

Washington University in St. Louis

Washington University Open Scholarship

All Theses and Dissertations (ETDs)

5-24-2012

Underwater Direction-of-Arrival Finding: Maximum Likelihood Estimation and Performance Analysis

Tao Li

Washington University in St. Louis

Follow this and additional works at: <https://openscholarship.wustl.edu/etd>

Recommended Citation

Li, Tao, "Underwater Direction-of-Arrival Finding: Maximum Likelihood Estimation and Performance Analysis" (2012). *All Theses and Dissertations (ETDs)*. 712.

<https://openscholarship.wustl.edu/etd/712>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS
School of Engineering and Applied Science
Department of Electrical & Systems Engineering

Dissertation Examination Committee:

Dr. Arye Nehorai, Chair

Dr. Nan Lin

Dr. Hiro Mukai

Dr. Joseph O'Sullivan

Dr. William Shannon

Dr. Jung-Tsung Shen

Underwater Direction-of-Arrival Finding:
Maximum Likelihood Estimation and Performance Analysis

by

Tao Li

A dissertation presented to the Graduate School of Arts and Sciences
of Washington University in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012
Saint Louis, Missouri

copyright by

Tao Li

2012

ABSTRACT OF THE DISSERTATION

Underwater Direction-of-Arrival Finding:
Maximum Likelihood Estimation and Performance Analysis

by

Tao Li

Doctor of Philosophy in Electrical Engineering

Washington University in St. Louis, May 2012

Research Advisor: Dr. Arye Nehorai

In this dissertation, we consider the problems of direction-of-arrival (DOA) finding using acoustic sensor arrays in underwater scenarios, and develop novel signal models, maximum likelihood (ML) estimation methods, and performance analysis results.

We first examine the underwater scenarios where the noise on sensor arrays are spatially correlated, for which we consider using sparse sensor arrays consisting of widely separated sub-arrays and develop ML DOA estimators based on the Expectation-Maximization scheme. We examine both zero-mean and non-zero-mean Gaussian incident signals and provide detailed estimation performance analysis. Our results show that non-zero means in signals improve the accuracy of DOA estimation.

Then we consider the problem of DOA estimation of marine vessel sources such as ships, submarines, or torpedoes, which emit acoustic signals containing both sinusoidal and random components. We propose a mixed signal model and develop an ML estimator for narrow-band DOA finding of such signals and then generalize the

results to the wide-band case. We provide thorough performance analysis for the proposed signal model and estimators. We show that our mixed signal model and ML estimators improve the DOA estimation performance in comparison with the typical stochastic ones assuming zero-mean Gaussian signals.

At last, we derive a Barankin-type bound (BTB) on the mean-square error of DOA estimation using acoustic sensor arrays. The typical DOA estimation performance evaluation are usually based on the Cramér-Rao Bound (CRB), which cannot predict the threshold region of signal-to-noise ratio (SNR), below which the accuracy of the ML estimation degrades rapidly. Identification of the threshold region has important applications for DOA estimation in practice. Our derived BTB provides an approximation to the SNR threshold region.

Acknowledgments

I am grateful to my PhD advisor, Dr. Arye Nehorai, for his valuable guidance throughout my doctoral research.

I would also like to thank my dissertation committee members, Dr. Nan Lin, Dr. Hiro Mukai, Dr. Joseph O'Sullivan, Dr. William Shannon, and Dr. Jung-Tsung Shen, for carefully revising and providing constructive suggestions to this dissertation.

I express my sincere gratitude to my instructors at Washington University in St. Louis, including Dr. Tzyh-Jong Tarn, Dr. Daniel Fuhrmann, Dr. Joseph O'Sullivan, Dr. Kenneth Kelton, Dr. Henric Krawczynski, and the late Dr. Christopher Byrnes, for helping me construct a solid knowledge background for my research work.

I offer my thanks and regards to all my labmates. Their friendships and aids made our lab a great place for work and research.

I convey my deepest gratitude to my parents and family in China for their endless love and emotional encouragement.

Tao Li

Washington University in Saint Louis
May 2012

To my parents.

Contents

Abstract	ii
Acknowledgments	iv
List of Figures	ix
1 Introduction	1
1.1 Formulation	1
1.2 Maximum Likelihood Estimation	3
1.3 Our Contributions	5
1.4 Outline of the Dissertation	7
2 Maximum Likelihood Direction-of-Arrival Estimation in Spatially Colored Noise Using Sparse Arrays	9
2.1 Introduction	10
2.2 Measurement Models	12
2.3 Maximum Likelihood Estimation	15
2.3.1 Deterministic ML Estimator	15
2.3.2 Stochastic ML Estimator	19
2.3.3 Mixed ML Estimator	24
2.4 Analytical Performance Analysis	27
2.4.1 Cramér-Rao Bounds and Asymptotic Errors	28
2.4.2 Analytical Performance Comparisons	32
2.5 Numerical Examples	34
2.6 Summary	41
3 Narrow-Band Direction-of-Arrival Estimation of Hydroacoustic Signals From Marine Vessels	42
3.1 Introduction	43
3.2 Measurement Model	45
3.3 Maximum Likelihood Estimation	46
3.4 Analytical Performance Analysis	51
3.5 Numerical Examples	55
3.6 Summary	59
4 Direction-of-Arrival Finding of Wide-Band Hydroacoustic Signals From Marine Vessels: An Extension from the Narrow-Band Case	60

4.1	Introduction	61
4.2	Measurement Model And Maximum Likelihood Estimation	62
4.2.1	Measurement Model	62
4.2.2	Maximum Likelihood Estimation	64
4.3	Analytical Performance Analysis	66
4.4	Numerical Example	68
4.5	Summary	70
5	Direction-of-Arrival Estimation of Hydroacoustic Signals From Marine Vessels: An Approach From Fourier Transform	71
5.1	Measurement Models And DOA Estimation	72
5.1.1	Measurement Model	72
5.1.2	DOA Estimation From A Modified Model	74
5.2	Analytical Performance Analysis	78
5.3	Numerical Example	80
5.4	Summary	82
6	A Barankin-Type Bound on Direction-of-Arrival Estimation	83
6.1	Introduction	84
6.2	Measurement Model	85
6.3	Barankin-Type Bound	87
6.4	Numerical Examples	93
6.4.1	Examples for Scalar-Sensor Array	93
6.4.2	Examples for Vector-Sensor Array	95
6.5	Summary	97
7	Conclusions	99
7.1	Key Contributions	99
7.2	Future Work	101
Appendix A	Proof of Theorem 1	102
Appendix B	Proof of Theorem 2	106
Appendix C	Proof of Lemma 2	112
Appendix D	Proof of Proposition 5	113
Appendix E	Proof of Proposition 6	115
Appendix F	Proof of Lemma 3	117
Appendix G	Proof of Equation (3.8)	119
Appendix H	Proof of Proposition 8	121

Appendix I	Proof of Proposition 9	124
Appendix J	Proof of Proposition 10	132
Appendix K	Proof of Proposition 11	139
Appendix L	Proof of Proposition 12	144
Appendix M	Proof of Proposition 13	147
Appendix N	Proof of Proposition 14	151
Appendix O	Proof of Proposition 16	156
Appendix P	Derivation of Equation (6.16)	160
References	162
Vita	171

List of Figures

2.1	Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from stochastic signal one-dimensional DOA estimation.	37
2.2	Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from mixed signal one-dimensional DOA estimation.	38
2.3	Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from stochastic signal two-dimensional DOA estimation with the diagonal blocks equal in the noise covariance matrix.	39
2.4	Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from mixed signal two-dimensional DOA estimation with the diagonal blocks equal in the noise covariance matrix.	40
3.1	Cramér-Rao bound and root-mean-square errors as functions of the number of measurements from mixed signal DOA estimation under spatially white noise using uniform linear array.	56
3.2	Cramér-Rao bound and root-mean-square errors as functions of the number of measurements from mixed signal DOA estimation under spatially colored noise using linear array consisting of separated sub-arrays.	58
4.1	Cramér-Rao bound and root-mean-square errors as functions of the number of measurements from wide-band mixed signal DOA estimation under spatially white noise.	69
5.1	Cramér-Rao bound and root-mean-square errors as functions of the number of subintervals N_d for mixed signal DOA estimation.	81
6.1	Square roots of the derived Barankin-type bound and maximum likelihood estimation mean-square error versus signal-to-noise ratio for narrow-band direction of arrival estimation using uniform linear scalar-sensor array.	94
6.2	Same as in Fig.1, but for the wide-band direction of arrival estimation.	94

