2004.
- [29] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, Mar 1986.
- [30] Guanling Su and M. Morf. The signal subspace approach for multiple wide-band emitter location. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(6):1502–1522, Dec 1983.
- [31] Yuya Sugimoyo, Shigeki Miyabe, Takeshi Yamada, Shoji Makino, and Biing-Hwang Juang. An extension of music exploiting higher-order moments via nonlinear mapping. *IEICE Transactions on Fundamentals of Electronics, Communications and*

5.7 References

- Computer Sciences*, E99.A(6):1152–1162, 2016.
- [32] B. Suksiri and M. Fukumoto. A computationally efficient wideband direction-of-arrival estimation method for l-shaped microphone arrays. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5, May 2018.
- [33] Bandhit Suksiri and Masahiro Fukumoto. Multiple frequency and source angle estimation by gaussian mixture model with modified microphone array data model. *Journal of Signal Processing*, 21(4):163–166, 2017.
- [34] Nizar Tayem. Azimuth/elevation directional finding with automatic pair matching. *International Journal of Antennas and Propagation*, 2016.
- [35] S. Valaee, B. Champagne, and P. Kabal. Localization of wideband signals using least-squares and total least-squares approaches. *IEEE Transactions on Signal Processing*, 47(5):1213–1222, May 1999.
- [36] S. Valaee and P. Kabal. Wideband array processing using a two-sided correlation transformation. *IEEE Transactions on Signal Processing*, 43(1):160–172, Jan 1995.
- [37] A. . Van Der Veen, E. F. Deprettere, and A. L. Swindlehurst. Subspace-based signal analysis using singular value decomposition. *Proceedings of the IEEE*, 81(9):1277–1308, Sep. 1993.
- [38] Charles F. Van Loan. Structured matrix computations from structured tensors: Lecture 6. the higher-order generalized singular value decomposition, June 2015.
- [39] M. Wajid, A. Kumar, and R. Bahl. Direction-of-arrival estimation algorithms using single acoustic vector-sensor. In *2017 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, pages 84–88, Nov 2017.
- [40] G. Wang, J. Xin, N. Zheng, and A. Sano. Computationally efficient subspace-based method for two-dimensional direction estimation with l-shaped array. *IEEE Transactions on Signal Processing*, 59(7):3197–3212, July 2011.

5.7 References

- [41] H. Wang and M. Kaveh. Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(4):823–831, Aug 1985.
- [42] Qing Wang, Hang Yang, Hua Chen, Yangyang Dong, and Laihua Wang. A low-complexity method for two-dimensional direction-of-arrival estimation using an l-shaped array. *Sensors*, 17(1), 2017.
- [43] Tao Wu, Zhenghong Deng, Yiwen Li, Zhengxin Li, and Yijie Huang. Estimation of two-dimensional non-symmetric incoherently distributed source with l-shape arrays. *Sensors*, 19(5), 2019.
- [44] J. Xin, N. Zheng, and A. Sano. Simple and efficient nonparametric method for estimating the number of signals without eigendecomposition. *IEEE Transactions on Signal Processing*, 55(4):1405–1420, April 2007.
- [45] Kunihiro Yokoi and Nozomu Hamada. Ica-based separation and doa estimation of analog modulated signals in multipath environment. *IEICE Transactions on Communications*, 88-B:4246–4249, 11 2005.
- [46] Yeo-Sun Yoon, L. M. Kaplan, and J. H. McClellan. Tops: new doa estimator for wideband signals. *IEEE Transactions on Signal Processing*, 54(6):1977–1989, June 2006.
- [47] H. Yu, J. Liu, Z. Huang, Y. Zhou, and X. Xu. A new method for wideband doa estimation. In *2007 International Conference on Wireless Communications, Networking and Mobile Computing*, pages 598–601, Sept 2007.
- [48] Reza Zekavat and R. Michael Buehrer. *Handbook of Position Location: Theory, Practice and Advances*. Wiley-IEEE Press, 1st edition, 2011.

Chapter 6

Complex-Valued Tensor Factorization for Estimating Direction of Wideband Sources and its Variance-of-Frequencies

This chapter presents an alternative way to estimate a variance each temporal frequency, azimuth and elevation of wideband sources for an L-shaped microphone array. Signal model is renovated into a tensor representation of three features, x-subarray angle, variance-of-frequency, and z-subarray manifold angle. Azimuth and elevation are estimated simultaneously by employing a tensor factorization on the proposed signal model. In addition, complex-valued parallel factor analysis is utilized as the tensor factorization. The performance is evaluated in terms of overall root-mean-squared error over a range of signal-to-noise ratio. The proposed method provides a promising alternative for intelligent source localization.

6.1 Problem of Pair Matching Method on Variance-of-Frequency

In the recent past, DOA estimation methods have been applied for human computer interaction in acoustic signal processing as mentioned in Chapters 4 and 5. A geometrical structure of L-shaped sensor array is widely applied to two-dimensional (2-D) DOA estimation method because it allows for simple implementation with low uncertainty estimation as well as circular array [3]. Several approaches have been proposed as potential solution for wideband DOA estimation, for example, incoherent signal subspace (ISS) [4], coherent signal subspace (CSS) [5], test of orthogonality of frequency subspace (TOFS) [6], and test of orthogonality of projected subspace (TOPS) [7]. These methods hold great promise in high efficient 2-D DOA estimation due to its superior robustness, accuracy, and efficiency in comparison to other conventional approaches. On the contrary, it was previously found in the recent works that frequency information, known as a variance of the source each frequency bins, are computed separately for each temporal frequency [4–10]. It implies that the pair matching method between the temporal frequencies is definitely required for practical applications, for example, to determine the incident source frequencies correctly. From the above statement, a common shortcoming is clearly recognized as the requirement of pair matching method on variance-of-frequency.

Therefore the purpose of this chapter is to investigate a new wideband 2-D DOA estimation in a highly efficient way for acoustic sources and make frequency information more usable. The proposed method constructs the sample cross-correlation of incidents sources for all frequency bins via third-order tensor representation. Each lateral slice of the proposed tensor represents a sample cross-correlation matrix of temporal frequency, where transformation matrices are employed in order to fit the sample cross-correlation

6.2 Proposed Method

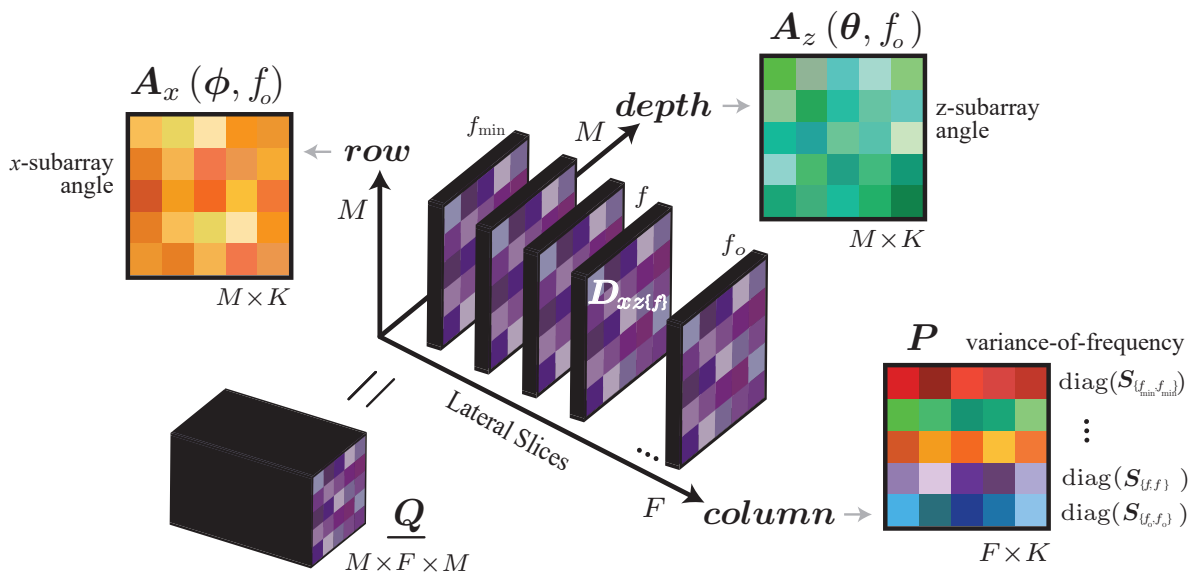


Fig. 6.1 The tensor representation of sample cross-correlation.

matrix into the proposed tensor correctly. Additionally, the transformation matrices are constructed by following a reported procedure [10]; performing a singular values decomposition (SVD) of a unique cross-correlation matrix, where elements in the row and column positions are sample cross-correlation matrices of two distinct frequencies from an identical coordinate axis. Complex-valued parallel factor analysis (COMFAC) is utilized as the tensor factorization [11,12], for estimating the following three features; x-subarray angle, self paired variance-of-frequency for all sources, and z-subarray manifold angle. Effectiveness of proposed method is substantiated through numerical simulations. In conclusion, the proposed method provides a promising alternative for intelligent source localization.

6.2 Proposed Method

Considers an acoustic source in the environment and an L-shaped array consisting of x and z axis as illustrated in Fig. 2.1, the received signal of each microphone is transformed into time-frequency representation via short-time Fourier transform [9,10].

6.2 Proposed Method

The received signal vectors from the sources is given as follows:

$$\begin{aligned}\mathbf{x}(t, f) &= \mathbf{A}_x(\boldsymbol{\phi}, f) \mathbf{s}(t, f) + \mathbf{w}_x(t, f), \\ \mathbf{z}(t, f) &= \mathbf{A}_z(\boldsymbol{\theta}, f) \mathbf{s}(t, f) + \mathbf{w}_z(t, f),\end{aligned}\tag{6.1}$$

where all variables are followed on Chapter 5: $\mathbf{x}(t, f) \in \mathbb{C}^M$ and $\mathbf{z}(t, f) \in \mathbb{C}^M$ are the summation of the received signal vectors for all sources corresponding to the x and z subarrays, $\mathbf{A}_x(\boldsymbol{\phi}, f) \in \mathbb{C}^{M \times K}$ and $\mathbf{A}_z(\boldsymbol{\theta}, f) \in \mathbb{C}^{M \times K}$ are the array manifold matrices in the x and z subarrays, $\mathbf{s}(t, f) \in \mathbb{C}^K$ is a signal source vector, and $\mathbf{w}_x(t, f) \in \mathbb{C}^M$ and $\mathbf{w}_z(t, f) \in \mathbb{C}^M$ are additive noise vectors for all the microphone elements corresponding to the x and z subarrays. Note that M, K represent the number of microphone elements each subarray, and the number of incident sources. ϕ_k, θ_k are x and z subarray angle. The parameters f, f_o denote source and reference frequencies, λ denotes a wavelength of sources, d denotes the spacing of the microphone elements any subarray, and t denotes captured time. Note that we followed the basic assumption guidelines as revealed in Chapter 5.