6.3	Square roots of the derived Barankin-type bound, Cramér-Rao bound, and maximum likelihood estimation mean-square error versus signal-to-noise ratio for narrow-band direction of arrival estimation using uniform linear acoustic vector-sensor arrays with $d = \lambda$, 2λ , and 3λ , respectively.	96
6.4	Variation of maximum likelihood ambiguity functions with respect to the inter-sensor distance d	97

Chapter 1

Introduction

Direction-of-arrival (DOA) estimation aims at finding the direction from which the signals impinge on the sensor array, which consists of a group of sensors arranged in a specific geometry that are able to measure the values of the impinging signals. In the underwater environment, the impinging signals are often acoustic signals, which are commonly called hydroacoustic signals whose DOAs can be estimated by acoustic sensor arrays. Underwater DOA estimation has important applications in the detecting, localizing, and tracking of marine vessels like ships, submarines, and torpedoes. In this dissertation, we propose new signal models, maximum likelihood (ML) estimation methods, and performance analysis results for underwater DOA estimation problems.

1.1 Formulation

According to their bandwidths, the incident signals can be classified into narrow-band and wide-band signals. The bandwidth of a narrow-band signal is small such that it

can be considered as sinusoidal. Usually a signal can be considered as narrow-band if

$$D/c \ll 1/B, \quad (1.1)$$

where D , c , and B are the array length, signal propagation speed, and signal bandwidth, respectively [1]-[3]. Suppose there are L signals incident on an array of M sensors from the far field, which means the incident signal waves are plane waves. If the signals are narrow-band, the array measurement model can be neatly formulated as [1]

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t) + \boldsymbol{\epsilon}(t), \quad t = 1, \dots, N, \quad (1.2)$$

where $\mathbf{y}(t)$ is a $M \times 1$ vector containing the array output at the t -th snapshot,

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_L)] \quad (1.3)$$

is the array steering matrix, $\mathbf{a}(\theta_l)$ is the $L \times 1$ steering vector corresponding to the l -th source, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_L]^T$ with θ_l the DOA of the l -th source, $\{\cdot\}^T$ denotes the matrix transpose, $\mathbf{x}(t)$ is the $L \times 1$ vector of signal values at the t -th snapshot, $\boldsymbol{\epsilon}(t)$ is an $M \times 1$ vector of noise values on the array sensors, and N is the total number of temporal measurements. The steering matrix $\mathbf{A}(\boldsymbol{\theta})$ is determined by the array geometry, signal carrier frequencies, and DOA (see [1] for the formulation of steering matrix). But since both the array geometry and carrier frequencies are known, the steering matrix becomes only a function of DOA. The aim of narrow-band DOA finding is to estimate $\boldsymbol{\theta}$ from the noise corrupted array output $\mathbf{y}(t)$, $t = 1, \dots, N$.

If the narrow-band condition in (1.1) is not satisfied, the signals are considered as wide-band, and the measurement model in (1.2) cannot be applied. In this case, we usually first decompose the wide frequency band into a set of narrow sub-bands [1], in

each of which the narrow-band condition is satisfied such that the model in (1.2) can be applied. For instance, if we decompose the wide frequency band into K sub-bands, in each of which (1.2) holds. Then the wide-band measurement model can be written as

$$\mathbf{y}_k(t) = \mathbf{A}_k(\boldsymbol{\theta})\mathbf{x}_k(t) + \boldsymbol{\epsilon}_k(t), \quad t = 1, \dots, N, \quad k = 1, \dots, K, \quad (1.4)$$

where $\mathbf{y}_k(t)$ is the $M \times 1$ measurement vector at the t -th snapshot from the k -th sub-band,

$$\mathbf{A}_k(\boldsymbol{\theta}) = [\mathbf{a}_k(\theta_1), \dots, \mathbf{a}_k(\theta_L)] \quad (1.5)$$

is the array steering matrix for the k -th sub-band, $\mathbf{x}_k(t)$ is the $L \times 1$ signal value vector from the k -th sub-band, $\boldsymbol{\epsilon}_k(t)$ is the $M \times 1$ noise vector from the k -th sub-band. Note that the DOA vector $\boldsymbol{\theta}$ is identical for all sub-bands but the carrier frequencies differ for different sub-bands. So the steering matrices from different sub-bands are different.

1.2 Maximum Likelihood Estimation

Typical DOA estimation methods include beamforming techniques [4], subspace-based methods such as MUSIC [6] and ESPRIT [2], and maximum likelihood (ML) methods (see [7] for examples). Among varieties of DOA estimation methods, the ML methods are often able to provide better performance than the others not only due to their asymptotic performance usually achieving the Cramer-Rao bound (CRB), which is a lower bound on estimation errors, but also because they can take advantage of better signal or noise models to provide better DOA estimation performance (see [8] for an example).

The ML methods aim at finding the DOA estimates by maximizing the log-likelihood (LL) functions over the unknown parameters including DOAs and unknown signal and noise parameters. Different signal or noise models may result in different LL functions and therefore different ML estimators. For example, the deterministic (conditional) signal model in [8] considers the signal values at all snapshots as deterministic unknown parameters and assumes the noise to be spatially and temporally white. Under this assumption, the unknown parameters in (1.2) include the DOA vector $\boldsymbol{\theta}$, the signal values $\boldsymbol{x}(t)$, $t = 1, \dots, N$, and the noise power σ^2 . By omitting constant terms, the LL function can be written as

$$L(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Q}) = -N \log |\mathbf{Q}| - \text{trace} \{ \mathbf{Q}^{-1} \mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \}, \quad (1.6)$$

where

$$\mathbf{X} = [\boldsymbol{x}(1), \dots, \boldsymbol{x}(N)], \quad (1.7)$$

$$\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) = [\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{X}][\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{X}]^H, \quad (1.8)$$

$$\mathbf{Y} = [\boldsymbol{y}(1), \dots, \boldsymbol{y}(N)], \quad (1.9)$$

$\mathbf{Q} = \sigma^2 \mathbf{I}_M$, \mathbf{I}_M is an $M \times M$ identity matrix, and $|\cdot|$ and $\text{trace}\{\cdot\}$ denote the determinant and the trace of a matrix, respectively.

Also in [8], the stochastic (unconditional) signal model considers the noise as spatially and temporally white as well but assumes the signal values are temporally independent and follow a zero-mean Gaussian distribution with unknown correlation matrix \mathbf{P} at each snapshot. According to this model and after omitting constant terms, the

LL function can be formulated as

$$L(\boldsymbol{\theta}, \mathbf{P}, \mathbf{Q}) = -\log |\mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q}| - \text{trace}\left\{[\mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q}]^{-1}\hat{\mathbf{R}}_{yy}\right\}, \quad (1.10)$$

where

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H. \quad (1.11)$$

The unknown parameters in (1.10) are $\boldsymbol{\theta}$, \mathbf{P} , and σ^2 .

We can see that the LL functions in (1.6) and (1.10) have different formulations and unknown parameters. Maximizing (1.6) and (1.10) with respect to their unknown parameters, respectively, will definitely results in different ML DOA estimators. It has been shown in [8] that the ML estimator based on the stochastic signal model provides better performance than the ML estimator based on the deterministic signal model for DOA estimation of zero-mean Gaussian signals. Therefore, to improve the performance of ML DOA estimation, we ought to design signal and noise models to be as accurate as possible.

1.3 Our Contributions

In this research, we develop new signal models, ML estimators, and performance analysis results for some underwater DOA estimation problems. We summarize our contributions as follows.