From the previous declaration, the sample cross-correlation matrix $\mathbf{R}_{xz\{f, f'\}} \in \mathbb{C}^{M \times M}$ is formulated between the x and z subarrays, as below

$$\begin{aligned}\mathbf{R}_{xz\{f, f'\}} &= E \{ \mathbf{x}(\mathbf{t}, f) \mathbf{z}^H(\mathbf{t}, f') \} \\ &= \mathbf{A}_x(\boldsymbol{\phi}, f) E \{ \mathbf{s}(\mathbf{t}, f) \mathbf{s}^H(\mathbf{t}, f') \} \mathbf{A}_z^H(\boldsymbol{\theta}, f') \\ &\quad + E \{ \mathbf{w}_x(\mathbf{t}, f) \mathbf{w}_z^H(\mathbf{t}, f') \} \\ &= \mathbf{A}_x(\boldsymbol{\phi}, f) \mathbf{S}_{\{f, f'\}} \mathbf{A}_z^H(\boldsymbol{\theta}, f'),\end{aligned}\tag{6.2}$$

where $\mathbf{S}_{\{f, f'\}} \in \mathbb{C}^{K \times K}$ is a sample cross-covariance matrix over two distinct frequencies. When $f' = f$, the diagonal elements of $\mathbf{S}_{\{f, f\}}$ are variance of the signal sources.

As mentioned in the introduction, transformation matrices are employed in order to fit the sample cross-correlation matrix into the proposed tensor appropriately. For the ideal case of transposition, the transformation matrices shape the sample cross-correlation matrix from any temporal frequency f to reference frequency f_o without

6.2 Proposed Method

changing the sample cross-covariance matrix. A transformed sample cross-correlation matrix $\mathbf{D}_{xz\{f\}}$ with the L-shaped structure is defined as

$$\begin{aligned}\mathbf{D}_{xz\{f\}} &= \mathbf{T}_{x\{f\}} \mathbf{R}_{xz\{f,f\}} \mathbf{T}_{z\{f\}}^H \\ &= \mathbf{A}_x(\phi, f_o) \mathbf{S}_{\{f,f\}} \mathbf{A}_z^H(\theta, f_o),\end{aligned}\tag{6.3}$$

where

$$\begin{aligned}\mathbf{A}_x(\phi, f_o) &= \mathbf{T}_{x\{f\}} \mathbf{A}_x(\phi, f), \\ \mathbf{A}_z(\theta, f_o) &= \mathbf{T}_{z\{f\}} \mathbf{A}_z(\theta, f),\end{aligned}\tag{6.4}$$

$\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}} \in \mathbb{C}^{M \times M}$ represent the transformation matrices in the x and z subarrays. The dual matrices $\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}}$ are originally designed by minimizing a Frobenius norm for the array manifold errors [5], and it can be alternatively designed along with the noise-free sample correlation matrices $\mathbf{R}_{xx\{f,f\}} - \mathbf{W}_{xx\{f,f\}}$ and $\mathbf{R}_{zz\{f,f\}} - \mathbf{W}_{zz\{f,f\}}$ where $\mathbf{W}_{xx\{f,f\}}, \mathbf{W}_{zz\{f,f\}}$ are noise covariance matrices. In our previous work, the alternative technique for transformation matrix estimation was proposed, and this alternative technique may hold promise to replace and improve an accuracy performance of the transformation matrix estimation as shown in Chapters 4 and 5.

After calculating the matrices $\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}}$ by the previously procedure, the sample matrix $\mathbf{D}_{xz\{f\}}$ is obtained each temporal frequency. We continue to introduce a new sample cross-correlation via tensor representation. Given an 3rd-order tensor $\underline{\mathbf{Q}} \in \mathbb{C}^{M \times F \times M}$ and the inner positive index K where F represents the number of frequency bins, the three-component matrices of the tensor $\underline{\mathbf{Q}}$ are expressed as

$$\underline{\mathbf{Q}} = [[\mathbf{A}_x(\phi, f_o), \mathbf{P}, \bar{\mathbf{A}}_z(\theta, f_o)]],\tag{6.5}$$

where $\bar{\mathbf{A}}_z(\theta, f_o) \in \mathbb{C}^{M \times K}$ denotes the complex conjugate of the elements of $\mathbf{A}_z(\theta, f_o)$. $\mathbf{P} \in \mathbb{R}_{\geq 0}^{F \times K}$ denotes the sample variance matrix for frequency range form f_{\min} to f_o ; we named as the variance-of-frequency. In particular, the variance-of-frequency matrix can

6.2 Proposed Method

be described as

$$\mathbf{P} = \begin{pmatrix} \sigma_{s_1\{f_{\min}\}}^2 & \sigma_{s_2\{f_{\min}\}}^2 & \cdots & \sigma_{s_K\{f_{\min}\}}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{s_1\{f\}}^2 & \sigma_{s_2\{f\}}^2 & \cdots & \sigma_{s_K\{f\}}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{s_1\{f_o\}}^2 & \sigma_{s_2\{f_o\}}^2 & \cdots & \sigma_{s_K\{f_o\}}^2 \end{pmatrix}, \quad (6.6)$$

where $\sigma_{s_k\{f\}}^2$ is a variance of the signal source $s_k(\mathbf{t}, f)$. Note that the matrices $\mathbf{D}_{xz\{f\}}$ are rearranged by lateral slices of the tensor $\underline{\mathbf{Q}}$ as illustrated in Fig. 6.1.

In fact, the method of calculating the transformation matrices $\mathbf{T}_{x\{f\}}, \mathbf{T}_{z\{f\}}$ from the following procedure [10], are based on SVD; therefore, the tensor $\underline{\mathbf{Q}}$ can be factorized into SVD form as $\underline{\mathbf{Q}} = [[\mathbf{U}, \mathbf{\Sigma}, \bar{\mathbf{V}}]]$ where $\mathbf{U}, \mathbf{V} \in \mathbb{C}^{M \times K}$ are the matrices of left and right singular vectors of $\mathbf{R}_{xz\{f_o, f_o\}}$ in the signal subspace. $\mathbf{\Sigma} \in \mathbb{R}_{\geq 0}^{F \times K}$ is the singular matrix, each column vectors of $\mathbf{\Sigma}$ are the singular values of $\mathbf{R}_{xz\{f, f\}}$ in the signal subspace under frequency range from f_{\min} to f_o . Although the singular value matrix $\mathbf{\Sigma}$ is not the variance-of-frequency matrix \mathbf{P} . When normalization processing is employed on $\mathbf{\Sigma}, \mathbf{P}$, the normalized matrix $\mathbf{\Sigma}$ can be considered as indicator of \mathbf{P} . Therefore, it is possible to factorize the tensor $\underline{\mathbf{Q}}$ as three-component matrices of $\mathbf{U}, \mathbf{\Sigma}$ and \mathbf{V} , respectively. Since the sample cross-correlation tensor $\underline{\mathbf{Q}}$ is now ready for extraction, parallel factor analysis structure is applied to isolate the variance-of-frequency component \mathbf{P} or $\mathbf{\Sigma}$ in $\underline{\mathbf{Q}}$. In this research work, we follow the procedure described by R. Bro and N.D. Sidiropoulos as shown in [11, 12], which is well-known as COMFAC structure.

Applying COMFAC to the tensor $\underline{\mathbf{Q}}$, we are now able to obtain the matrices \mathbf{U}, \mathbf{V} and the variance-of-frequency $\mathbf{\Sigma}$. According to useful technique [10], we form new matrices $\mathbf{\Gamma}_x$ and $\mathbf{\Gamma}_z$ to obtain the x and z subarray angles, as follows:

$$\begin{aligned} \mathbf{\Gamma}_x &= \mathbf{U}_1^\dagger \mathbf{U}_2 \\ &= \mathbf{F} \mathbf{\Phi}_x \mathbf{F}^{-1}, \\ \mathbf{\Gamma}_z &= \mathbf{V}_1^\dagger \mathbf{V}_2 \\ &= \mathbf{G} \mathbf{\Phi}_z \mathbf{G}^{-1}, \end{aligned} \quad (6.7)$$

6.3 Numerical Simulations

where

$$\begin{aligned}\mathbf{\Phi}_x &= \text{diag}(\alpha_x(\phi_1, f_o), \alpha_x(\phi_2, f_o), \dots, \alpha_x(\phi_K, f_o)), \\ \mathbf{\Phi}_z &= \text{diag}(\alpha_z(\theta_1, f_o), \alpha_z(\theta_2, f_o), \dots, \alpha_z(\theta_K, f_o)),\end{aligned}\tag{6.8}$$

$\mathbf{U}_1, \mathbf{U}_2, \mathbf{V}_1, \mathbf{V}_2 \in \mathbb{C}^{M-1 \times K}$ are the first and the last $M - 1$ rows of singular matrices \mathbf{U} and \mathbf{V} , respectively, $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{K \times K}$ denote the invertible matrices corresponding to $\mathbf{A}_x(\phi, f_o)$, $\mathbf{A}_z(\theta, f_o)$, and $\mathbf{\Phi}_x, \mathbf{\Phi}_z \in \mathbb{C}^{K \times K}$ are the diagonal matrices, which are structured by the conjugate symmetry property of array manifold matrices $\mathbf{A}_x(\phi, f_o)$, $\mathbf{A}_z(\theta, f_o)$. Applying Eigendecomposition to $\mathbf{\Gamma}_x, \mathbf{\Gamma}_z$, the x and z subarray angles can be estimated by the eigenvalues matrices $\mathbf{\Lambda}_x$ and $\mathbf{\Lambda}_z$ of $\mathbf{\Gamma}_x$ and $\mathbf{\Gamma}_z$, respectively;

$$\begin{aligned}\phi_k &= \cos^{-1}\left(\text{angle}(\lambda_{x_k}) \frac{\lambda}{2\pi d}\right), \\ \theta_k &= \cos^{-1}\left(\text{angle}(\lambda_{z_k}) \frac{\lambda}{2\pi d}\right),\end{aligned}\tag{6.9}$$

where $\lambda_{x_k}, \lambda_{z_k} \in \mathbb{C}$ are the k^{th} eigenvalue in $\mathbf{\Lambda}_x$ and $\mathbf{\Lambda}_z$. In the end, relation between the azimuth, elevation, x subarray, and z subarray angles can be explained and calculated by Euler's rotation theorem.

6.3 Numerical Simulations

The performance of proposed method is demonstrated via a computer simulation. Three uncorrelated source angles $(\theta_k^{\text{DOA}}, \phi_k^{\text{DOA}})$ were placed at $(41.41^\circ, 60^\circ)$, $(60^\circ, 45^\circ)$, and $(75.53^\circ, 30^\circ)$, where θ_k^{DOA} and ϕ_k^{DOA} are true DOA x and z subarray angles. The parameters were chosen as follows: the sampling frequency was 48 kHz, the microphone captured time was 2 second, the speed of sound c was assumed as 340 m/s, the spacing of microphone elements d was 5 cm, the minimum frequency f_{\min} was 100 Hz, the reference frequency f_o was 3.4 kHz.

Two scenarios are considered. In the first scenario, accuracy performance on variance-of-frequency estimation is considered. The three sources were the following