ML DOA Estimation in Spatially Colored Noise Using Sparse Arrays

Sparse sensor arrays have been explored as an effective solution to DOA estimation in spatially colored noise, which is quite common in underwater scenarios. We consider the narrow-band DOA estimation in spatially colored noise using sparse sensor arrays and develop new ML DOA estimators under the assumptions of zero-mean and non-zero-mean Gaussian signals based on an Expectation-Maximization (EM) framework. For the DOA finding of non-zero-mean Gaussian signals, we compute the CRB as well as the asymptotic error covariance matrix of the ML estimator that improperly assumes zero-mean Gaussian signals. We provide both analytical and numerical comparisons for the existing deterministic and the proposed ML estimators. The results show that the proposed estimators provide better accuracy than the existing deterministic estimator, and that the non-zero means in the signals improve the accuracy of DOA estimation.

DOA Finding for Hydroacoustic Signals From Marine Vessels

The hydroacoustic signals from marine vessels are known to consist of two parts: the noise-like part with continuous spectra and the sinusoidal part with discrete frequencies, which can be exploited to improve the DOA estimation accuracy. We consider the DOA estimation of hydroacoustic signals from marine vessel sources by modeling such signals as the mixture of deterministic sinusoidal signals and stochastic Gaussian signals, and derive the ML DOA estimator. We compute the asymptotic error covariance matrix of the proposed ML estimator, as well as that of the typical ML estimator

assuming zero-mean Gaussian signals, for DOA estimation of such signals. Our analytical comparisons and numerical examples show that compared with the typical ML estimator, the proposed ML estimator enhances the DOA estimation accuracy for the hydroacoustic signals from marine vessels.

A Barankin-Type Bound

Identification of the signal-to-noise ratio (SNR) threshold region, below which the accuracy of the ML estimation degrades rapidly, has important applications in the DOA estimation practice. The Barankin bound is a useful tool in estimation problems for predicting this threshold region of SNR. We derive a Barankin-type bound on the mean-square error (MSE) in estimating the DOAs of far-field sources using acoustic sensor arrays. We consider both narrow-band and wide-band deterministic signals, and scalar or vector sensors. Our results provide an approximation to the threshold of the SNR below which the ML estimation performance degrades rapidly. For narrow-band DOA estimation using uniform linear acoustic vector-sensor arrays, we show that this threshold increases with the inter-sensor distance. As a result, for medium SNR values, the performance does not necessarily improve with the inter-sensor distance.

1.4 Outline of the Dissertation

The rest of the dissertation is organized as follows. In Chapter 2, we present the results for the ML DOA estimation in spatially colored noise using sparse arrays. In Chapter 3, we develop the narrow-band DOA estimation models and results for hydroacoustic signals from marine vessels. Chapters 4 and 5 generalize the narrow-band results in

Chapter 3 to the wide-band case. In Chapter 6, we derive a Barankin-type bound on DOA estimation. At last, in Chapter 7, we summarize our contributions and discuss possible topics for future work.

Chapter 2

Maximum Likelihood

Direction-of-Arrival Estimation in

Spatially Colored Noise Using

Sparse Arrays¹

Spatially colored noise is quite common on sensor arrays in underwater direction-of-arrival (DOA) estimation scenarios. In this chapter, we consider the problem of maximum likelihood (ML) DOA estimation of narrow-band signals in spatially colored noise using sparse sensor arrays, which consist of widely separated sub-arrays such that the unknown spatially colored noise field is uncorrelated between different sub-arrays. We develop ML DOA estimators under the assumptions of zero-mean and non-zero-mean Gaussian signals based on an Expectation-Maximization (EM) framework. For DOA estimation of non-zero-mean Gaussian signals, we derive the Cramer-Rao bound (CRB) as well as the asymptotic error covariance matrix of the

¹Based on T. Li and A. Nehorai, "Maximum Likelihood Direction Finding in Spatially Colored Noise Fields Using Sparse Sensor Arrays," *IEEE Trans. Signal Process.*, vol. 59, pp. 1048-1062, Mar. 2011. ©[2011] IEEE.

ML estimator that improperly assumes zero-mean Gaussian signals. We provide analytical and numerical performance comparisons for the existing deterministic and the proposed ML estimators. The results show that the proposed estimators normally provide better accuracy than the existing deterministic estimator, and that the non-zero means in the signals improve the accuracy of DOA estimation.

2.1 Introduction

Array processing for DOA estimation has been a topic of intensive research interest during the past two decades. Many proposed estimators assume spatially white noise (see [7]-[11] for examples) such that the array noise covariance matrix is proportional to an identity matrix. However, this assumption is not realistic in many practical applications [12]-[18] where the noise fields are spatially colored. The spatial correlation or nonuniformity in the colored noise may significantly degrade the performance of the estimators assuming spatially white noise [19], [20]. In these applications, it is beneficial to take the spatial color of the noise into account to improve the resolution of DOA estimation.

Unfortunately, the problem of DOA estimation under spatially colored noise is not solvable unless special constraints are imposed on signals or noise. For instance, in [14], the noise field is assumed to satisfy a spatially autoregressive model. In [16], the signals are required to be partially known as a linear combination of a set of basis functions. The estimator proposed in [17] requires the temporal correlation length of the signals to be larger than that of the noise. However, these assumptions do not always hold in practice, and the performances of these estimators may deteriorate when the required constraints are not satisfied.

To avoid the constraints on signals and noise, sparse arrays consisting of separated sub-arrays were exploited for DOA estimation in spatially colored noise [21]-[24]. This technique sets the sub-arrays to be well separated such that noise is uncorrelated between different sub-arrays. As a result, the array noise covariance matrix presents a block-diagonal structure, which guarantees the identifiability of DOA information [24]. The early works on this topic [21]-[23] explored estimation methods using two separated sub-arrays. Recently, a deterministic ML estimator was proposed for the case of multiple sub-arrays [24].

In this chapter, under the assumption of Gaussian signals, we develop ML estimators based on an Expectation-Maximization (EM) framework for narrow-band DOA estimation in spatially colored noise using sparse arrays consisting of multiple sub-arrays. To the best of our knowledge, no similar estimator has been proposed so far. Many existing ML estimators assume the means of the Gaussian signals are zero (see [8]-[11] for examples). In this chapter, we consider both zero-mean and non-zero-mean Gaussian signals. The ML DOA estimation of non-zero mean signals under spatially colored noise was addressed in [14], [25]. The estimator in [14] is developed based on an autoregressive noise model, and is not a rigorous ML DOA estimator. The work in [25] shows that the non-zero mean component can be used to extract DOA information under spatially colored noise without any constraint needed on the sensor arrays, signals or the noise field. Some applications with non-zero mean signals were also given in [25], such as short-term MEG and EEG [26], [27], communication using amplitude-modulated or frequency-shift keying (with large frequency deviation) signals [28], [29], and underground source localization using gradiometer arrays [30]. Non-zero mean signals can also be used to describe acoustic waves from ships and submarines, which normally consist of sinusoidal components and noise with continuous

spectrums [31]. Since the frequency of a sinusoidal component can often be estimated accurately (see [32]-[36] for examples), we can obtain a non-zero-mean complex amplitude from a narrow-band acoustic wave by focusing the carrier frequency on the frequency of its sinusoidal component.

The ML estimator in [25] uses only the non-zero mean component in signals for DOA estimation. Our proposed ML estimator takes advantage of the block-diagonal structure of the noise covariance matrix and makes use of the total signal power for DOA estimation. We present relevant performance analysis results and give both analytical and numerical comparisons for estimators based on different signal models. We show that with the same correlation matrices of signals and noise, the non-zero-mean signals improve the accuracy of DOA estimation compared with the zero-mean ones.

The remainder of this chapter is organized as follows. We present the measurement models in Section 2.2, and derive our EM-based ML estimators in Section 2.3. Section 2.4 gives the results of analytical performance analysis. Numerical examples appear in Section 2.5. We give our conclusions in Section 2.6.

2.2 Measurement Models

In this section, we give the narrow-band measurement models for DOA estimation using sparse arrays under spatially colored noise.

Consider narrow-band signals from L distant sources impinging on a sparse array composed of K separated sub-arrays, the k -th of which consists of M_k sensors. Let $M = \sum_{k=1}^K M_k$ be the total number of sensors in the array. For one-dimensional (1D)

DOA estimation, the array output can be written as

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t) + \mathbf{e}(t), \quad t = 1, \dots, N, \quad (2.1)$$

where $\mathbf{y}(t)$ is the $M \times 1$ measurement vector at the t -th snapshot,

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_L)] \quad (2.2)$$

is the array steering matrix, $\mathbf{a}(\theta_l)$ is the steering vector corresponding to the l -th source, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_L]^T$ is the vector containing the DOAs of all sources, $\{\cdot\}^T$ denotes the matrix transpose, θ_l is the DOA of the l -th source, $\mathbf{x}(t)$ is the $L \times 1$ vector of all signal values at the t -th snapshot, N is the total number of temporal measurements, and $\mathbf{e}(t)$ is the $M \times 1$ noise vector following the zero-mean circular complex Gaussian distribution with covariance matrix

$$\mathbf{Q} = \text{blkdiag}\{\mathbf{Q}_1, \dots, \mathbf{Q}_K\}, \quad (2.3)$$

where $\text{blkdiag}\{\cdot\}$ denotes the block-diagonal matrix operator, and \mathbf{Q}_k is the $M_k \times M_k$ noise covariance matrix on the k -th sub-array. We assume the noise $\mathbf{e}(t)$ is temporally white, wide sense stationary, and uncorrelated with the signals.

For 2D DOA estimation problems, we have

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\boldsymbol{\theta}_1), \dots, \mathbf{a}(\boldsymbol{\theta}_L)], \quad (2.4)$$

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \dots, \boldsymbol{\theta}_L^T]^T, \quad (2.5)$$

and $\boldsymbol{\theta}_l = [\phi_l, \psi_l]^T$ in (2.1), where ϕ_l and ψ_l are the elevation and azimuth angles of the l -th source, respectively.

We now present the three types of signal models that will be considered in this chapter.

The deterministic signal model [8], [24] considers the signal values at all snapshots as deterministic unknown parameters. Under the assumption of deterministic signals, the unknown parameters in (2.1) consist of the DOA vector $\boldsymbol{\theta}$, the signal parameters $\mathbf{x}(t)$, $t = 1, \dots, T$, and the noise covariance matrix \mathbf{Q} .

In contrast, stochastic signal models consider $\mathbf{x}(t)$ to be a random process generated from a specific probability density function (pdf), which is normally assumed to be Gaussian. In this chapter, we consider both zero-mean and non-zero-mean circular complex Gaussian signals.

For zero-mean complex Gaussian signals, we assume

$$\mathbb{E} \{ \mathbf{x}(t) \mathbf{x}(s)^H \} = \mathbf{P} \delta_{t,s}, \quad (2.6)$$

where $\mathbb{E} \{ \cdot \}$ denotes the expectation operator, $\{ \cdot \}^H$ denotes the conjugate transpose, \mathbf{P} is the signal covariance matrix, and $\delta_{t,s}$ is the Kronecker delta function. Under the assumption of zero-mean Gaussian signals, the unknown parameters in (2.1) consist of the DOA vector $\boldsymbol{\theta}$, the signal covariance matrix \mathbf{P} , and the noise covariance matrix \mathbf{Q} .

For non-zero-mean complex Gaussian signals, we assume

$$\mathbb{E} \{ \mathbf{x}(t) \} = \mathbf{b}, \quad (2.7)$$

$$\mathbb{E} \{ [\mathbf{x}(t) - \mathbf{b}] [\mathbf{x}(s) - \mathbf{b}]^H \} = \mathbf{P} \delta_{t,s}, \quad (2.8)$$

where \mathbf{b} is the signal mean. Under the assumption of non-zero-mean Gaussian signals, the unknown parameters consist of the DOA vector $\boldsymbol{\theta}$, the signal mean \mathbf{b} , the signal covariance matrix \mathbf{P} , and the noise covariance matrix \mathbf{Q} .

Stochastic signals have been modeled with zero-mean Gaussian distributions in most existing work (see [8], [10], [20] for examples). Non-zero-mean Gaussian signals can be considered as mixtures of zero-mean Gaussian signals with deterministic unknown constants. In the remainder of this chapter, for simplicity of notation and presentation, we use “stochastic signals” to represent zero-mean Gaussian signals and use “mixed signals” to represent non-zero-mean Gaussian signals. Similarly, we use “stochastic” and “mixed” estimators to represent the estimators developed under the assumptions of zero-mean and non-zero-mean Gaussian signals, respectively.

2.3 Maximum Likelihood Estimation

In this section, we present our stochastic and mixed ML estimators for DOA finding under spatially colored noise using sparse sensor arrays. For convenience of comparison and further analysis, we first summarize the results in [24] for the deterministic ML DOA estimator.

2.3.1 Deterministic ML Estimator

After neglecting constant terms, the log-likelihood function based on the deterministic signal model can be written as

$$L(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Q}) = -N \log |\mathbf{Q}| - \text{trace} \{ \mathbf{Q}^{-1} \mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \}, \quad (2.9)$$

where

$$\mathbf{X} = [\mathbf{x}(1), \dots, \mathbf{x}(N)], \quad (2.10)$$

$$\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) = [\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{X}][\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{X}]^H, \quad (2.11)$$

$$\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(N)], \quad (2.12)$$

and $|\cdot|$ and $\text{trace}\{\cdot\}$ denote the determinant and the trace of a matrix, respectively.

By fixing $\boldsymbol{\theta}$ and \mathbf{X} , the ML estimate of \mathbf{Q} can be obtained as

$$\hat{\mathbf{Q}}(\boldsymbol{\theta}, \mathbf{X}) = \frac{1}{N} \mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E}, \quad (2.13)$$

where \odot denotes the Hadamard product,

$$\mathbf{E} = \text{blkdiag}\{\mathbf{E}_1, \dots, \mathbf{E}_K\}, \quad (2.14)$$

and \mathbf{E}_k is an $M_k \times M_k$ matrix with all entries equal to one. By inserting (2.13) into (2.9), the log-likelihood function can be simplified into

$$L(\boldsymbol{\theta}, \mathbf{X}) = -\log \left| \frac{1}{N} \mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E} \right|. \quad (2.15)$$

Similarly, by fixing $\boldsymbol{\theta}$ and \mathbf{Q} , the ML estimate of \mathbf{X} can be expressed as

$$\hat{\mathbf{X}}(\boldsymbol{\theta}, \mathbf{Q}) = \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta}) \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1} \tilde{\mathbf{A}}^H(\boldsymbol{\theta}) \tilde{\mathbf{Y}}, \quad (2.16)$$

where $\tilde{\mathbf{A}}(\boldsymbol{\theta}) = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A}(\boldsymbol{\theta})$ and $\tilde{\mathbf{Y}} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{Y}$. By substituting (2.16) into (2.15), the log-likelihood function can be rewritten as

$$L(\boldsymbol{\theta}, \mathbf{Q}) = -\log \left| \left[\mathbf{Q}^{\frac{1}{2}} \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \hat{\mathbf{R}}_{yy} \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \mathbf{Q}^{\frac{1}{2}} \right] \odot \mathbf{E} \right|, \quad (2.17)$$

where

$$\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) = \mathbf{I} - \tilde{\mathbf{A}}(\boldsymbol{\theta}) \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta}) \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1} \tilde{\mathbf{A}}^H(\boldsymbol{\theta}), \quad (2.18)$$

\mathbf{I} is the identity matrix, and

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N} \mathbf{Q}^{-\frac{1}{2}} \mathbf{Y} \mathbf{Y}^H \mathbf{Q}^{-\frac{1}{2}}. \quad (2.19)$$

The deterministic ML DOA estimator proposed in [24] is implemented in an iterative manner as follows.

Algorithm 1: Deterministic ML DOA Estimator

Step 1: Initialize $\hat{\mathbf{Q}}$ at $\hat{\mathbf{Q}} = \mathbf{I}$, an identity matrix.