6.3 Numerical Simulations

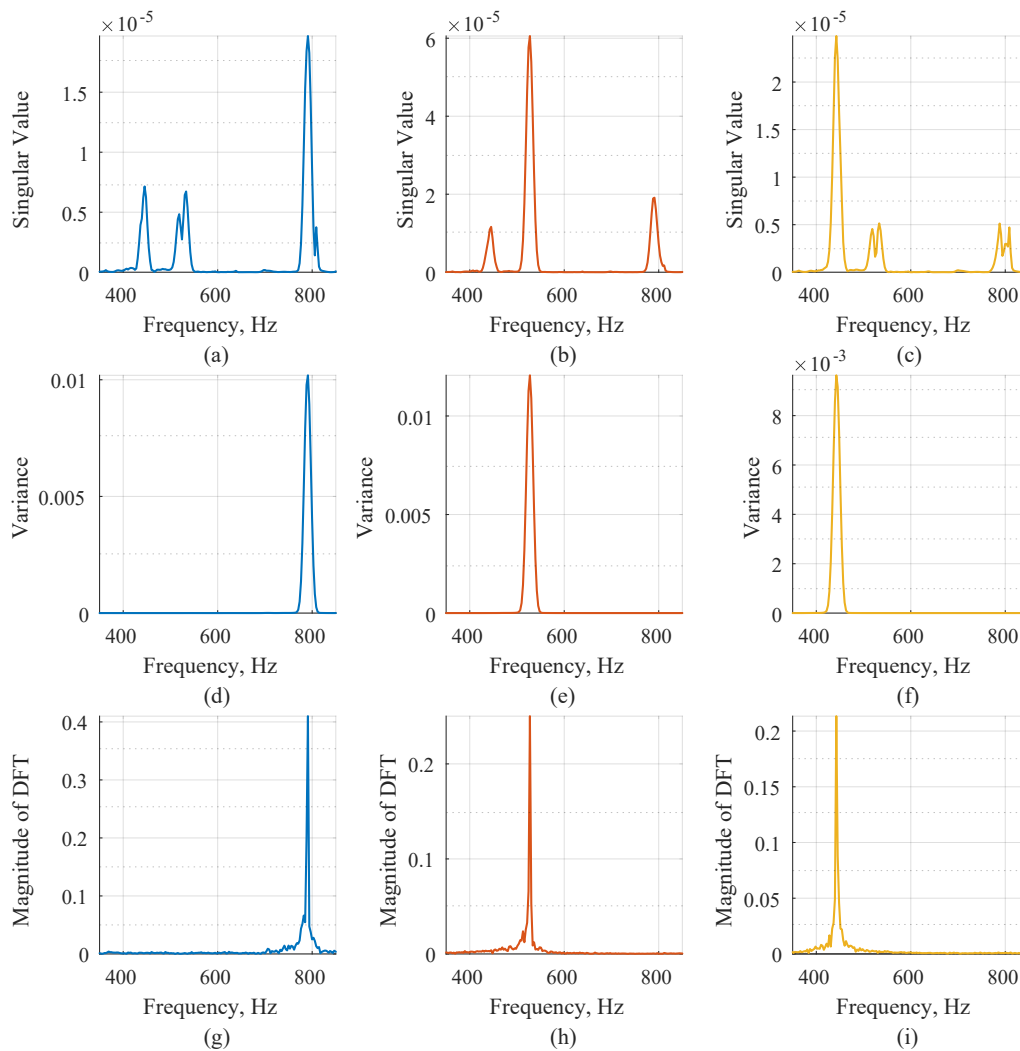


Fig. 6.2 Numerical simulation on variance-of-frequency estimation; (a), (b), (c) singular values or estimated variance-of-frequency, (d), (e), (f) actual variance-of-frequency, (g), (h), (i) actual magnitude of DFT, where (a), (d), (g) are the first source, (b), (e), (h) are the second source, (c), (f), (i) are the third source.

piano notes; G5 containing 783.99 kHz, C5 or Tenor C containing 523.25 kHz, A4 or A440 containing 440 kHz of main frequency, respectively. Fig. 6.2 shows numerical simulation on variance-of-frequency estimation; estimated variance-of-frequency, actual variance-of-frequency, and actual magnitude of discrete Fourier transform (DFT). When normalization processing is employed, it can be seen that the frequency at maximum peak of estimated variance-of-frequency exhibits extremely similar to actual variance-of-

6.4 Conclusion

frequency. Specifically, some weak false-peaks still remain at other frequency-dependent sources. This phenomenon can be explained by two factors; the estimation of transformation matrices, and the tensor modeling. Overall, it is possible to show that the estimated variance-of-frequency can be considered as indicator of the actual magnitude of DFT.

In the second scenario, we provided the performance evaluation of the proposed method in comparison with following methods: IMUSIC [4]; TOPS [7]; weighted squared TOPS (WS-TOPS) [8]; TOFS [6]; and CSS based on the dual optimization problems (CSS-DOP) [9]. The sources were human voices containing the frequency range from 0.1 to 16 kHz, and the number of Monte Carlo trials each scenario was 100 runs. The performance is evaluated in term of overall root-mean-squared error (RMSE) and standard division (SD) respect to the number of microphone elements each subarray M over a range of the signal-to-noise ratio (SNR) as shown in Fig. 6.3. Simulation results indicated that the proposed method exhibits extremely similar RMSE performance to CSS-DOP for all SNR range. Since SNR decrease from 20 dB, the proposed method and CSS-DOP exhibit better RMSE performance than other methods with respect to the number of microphone elements. This substantial ability is more meaningful for many practical applications.

6.4 Conclusion

A novel 2D wideband DOA estimation and variance-of-frequency methods for an L-shaped microphone array model were presented. We addressed a problem of estimating both x and z subarray angles and self paired variance-of-frequency by renovating the signal model into a tensor representation. COMFAC is utilized as the tensor factorization to find the solution of proposed signal model. Effectiveness of proposed method