Step 2: Find the DOA estimate as

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \log \left| \left[\hat{\mathbf{Q}}^{\frac{1}{2}} \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \hat{\mathbf{R}}_{yy} \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \hat{\mathbf{Q}}^{\frac{1}{2}} \right] \odot \mathbf{E} \right|. \quad (2.20)$$

Step 3: Compute $\hat{\mathbf{X}}$ using equation (2.16) with the $\hat{\boldsymbol{\theta}}$ value obtained in step 2.

Refine $\hat{\mathbf{Q}}$ using equation (2.13) and the obtained values of $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{X}}$.

Iterate Steps 2 and 3 several times to obtain the final ML DOA estimate.

We now modify this estimator for the special case when all blocks in \mathbf{Q} are equal.

Suppose $\mathbf{Q}_1 = \mathbf{Q}_2 = \dots = \mathbf{Q}_K = \mathbf{Q}_0$. Then we have the log-likelihood function

$$L(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Q}_0) = -NK \log |\mathbf{Q}_0| - \text{trace} \left\{ \mathbf{Q}_0^{-1} \mathbf{C}_0(\boldsymbol{\theta}, \mathbf{X}) \right\}, \quad (2.21)$$

where

$$\mathbf{C}_0(\boldsymbol{\theta}, \mathbf{X}) = \sum_{k=1}^K \mathbf{C}_k(\boldsymbol{\theta}, \mathbf{X}) = \mathbf{F}[\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E}]\mathbf{F}^T, \quad (2.22)$$

$$\mathbf{F} = \mathbf{1}_K^T \otimes \mathbf{I}_{M_0}, \quad (2.23)$$

$\mathbf{C}_k(\boldsymbol{\theta}, \mathbf{X})$ is the k -th diagonal block of $\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E}$, $\mathbf{1}_K$ is a $K \times 1$ vector with all entries 1, M_0 is the dimension of \mathbf{Q}_0 , \mathbf{I}_{M_0} is the $M_0 \times M_0$ identity matrix, and \otimes denotes the Kronecker product.

We use the following lemma [37] to obtain the ML estimate for \mathbf{Q}_0 directly.

Lemma 1. *Let \mathbf{C} be an $M \times M$ positive definite matrix. Then, for $a > 0$ and $b > 0$,*

$$|\mathbf{Q}|^{-b} \exp\{-a \text{trace}\{\mathbf{Q}^{-1}\mathbf{C}\}\} \leq |a\mathbf{C}/b|^{-b} \exp\{-Mb\} \quad (2.24)$$

for all $M \times M$ positive definite matrices \mathbf{Q} . The equality holds if and only if $\mathbf{Q} = a\mathbf{C}/b$.

According to equation (2.21) and Lemma 1, we obtain the ML estimate for \mathbf{Q}_0 as

$$\hat{\mathbf{Q}}_0(\boldsymbol{\theta}, \mathbf{X}) = \frac{1}{NK} \mathbf{C}_0(\boldsymbol{\theta}, \mathbf{X}) = \frac{1}{NK} \mathbf{F}[\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E}]\mathbf{F}^T. \quad (2.25)$$

Inserting (2.25) into (2.21) and omitting constant terms, we rewrite the log-likelihood function as

$$L(\boldsymbol{\theta}, \mathbf{X}) = -\log \left| \frac{1}{NK} \mathbf{F}[\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E}]\mathbf{F}^T \right|. \quad (2.26)$$

When $\boldsymbol{\theta}$ and \mathbf{Q} are fixed, the ML estimate of \mathbf{X} is the same as that in equation (2.16). Substituting (2.16) into (2.26), we have

$$L(\boldsymbol{\theta}, \mathbf{Q}_0) = -\log \left| \frac{1}{K} \mathbf{F} \left[\left(\mathbf{Q}^{\frac{1}{2}} \boldsymbol{\Pi}_{\hat{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \hat{\mathbf{R}}_{yy} \boldsymbol{\Pi}_{\hat{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \mathbf{Q}^{\frac{1}{2}} \right) \odot \mathbf{E} \right] \mathbf{F}^T \right|. \quad (2.27)$$

Therefore, when all the blocks of \mathbf{Q} are equal, the iterative ML DOA estimation can be implemented following steps similar to those in Algorithm 1, except that the estimate of $\boldsymbol{\theta}$ in Step 2 should be replaced with

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \log \left| \frac{1}{K} \mathbf{F} \left[\left(\hat{\mathbf{Q}}^{\frac{1}{2}} \boldsymbol{\Pi}_{\hat{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \hat{\mathbf{R}}_{yy} \boldsymbol{\Pi}_{\hat{\mathbf{A}}}^{\perp}(\boldsymbol{\theta}) \hat{\mathbf{Q}}^{\frac{1}{2}} \right) \odot \mathbf{E} \right] \mathbf{F}^T \right|, \quad (2.28)$$

and the estimate $\hat{\mathbf{Q}}$ in Step 3 should be replaced with $\hat{\mathbf{Q}}_0$ using equation (2.25).

2.3.2 Stochastic ML Estimator

For stochastic signals, the log-likelihood function can be formulated as follows after omitting constant terms.

$$L(\boldsymbol{\theta}, \mathbf{P}, \mathbf{Q}) = -\log \left| \mathbf{A}(\boldsymbol{\theta}) \mathbf{P} \mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q} \right| - \text{trace} \left\{ \left[\mathbf{A}(\boldsymbol{\theta}) \mathbf{P} \mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q} \right]^{-1} \hat{\mathbf{R}}_{yy} \right\}, \quad (2.29)$$

where

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N} \mathbf{Y} \mathbf{Y}^H. \quad (2.30)$$

Closed-form ML estimates for \mathbf{P} and \mathbf{Q} as functions of $\boldsymbol{\theta}$ are generally not available except for some special cases such as $\mathbf{Q} = \sigma^2 \mathbf{I}$. Herein, we present an iterative ML DOA estimation procedure based on the EM framework [38]-[41].

Let $\boldsymbol{\alpha} = \{\boldsymbol{\theta}, \mathbf{P}, \mathbf{Q}\}$ denote all the unknown parameters. Suppose our current parameter estimates are $\hat{\boldsymbol{\alpha}} = \{\hat{\boldsymbol{\theta}}, \hat{\mathbf{P}}, \hat{\mathbf{Q}}\}$, and let $\{\mathbf{Y}, \mathbf{X}\}$ and \mathbf{Y} be the complete and

incomplete data, respectively. According to the EM algorithm, the improved estimates can be found in the next step as

$$\hat{\boldsymbol{\alpha}}^{\text{new}} = \arg \max_{\boldsymbol{\alpha}} \mathbb{E}_{\mathbf{X}|\mathbf{Y}, \hat{\boldsymbol{\alpha}}} \{ \log f(\mathbf{Y}, \mathbf{X}; \boldsymbol{\alpha}) \}, \quad (2.31)$$

where $\mathbb{E}_{\mathbf{X}|\mathbf{Y}, \hat{\boldsymbol{\alpha}}} \{ \cdot \}$ is the expectation under the conditional distribution $f(\mathbf{X}|\mathbf{Y}, \hat{\boldsymbol{\alpha}})$.

The joint pdf $f(\mathbf{Y}, \mathbf{X}; \boldsymbol{\alpha})$ parameterized by $\boldsymbol{\alpha}$ can be written as

$$\begin{aligned} f(\mathbf{Y}, \mathbf{X}; \boldsymbol{\alpha}) &= f(\mathbf{Y}|\mathbf{X}, \boldsymbol{\theta}, \mathbf{Q})f(\mathbf{X}; \mathbf{P}) = \prod_{t=1}^N \frac{1}{|\pi \mathbf{Q}|} \exp \left\{ - [\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t)]^H \right. \\ &\quad \left. \times \mathbf{Q}^{-1} [\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t)] \right\} \frac{1}{|\pi \mathbf{P}|} \exp \left\{ - \mathbf{x}^H(t) \mathbf{P}^{-1} \mathbf{x}(t) \right\}. \end{aligned} \quad (2.32)$$

Note that in (2.32) we assume the signals are not fully coherent such that $\mathbf{P} > 0$ (i.e., \mathbf{P} is positive definite as a matrix), which is true in most practical applications [21].