6.4 Conclusion

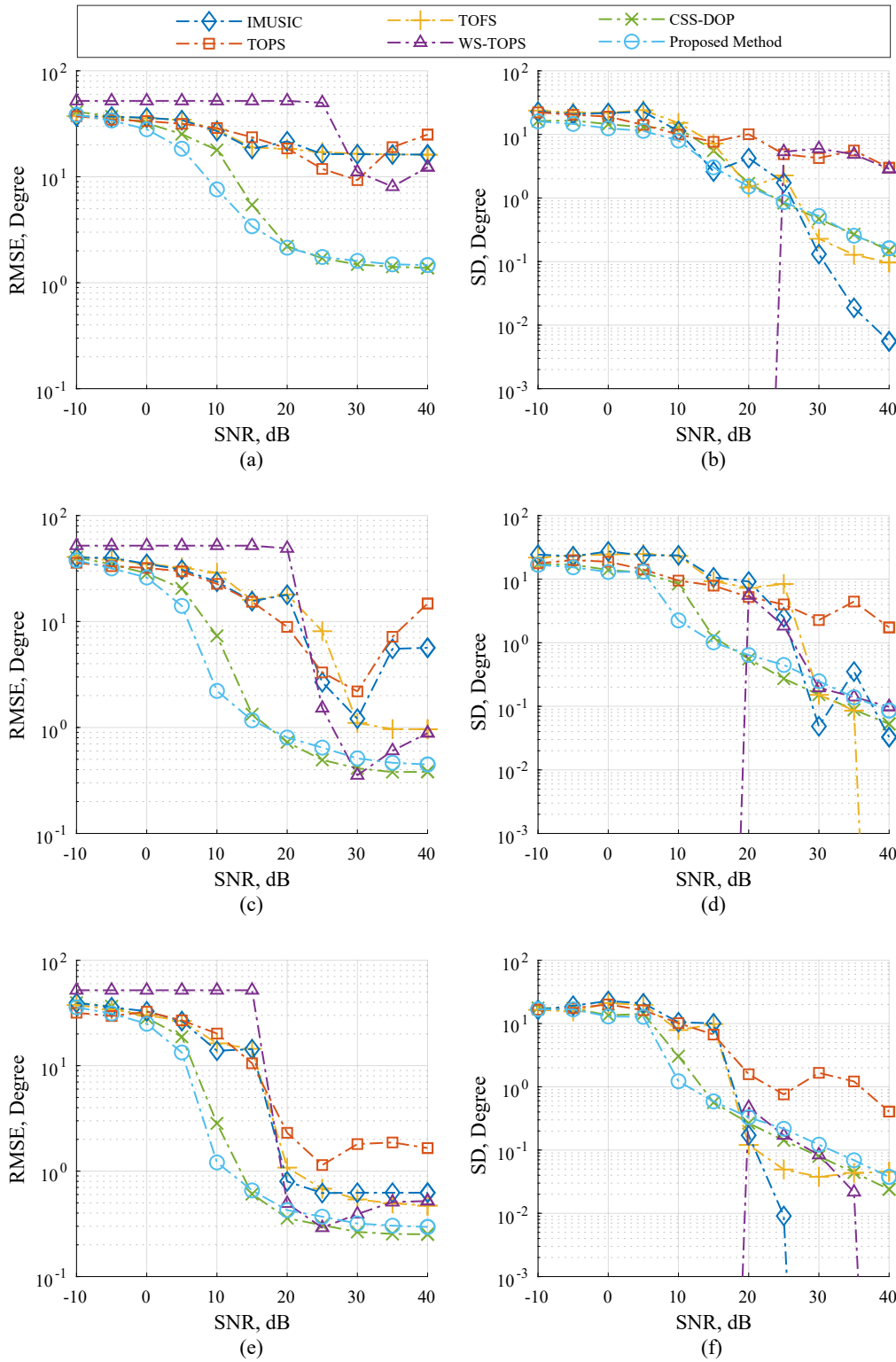


Fig. 6.3 Performance evaluations via simulation; (a), (c), (e) RMSE versus SNR and (b), (d), (f) SD versus SNR where the number of microphone elements each subarray on (a), (b) $M = 6$, (c), (d) $M = 8$, and (e), (f) $M = 10$.

6.5 References

were substantiated through numerical simulations, and the simulation results showed that the proposed method exhibited especially effective performance than other methods. Furthermore, the self paired variance-of-frequency for all sources can be estimated without a requirement of pair matching method.

6.5 References

- [1] S. Haykin and K.R. Liu, Handbook on Array Processing and Sensor Networks, Wiley-IEEE Press, 2010.
- [2] R. Zekavat and R.M. Buehrer, Handbook of Position Location: Theory, Practice and Advances, 1st ed., Wiley-IEEE Press, 2011.
- [3] Y. Hua, T.K. Sarkar, and D.D.Weiner, “ An l-shaped array for estimating 2-d directions of wave arrival, ” IEEE Transactions on Antennas and Propagation, vol.39, no.2, pp.143–146, Feb 1991.
- [4] G. Su and M. Morf, “ The signal subspace approach for multiple wideband emitter location, ” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol.31, no.6, pp.1502–1522, Dec 1983.
- [5] H. Wang and M. Kaveh, “ Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wideband sources, ” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol.33, no.4, pp.823–831, Aug 1985.
- [6] H. Yu, J. Liu, Z. Huang, Y. Zhou, and X. Xu, “ A new method for wideband doa estimation, ” 2007 International Conference on Wireless Communications, Networking and Mobile Computing, pp.598–601, Sept 2007.
- [7] Y.S.Yoon, L.M. Kaplan, and J.H. McClellan, “Tops: newdoa estimator for wideband signals, ” IEEE Transactions on Signal Processing, vol.54, no.6, pp.1977–1989, June 2006.

6.5 References

- [8] H. Hirota and O. Tomoaki, “Doa estimation for wideband signals based on weighted squared tops,” *EURASIP Journal on Wireless Communications and Networking*, vol.2016, no.1, p.243, Oct 2016.
- [9] B. Suksiri and M. Fukumoto, “A computationally efficient wideband direction-of-arrival estimation method for l-shaped microphone arrays,” *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp.1–5, May 2018.
- [10] B. Suksiri and M. Fukumoto, “A highly efficient wideband two-dimensional direction estimation method with l-shaped microphone array,” manuscript in preparation.
- [11] R. Bro, N.D. Sidiropoulos, and G.B. Giannakis, “A fast least squares algorithm for separating trilinear mixtures,” *International Workshop Independent Component Analysis & Signal Separation*, Jan 1999.
- [12] N.D. Sidiropoulos, G.B. Giannakis, and R. Bro, “Blind PARAFAC receivers for DS-CDMA systems,” *IEEE Transactions on Signal Processing*, vol.48, no.3, Mar 2000.

Chapter 7

Conclusion

An extension of techniques, new framework and suitable theory for estimating acoustic direction-of-arrivals (DOAs) have been presented. This dissertation aimed to bridge a research gap of acoustic source compatibility on the recent narrowband and wideband subspace methods to estimate DOA of the acoustic sources directly and effectively, which is a major contribution of this dissertation.

Chapter 2 presented an alternative signal modeling for wideband sources with L-shaped microphone array configuration. The problem of estimating multi-narrowband DOAs was resolved by using the proposed signal model along with remodeling the array manifold matrices. Extension of most classical subspace-based methods for multi-narrowband DOA estimation were given. Simulation results indicated that the new array manifold matrices by using the proposed signal model along with the classical subspace-based methods have achieved the multi-narrowband DOA estimation with acceptable signal-to-noise ratio (SNR) levels.

Chapter 3 presented an alternative DOA estimation method for wideband sources by using a Gaussian mixture model with a maximum likelihood estimation algorithm for multiple frequencies and source angle estimation. The problem of estimating DOA of wideband signals have been addressed by focusing the entire observation subspace in each frequency bin along with employing a multi-narrowband signal model. The performance is evaluated in terms of the root-mean-squared error (RMSE) over a range of SNR, and the results suggested that the proposed method is a particularly effective

method of DOA estimation. Furthermore, the proposed method enables the synthesis of signal sources and provides an alternative to intelligent source localization systems.

Chapter 4 presented an improved version of the previous method for estimating wideband two-dimensional (2D) DOA much more efficiently than the previously. We proposed a way to construct a wideband sample cross-correlation matrix and addressed a problem of estimating transformation matrices without any process of DOA preliminary estimation. The proposed transformation matrices were constructed by performing singular value decomposition (SVD) of a new unique cross-correlation matrix, where elements in the row and column positions are sample cross-correlation matrices of two different frequencies. Wideband DOAs can be estimated by using this wideband sample cross-correlation matrix along with a scheme of estimating DOA in a narrowband subspace method. Theoretical analysis and effectiveness of the proposed method are substantiated through numerical simulations, Furthermore, the results show that the proposed method exhibited effective performance than other wideband methods over a range of SNR with just a few microphones in reverberating environments.

In Chapter 5, an efficient framework for estimating DOA of wideband sound sources was presented by integrating our previous findings in Chapter 4. The issue of transforming multiple narrowband cross-correlation matrices for all frequency bins into a wideband cross-correlation matrix has been addressed successfully by employing the extension theory of Orthogonal Procrustes analysis along with performing high order singular value decomposition (HOGSVD) of array of the novel cross-correlation matrices, where elements in the row and column positions are a sample cross-correlation matrix between received signal and itself on two distinct frequencies. It was shown in the theoretical analysis that the proposed transformation procedure provided the best solution under appropriate constraints, and no longer require any process of DOA preliminary estimation. A major contribution of this research work is that the proposed

framework enables cutting-edge researches in the recent narrowband subspace methods to estimate DOA of the wideband sources directly, which result in the reducing computational complexity and facilitating the estimation algorithm. We also have performed several examples of using the proposed framework, such as, 2D-MUSIC, MUSIC, and ESPRIT method integration with the L-shaped microphone arrays. Furthermore, the simulation and experimental results showed that the fusion methods by using the proposed framework exhibited especially effective performance than other wideband DOA estimation methods over a range of SNR with much fewer sensors, high noise and reverberation conditions.

Finally, a novel 2D wideband DOA and variance-of-frequency estimation methods for an L-shaped microphone array model were presented in Chapter 6. We addressed a problem of estimating both x and z subarray angles and self paired variance-of-frequency by renovating the signal model into a tensor representation. Effectiveness of proposed method were substantiated through numerical simulations, and the simulation results showed that the proposed method exhibited especially effective performance than other methods. Furthermore, the self paired variance-of-frequency for all sources can be estimated without a requirement of pair matching method.

We believe that our researches presented in this dissertation will appeal to researchers who wish to develop a sound source based navigation system and improve its robust estimation. We also hope that this new framework and theory can be a good alternative for estimating DOA of acoustic sources, especially human speeches and musical sounds.

Acknowledgement

The author is most grateful and foremost to his advisor, Professor Masahiro Fukumoto, for his valuable supervision, supports, encouragements throughout the study and mentoring me over the course of my graduate studies. He has helped me through extremely difficult times over the course and the revising of manuscripts, and for that I sincerely thank him for his confidence in me. I would additionally like to thank Associate Professor Shinichi Yoshida, and Associate Professor Toru Kurihara for his support in both the research and especially the life in Japan. His knowledge and understanding of the machine learning fields and signal processing have allowed me to fully express the concepts behind this research. Special thanks also to Professor Masayoshi Tachibana, and Associate Professor Yukinobu Hoshino, members of dissertation committee, for their valuable suggestions and comments. The author wishes to acknowledge Professor Emeritus Lawrie Hunter, Professor Paul Daniels, and Professor Gordon Bateson for his valuable guidance in research writing.

My special thanks goes to members of International Relations Center (IRC) for their kindly support. The author wishes to acknowledge Kochi University of Technology (KUT), Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Japan Student Services Organization (JASSO) for a great opportunity of financial supports over a period of three years.

Sincere appreciation is extended to all Japanese friends in Kochi and to colleagues, including in the Signal Processing & New Generation Network Laboratory for their useful technical experience sharing and kind technical assistance. This research would not have been possible without the assistance of laboratory member who constructed the experimental apparatus and built the foundations for data analysis.

Acknowledgement

The author sincerely appreciates all of his Thai friends in Kochi for their friendships and goodwill. Especially, I would like to thank my best friend, Dr. Chiramathe Nami who shared this journey with me. Without him, I could never have completed many research works. Thank you for inspiring me, brother; “Allahu Akbar!” .