Neglecting constant terms, we obtain

$$\begin{aligned} \log f(\mathbf{Y}, \mathbf{X}; \boldsymbol{\alpha}) &\propto -N \log |\mathbf{Q}| - N \log |\mathbf{P}| - \sum_{t=1}^N \left\{ \mathbf{y}^H(t) \mathbf{Q}^{-1} \mathbf{y}(t) \right. \\ &\quad \left. - \mathbf{y}^H(t) \mathbf{Q}^{-1} \mathbf{A}(\boldsymbol{\theta}) \mathbf{x}(t) - \mathbf{x}^H(t) \mathbf{A}^H(\boldsymbol{\theta}) \mathbf{Q}^{-1} \mathbf{y}(t) \right. \\ &\quad \left. + \mathbf{x}^H(t) [\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{Q}^{-1} \mathbf{A}(\boldsymbol{\theta}) + \mathbf{P}^{-1}] \mathbf{x}(t) \right\} \\ &= -N \log |\mathbf{Q}| - N \log |\mathbf{P}| - \text{trace} \left\{ \mathbf{Q}^{-1} \mathbf{Y} \mathbf{Y}^H \right. \\ &\quad \left. - \mathbf{Q}^{-1} \mathbf{A}(\boldsymbol{\theta}) \mathbf{X} \mathbf{Y}^H - \mathbf{Q}^{-1} \mathbf{Y} \mathbf{X}^H \mathbf{A}^H(\boldsymbol{\theta}) \right. \\ &\quad \left. + [\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{Q}^{-1} \mathbf{A}(\boldsymbol{\theta}) + \mathbf{P}^{-1}] \mathbf{X} \mathbf{X}^H \right\} \\ &= -N \log |\mathbf{P}| - \text{trace} \left\{ \mathbf{P}^{-1} \mathbf{X} \mathbf{X}^H \right\} - N \log |\mathbf{Q}| \\ &\quad - \text{trace} \left\{ \mathbf{Q}^{-1} \left[\mathbf{Y} \mathbf{Y}^H - \mathbf{A}(\boldsymbol{\theta}) \mathbf{X} \mathbf{Y}^H - \mathbf{Y} \mathbf{X}^H \mathbf{A}^H(\boldsymbol{\theta}) \right. \right. \\ &\quad \left. \left. + \mathbf{A}(\boldsymbol{\theta}) \mathbf{X} \mathbf{X}^H \mathbf{A}^H(\boldsymbol{\theta}) \right] \right\}. \end{aligned} \quad (2.33)$$

From the joint complex Gaussian distribution of $\mathbf{y}(t)$ and $\mathbf{x}(t)$, we can obtain the posterior distribution [42]

$$\mathbf{x}(t)|\mathbf{y}(t), \hat{\boldsymbol{\alpha}} \sim \mathcal{CN}(\hat{\mathbf{g}}(t), \hat{\boldsymbol{\Phi}}), \quad (2.34)$$

where $\mathcal{CN}(\cdot)$ denotes the complex Gaussian distribution,

$$\hat{\boldsymbol{\Phi}} = \left[\mathbf{A}^H(\hat{\boldsymbol{\theta}}) \hat{\mathbf{Q}}^{-1} \mathbf{A}(\hat{\boldsymbol{\theta}}) + \hat{\mathbf{P}}^{-1} \right]^{-1}, \quad (2.35)$$

$$\hat{\mathbf{g}}(t) = \hat{\mathbf{P}} \mathbf{A}^H(\hat{\boldsymbol{\theta}}) \left[\mathbf{A}(\hat{\boldsymbol{\theta}}) \hat{\mathbf{P}} \mathbf{A}^H(\hat{\boldsymbol{\theta}}) + \hat{\mathbf{Q}} \right]^{-1} \mathbf{y}(t). \quad (2.36)$$

Let $\hat{\mathbf{G}} = [\hat{\mathbf{g}}(1), \dots, \hat{\mathbf{g}}(N)]$, then we have

$$\mathbb{E}_{\mathbf{X}|\mathbf{Y}, \hat{\boldsymbol{\alpha}}} \{ \mathbf{X} \} = \hat{\mathbf{G}}, \quad (2.37)$$

$$\mathbb{E}_{\mathbf{X}|\mathbf{Y}, \hat{\boldsymbol{\alpha}}} \{ \mathbf{X} \mathbf{X}^H \} = N \hat{\boldsymbol{\Phi}} + \hat{\mathbf{G}} \hat{\mathbf{G}}^H. \quad (2.38)$$

Using the results from (2.33), (2.37), and (2.38), we have

$$\begin{aligned} & \mathbb{E}_{\mathbf{X}|\mathbf{Y}, \hat{\boldsymbol{\alpha}}} \{ \log f(\mathbf{Y}, \mathbf{X}; \boldsymbol{\alpha}) \} \\ &= -N \log |\mathbf{P}| - \text{trace} \left\{ \mathbf{P}^{-1} \left(N \hat{\boldsymbol{\Phi}} + \hat{\mathbf{G}} \hat{\mathbf{G}}^H \right) \right\} \\ & \quad - N \log |\mathbf{Q}| - \text{trace} \left\{ \mathbf{Q}^{-1} \left[\mathbf{Y} \mathbf{Y}^H - \mathbf{A}(\boldsymbol{\theta}) \hat{\mathbf{G}} \mathbf{Y}^H \right. \right. \\ & \quad \left. \left. - \mathbf{Y} \hat{\mathbf{G}}^H \mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{A}(\boldsymbol{\theta}) \left(N \hat{\boldsymbol{\Phi}} + \hat{\mathbf{G}} \hat{\mathbf{G}}^H \right) \mathbf{A}^H(\boldsymbol{\theta}) \right] \right\} \\ & \propto -\log |\mathbf{P}| - \text{trace} \left\{ \mathbf{P}^{-1} \left(\hat{\boldsymbol{\Phi}} + \hat{\mathbf{R}}_{\hat{\mathbf{g}}\hat{\mathbf{g}}} \right) \right\} \\ & \quad - \log |\mathbf{Q}| - \text{trace} \left\{ \mathbf{Q}^{-1} \boldsymbol{\Gamma}(\boldsymbol{\theta}) \right\} \end{aligned} \quad (2.39)$$

$$\begin{aligned} &= -\log |\mathbf{P}| - \text{trace} \left\{ \mathbf{P}^{-1} \left(\hat{\boldsymbol{\Phi}} + \hat{\mathbf{R}}_{\hat{\mathbf{g}}\hat{\mathbf{g}}} \right) \right\} \\ & \quad - \sum_{k=1}^K \left(\log |\mathbf{Q}_k| + \text{trace} \left\{ \mathbf{Q}_k^{-1} \boldsymbol{\Gamma}_k(\boldsymbol{\theta}) \right\} \right), \end{aligned} \quad (2.40)$$

where

$$\hat{\mathbf{R}}_{\hat{\mathbf{g}}\hat{\mathbf{g}}} = \frac{1}{N} \hat{\mathbf{G}} \hat{\mathbf{G}} = \frac{1}{N} \sum_{t=1}^N \hat{\mathbf{g}}(t) \hat{\mathbf{g}}^H(t), \quad (2.41)$$

$$\begin{aligned} \mathbf{\Gamma}(\boldsymbol{\theta}) &= \frac{1}{N} \left[\mathbf{Y} \mathbf{Y}^H - \mathbf{A}(\boldsymbol{\theta}) \hat{\mathbf{G}} \mathbf{Y}^H - \mathbf{Y} \hat{\mathbf{G}}^H \mathbf{A}^H(\boldsymbol{\theta}) \right. \\ &\quad \left. + \mathbf{A}(\boldsymbol{\theta}) \left(N \hat{\boldsymbol{\Phi}} + \hat{\mathbf{G}} \hat{\mathbf{G}} \right) \mathbf{A}^H(\boldsymbol{\theta}) \right] \\ &= \mathbf{A}(\boldsymbol{\theta}) \hat{\boldsymbol{\Phi}} \mathbf{A}^H(\boldsymbol{\theta}) + \frac{1}{N} \left(\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta}) \hat{\mathbf{G}} \right) \left(\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta}) \hat{\mathbf{G}} \right)^H, \end{aligned} \quad (2.42)$$

and $\mathbf{\Gamma}_k(\boldsymbol{\theta})$ is the k -th diagonal block of $\mathbf{\Gamma}(\boldsymbol{\theta}) \odot \mathbf{E}$.