I also would like to thank Franz Vonlichten and Helmut Vonlichten from “E.S. Posthumus” for his immeasurable bravery music during these three years of studies in university. Thank you for makes me braver when I have fallen in the dark.

Finally, I would like to express my sincere gratitude to my mother and little cute brother. Without whose love, support and understanding I could never have completed this program. Thank you, Mom, brother, always and forever.

Author's Biography



Bandhit Suksiri received B.E. in Computer Engineering from Thai-nichi Institute of Technology, Thailand, in 2014, and M.S. in Information Engineering from Kochi University of Technology, Japan in 2016. He is currently working toward a D.Eng. in Information System Engineering, Kochi University of Technology, Japan, for a period of three years. His research interest includes applied mathematics, embedded system, acoustic, DSP, array processing, and machine learning in acoustic signal processing. He is a student member of IEEE SPS and CAS. Currently, he is planning to work in Tokyo as a signal processing engineer in Japanese robotics company.

List of Publications

Journals

- Bandhit Suksiri and Masahiro Fukumoto, “An Efficient Framework for Estimating Direction of Multiple Sound Sources using Higher-Order Generalized Singular Value Decomposition,” *Sensors* (Submitted for publication: May 28, 2019).

Acknowledgement

- Bandhit Suksiri and Masahiro Fukumoto, “A Highly Efficient Wideband Two-Dimensional Direction Estimation Method with L-Shaped Microphone Array,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (IEICE-EA)* (Conditional accepted: October 1, 2018; Latest revised: May 28, 2019; Expected publication: mid July 2019).
- Bandhit Suksiri and Masahiro Fukumoto, “Multiple Frequency and Source Angle Estimation by Gaussian Mixture Model with Modified Microphone Array Data Model,” *Journal of Signal Processing*, vol.21, no.4, pp.163–166, 2017, doi: 10.2299/jsp.21.163 (Revised: April 10, 2017; Accepted: June 21, 2017; Published: July 20, 2017).

International Conference Proceedings

- Bandhit Suksiri and Masahiro Fukumoto, “Wideband Direction-of-Arrival Estimation with Cross-Sample Matching Technique on L-Shaped Microphone Arrays,” 2019 (16th) International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Pattaya, Thailand, 2019, pp. 1-4 (planned to make presentation on July 10 – 13, 2019). The conference proceedings will appear in IEEE Xplore Digital Library.
- Bandhit Suksiri and Masahiro Fukumoto, “A Computationally Efficient Wideband Direction-of-Arrival Estimation Method for L-Shaped Microphone Arrays,” 2018 (50th) IEEE International Symposium on Circuits and Systems (ISCAS, 50th anniversary), Florence, Italy, 2018, pp. 1-5, doi: 10.1109/ISCAS.2018.8350972 (Acceptance rate: 53%). We received “Student Travel Grant” prize; offering a small number of supplemental travel grants to eligible and selected student authors who plan to present and publish an accepted paper at the conference. The conference proceedings appeared on IEEE Xplore Digital Library.

Acknowledgement

- Bandhit Suksiri and Masahiro Fukumoto, “Enhanced array manifold matrices for L-shaped microphone array-based 2-D DOA estimation,” 2017 (9th) Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Malaysia, 2017, pp. 955-960, doi: 10.1109/APSIPA.2017.8282160 (Acceptance rate: 70%). The conference proceedings appeared on IEEE Xplore Digital Library.
- Bandhit Suksiri and Masahiro Fukumoto, “Multiple Frequency and Source Angle Estimation by Gaussian Mixture Model with Modified Microphone Array Data Model,” 2017 (14th) RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP’17), Guam, 2017. We received “Student Paper Award” ; recognizing excellence in a paper presented at the conference.
- Bandhit Suksiri and Masahiro Fukumoto, “Implementation of Artificial Neural Network and Multilevel of Discrete Wavelet Transform for Voice Recognition”, has been selected for publishing as chapter in Computer and Information Science, Springer, 2016, and delivered to make presentation on 2016 (15th) IEEE/ACIS International Conference on Computer and Information Science, Japan, 2016 (during master ’ s degree at Kochi University of Technology, Japan).

Domestic Conference Proceedings

- Bandhit Suksiri and Masahiro Fukumoto, “Complex-valued Tensor Factorization on Wideband Two-dimensional Direction Estimation Method with L-Shaped Microphone Array,” 33rd SIP SYMPOSIUM, Japan, 2018 (presenting in Japanese).
- Bandhit Suksiri and Masahiro Fukumoto, “A Novel L-Shaped Microphone Array-based Wideband Direction of Arrival Estimation Method using the Special Cross-correlation Matrix,” 32nd SIP SYMPOSIUM, Japan, 2017 (presenting in English).
- Bandhit Suksiri and Masahiro Fukumoto, “Wavelet Analysis for Multiple Frequency

Acknowledgement

and Signal Classification in Linear Phased Array Model,” 31st SIP SYMPOSIUM, Japan, 2016 (presenting in English).

- Bandhit Suksiri and Masahiro Fukumoto, “Voice Recognition using Signal Clustering Neural Network with Wavelet Transform Feature Extraction,” Technical Report, IEICE SIP2015-143, Japan, 2016 (presenting in English, during master ’ s degree at Kochi University of Technology, Japan).

Awards & Achievements

- Role of reviewer on “Signal Processing” section (be a reviewer), 2019 (51st) IEEE International Symposium on Circuits and Systems (ISCAS), Hokkaido, Japan, 2019
- Student Travel Grant, 2018 (50th) IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 2018
- Student Paper Award, 2017 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP’17), Guam, 2017
- Sakuma Award, Perfect Research & Academic Record, Kochi University of Technology, 2016 (during master ’ s degree at Kochi University of Technology, Japan)
- Summa Cum Laude (First Class Honors), Thai Nichi Institute of Technology (during undergraduate degree at Thai-Nichi Institute of Technology, Thailand, Thailand)