According to Lemma 1, we have

$$\hat{\mathbf{P}}^{\text{new}} = \hat{\boldsymbol{\Phi}} + \hat{\mathbf{R}}_{\hat{\mathbf{g}}\hat{\mathbf{g}}}, \quad (2.43)$$

$$\hat{\mathbf{Q}}_k^{\text{new}} = \mathbf{\Gamma}_k(\boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}^{\text{new}}}, \quad (2.44)$$

$$\hat{\mathbf{Q}}^{\text{new}} = \mathbf{\Gamma}(\boldsymbol{\theta}) \odot \mathbf{E} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}^{\text{new}}}. \quad (2.45)$$

Inserting (2.43) and (2.45) into (2.39), we can obtain $\hat{\boldsymbol{\theta}}^{\text{new}}$ as

$$\hat{\boldsymbol{\theta}}^{\text{new}} = \arg \min_{\boldsymbol{\theta}} \log |\mathbf{\Gamma}(\boldsymbol{\theta}) \odot \mathbf{E}|. \quad (2.46)$$

Though we may implement the DOA estimation completely using the EM results in (2.43)-(2.46), we note that when \mathbf{Q} is fixed, the ML estimate for $\boldsymbol{\theta}$ is [10]

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \arg \min_{\boldsymbol{\theta}} \left\{ \log \left| \tilde{\mathbf{A}}(\boldsymbol{\theta}) \hat{\mathbf{P}}(\boldsymbol{\theta}) \tilde{\mathbf{A}}^H(\boldsymbol{\theta}) + \mathbf{I} \right| \right. \\ &\quad \left. + \text{trace} \left\{ \left[\tilde{\mathbf{A}}(\boldsymbol{\theta}) \hat{\mathbf{P}}(\boldsymbol{\theta}) \tilde{\mathbf{A}}^H(\boldsymbol{\theta}) + \mathbf{I} \right]^{-1} \hat{\mathbf{R}}_{\mathbf{y}\mathbf{y}} \right\} \right\}, \end{aligned} \quad (2.47)$$

where

$$\begin{aligned} \hat{\mathbf{P}}(\boldsymbol{\theta}) &= \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta}) \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1} \tilde{\mathbf{A}}^H(\boldsymbol{\theta}) \hat{\mathbf{R}}_{yy} \tilde{\mathbf{A}}(\boldsymbol{\theta}) \\ &\quad \times \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta}) \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1} - \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta}) \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1}. \end{aligned} \quad (2.48)$$

The ML estimate for \mathbf{P} can be obtained as $\hat{\mathbf{P}}(\hat{\boldsymbol{\theta}})$. Equations (2.47) and (2.48) provide the optimal results for $\boldsymbol{\theta}$ and \mathbf{P} when \mathbf{Q} is fixed. They also demonstrate that the EM updates $\hat{\boldsymbol{\theta}}^{\text{new}}$ and $\hat{\mathbf{P}}^{\text{new}}$ are not the best match for $\mathbf{Q} = \hat{\mathbf{Q}}^{\text{new}}$.

When $\boldsymbol{\theta}$ and \mathbf{P} are fixed, a closed-form ML estimate for \mathbf{Q} is normally not available [20]. However, we can update the estimate for \mathbf{Q} using the EM result. Assuming the existing parameter estimates are $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{P}}$, and $\hat{\mathbf{Q}}$, according to the EM result in (2.45), an improved estimate for \mathbf{Q} can be found as

$$\hat{\mathbf{Q}}^{\text{new}} = \Gamma(\hat{\boldsymbol{\theta}}) \odot \mathbf{E}. \quad (2.49)$$

We now consider the special case when $\mathbf{Q}_1 = \mathbf{Q}_2 = \dots = \mathbf{Q}_K = \mathbf{Q}_0$. For this special case, the results in (2.47) and (2.48) still hold when \mathbf{Q} is fixed. To obtain the EM update equation for \mathbf{Q}_0 , we rewrite (2.40) as

$$\begin{aligned} & \mathbb{E}_{\mathbf{X}|\mathbf{Y}, \boldsymbol{\alpha}} \{ \log f(\mathbf{Y}, \mathbf{X}; \boldsymbol{\alpha}) \} \\ & \propto -\log |\mathbf{P}| - \text{trace} \left\{ \mathbf{P}^{-1} \left(\hat{\boldsymbol{\Phi}} + \hat{\mathbf{R}}_{\hat{g}\hat{g}} \right) \right\} \\ & \quad - K \left(\log |\mathbf{Q}_0| + \text{trace} \left\{ \mathbf{Q}_0^{-1} \Gamma_0(\boldsymbol{\theta}) \right\} \right), \end{aligned} \quad (2.50)$$

where

$$\Gamma_0(\boldsymbol{\theta}) = \frac{1}{K} \sum_{k=1}^K \Gamma_k(\boldsymbol{\theta}) = \frac{1}{K} \mathbf{F} [\Gamma(\boldsymbol{\theta}) \odot \mathbf{E}] \mathbf{F}^T. \quad (2.51)$$

So if the existing parameter estimates are $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{P}}$, and $\hat{\mathbf{Q}}$, an improved estimate for \mathbf{Q}_0 is

$$\hat{\mathbf{Q}}_0^{\text{new}} = \frac{1}{K} \mathbf{F} [\boldsymbol{\Gamma}(\hat{\boldsymbol{\theta}}) \odot \mathbf{E}] \mathbf{F}^T. \quad (2.52)$$

Based on the results in (2.47)-(2.49) and (2.52), we propose our EM-based stochastic DOA estimator as follows.

Algorithm 2: EM-Based Stochastic ML DOA Estimator

Step 1: Initialize the parameter estimates at $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_{\text{init}}$, $\hat{\mathbf{P}} = \mathbf{P}_{\text{init}}$, and $\hat{\mathbf{Q}} = \mathbf{Q}_{\text{init}}$.

Step 2: Update $\hat{\mathbf{Q}}$ using equation (2.49).

Step 3: Fixing \mathbf{Q} at the $\hat{\mathbf{Q}}$ value obtained in Step 2, update $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{P}}$ using (2.47) and (2.48).

Iterate Steps 2 and 3 until convergence to obtain the final ML DOA estimate. Replace (2.49) in Step 2 with (2.52) for the special case $\mathbf{Q}_1 = \mathbf{Q}_2 = \dots = \mathbf{Q}_K = \mathbf{Q}_0$.

2.3.3 Mixed ML Estimator

Omitting constant terms, we have the log-likelihood function for mixed signals as

$$L(\boldsymbol{\theta}, \mathbf{b}, \mathbf{P}, \mathbf{Q}) = -\log |\mathbf{A}(\boldsymbol{\theta}) \mathbf{P} \mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q}| - \text{trace} \left\{ [\mathbf{A}(\boldsymbol{\theta}) \mathbf{P} \mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q}]^{-1} \hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta}, \mathbf{b}) \right\}, \quad (2.53)$$

where

$$\hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta}, \mathbf{b}) = \frac{1}{N} \sum_{t=1}^N [\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{b}][\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{b}]^H. \quad (2.54)$$

When \mathbf{Q} is fixed, we can obtain the closed-form ML estimate for \mathbf{b} as a function of $\boldsymbol{\theta}$ as

$$\hat{\mathbf{b}}(\boldsymbol{\theta}) = \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta})\tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1} \tilde{\mathbf{A}}^H(\boldsymbol{\theta})\bar{\mathbf{y}}, \quad (2.55)$$

where $\bar{\mathbf{y}} = \frac{1}{N} \sum_{t=1}^N \mathbf{y}(t)$. Substituting (2.55) into (2.53), we can see that the resulting log-likelihood function $L(\boldsymbol{\theta}, \hat{\mathbf{b}}(\boldsymbol{\theta}), \mathbf{P}, \mathbf{Q})$ is similar to the stochastic log-likelihood function in (2.29). The only difference is that the matrix $\hat{\mathbf{R}}_{\mathbf{y}\mathbf{y}}$ in (2.29) is replaced by $\hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta}) = \hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta}, \hat{\mathbf{b}}(\boldsymbol{\theta}))$ in $L(\boldsymbol{\theta}, \hat{\mathbf{b}}(\boldsymbol{\theta}), \mathbf{P}, \mathbf{Q})$. Consequently, the results for stochastic signals in (2.47) and (2.48) still hold for mixed signals by replacing $\hat{\mathbf{R}}_{\mathbf{y}\mathbf{y}}$ in them with $\hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta})$ [10], or equivalently, by replacing $\mathbf{y}(t)$ in them with $\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\hat{\mathbf{b}}(\boldsymbol{\theta})$. Therefore, for mixed signals, by fixing \mathbf{Q} and using equations (2.47), (2.48) and (2.55), we can obtain the ML estimate for $\boldsymbol{\theta}$ as

$$\begin{aligned} \hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \left\{ \log \left| \tilde{\mathbf{A}}(\boldsymbol{\theta})\hat{\mathbf{P}}(\boldsymbol{\theta})\tilde{\mathbf{A}}^H(\boldsymbol{\theta}) + \mathbf{I} \right| \right. \\ \left. + \text{trace} \left\{ \left[\tilde{\mathbf{A}}(\boldsymbol{\theta})\hat{\mathbf{P}}(\boldsymbol{\theta})\tilde{\mathbf{A}}^H(\boldsymbol{\theta}) + \mathbf{I} \right]^{-1} \hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta}) \right\} \right\}, \end{aligned} \quad (2.56)$$

where $\hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta}) = \mathbf{Q}^{-\frac{1}{2}}\hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta})\mathbf{Q}^{-\frac{1}{2}}$ and

$$\begin{aligned} \hat{\mathbf{P}}(\boldsymbol{\theta}) = \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta})\tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1} \tilde{\mathbf{A}}^H(\boldsymbol{\theta})\hat{\mathbf{C}}_{\mathbf{y}\mathbf{y}}(\boldsymbol{\theta})\tilde{\mathbf{A}}(\boldsymbol{\theta}) \\ \times \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta})\tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1} - \left[\tilde{\mathbf{A}}^H(\boldsymbol{\theta})\tilde{\mathbf{A}}(\boldsymbol{\theta}) \right]^{-1}. \end{aligned} \quad (2.57)$$

The ML estimates for \mathbf{b} and \mathbf{P} are $\hat{\mathbf{b}}(\hat{\boldsymbol{\theta}})$ and $\hat{\mathbf{P}}(\hat{\boldsymbol{\theta}})$, respectively. Similarly, if the existing estimates are $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{b}}$, $\hat{\mathbf{P}}$, and $\hat{\mathbf{Q}}$, then the update equations for \mathbf{Q} in (2.49) and

(2.52) can still be used for mixed signals by replacing $\mathbf{y}(t)$ with $\mathbf{y}(t) - \mathbf{A}(\hat{\boldsymbol{\theta}})\hat{\mathbf{b}}$. We thus propose the following algorithm for mixed ML DOA estimation.

Algorithm 3: EM-Based Mixed ML DOA Estimator

Step 1: Initialize the parameter estimates at $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_{\text{init}}$, $\hat{\mathbf{b}} = \mathbf{b}_{\text{init}}$, $\hat{\mathbf{P}} = \mathbf{P}_{\text{init}}$, and $\hat{\mathbf{Q}} = \mathbf{Q}_{\text{init}}$.

Step 2: Update $\hat{\mathbf{Q}}$ using equation (2.49), with $\mathbf{y}(t)$ replaced by $\mathbf{y}(t) - \mathbf{A}(\hat{\boldsymbol{\theta}})\hat{\mathbf{b}}$, $t = 1, \dots, N$.

Step 3: Fixing \mathbf{Q} at the $\hat{\mathbf{Q}}$ value obtained in Step 2, update $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{b}}$, and $\hat{\mathbf{P}}$ using (2.56), (2.55) and (2.57), respectively.

Iterate Steps 2 and 3 until convergence to obtain the final ML DOA estimate. Replace (2.49) in Step 2 with (2.52) for the special case $\mathbf{Q}_1 = \mathbf{Q}_2 = \dots = \mathbf{Q}_K = \mathbf{Q}_0$.

We now consider the mixed ML estimator for the special case of spatially white noise with $\mathbf{Q} = \sigma^2 \mathbf{I}$. When $\mathbf{Q} = \sigma^2 \mathbf{I}$, the ML estimate of \mathbf{b} becomes

$$\hat{\mathbf{b}}(\boldsymbol{\theta}) = [\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})]^{-1} \mathbf{A}^H(\boldsymbol{\theta})\bar{\mathbf{y}}. \quad (2.58)$$

Using (2.58) and the well-known results of DOA estimation for stochastic signals [8], [10] under spatially white noise, we obtain the ML estimate of $\boldsymbol{\theta}$ as

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \left\{ \log \left| \mathbf{A}(\boldsymbol{\theta})\hat{\mathbf{P}}(\boldsymbol{\theta})\mathbf{A}^H(\boldsymbol{\theta}) + \hat{\sigma}^2(\boldsymbol{\theta})\mathbf{I} \right| \right\}, \quad (2.59)$$

where

$$\hat{\sigma}^2(\boldsymbol{\theta}) = \frac{1}{M-L} \text{trace} \left\{ \left[\mathbf{I} - \mathbf{A}(\boldsymbol{\theta}) (\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta}))^{-1} \mathbf{A}^H(\boldsymbol{\theta}) \right]^{-1} \hat{\mathbf{C}}_{yy}(\boldsymbol{\theta}) \right\}, \quad (2.60)$$

$$\begin{aligned} \hat{\mathbf{P}}(\boldsymbol{\theta}) &= [\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta})]^{-1} \mathbf{A}^H(\boldsymbol{\theta}) \hat{\mathbf{C}}_{yy}(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta}) \\ &\times [\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta})]^{-1} - \hat{\sigma}^2(\boldsymbol{\theta}) [\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta})]^{-1}. \end{aligned} \quad (2.61)$$

2.4 Analytical Performance Analysis

In this section, we present analytical results on the performances of the deterministic, stochastic, and mixed ML DOA estimators. We also extend some well-known CRB and asymptotic error results for 1D DOA estimation to the 2D case. Our theorems and proportions derived in this section hold for both 1D and 2D DOA estimation under arbitrary proper noise covariance matrices unless the DOA dimension or noise covariance matrix structure is clearly specified for the theorem or proposition. For convenience of formulation, we define the following notations.

$$\mathbf{R} = \mathbf{A}(\boldsymbol{\theta}) \mathbf{P} \mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q}, \quad (2.62)$$

$$\mathbf{D} = \left[\frac{d\mathbf{a}(\theta_1)}{d\theta_1}, \frac{d\mathbf{a}(\theta_2)}{d\theta_2}, \dots, \frac{d\mathbf{a}(\theta_L)}{d\theta_L} \right], \quad (2.63)$$

$$\tilde{\mathbf{D}} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}, \quad (2.64)$$

$$\bar{\mathbf{D}} = \mathbf{R}^{-\frac{1}{2}} \mathbf{D}, \quad (2.65)$$

$$\mathbf{D}_2 = \left[\frac{\partial \mathbf{a}(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1^T}, \frac{\partial \mathbf{a}(\boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2^T}, \dots, \frac{\partial \mathbf{a}(\boldsymbol{\theta}_L)}{\partial \boldsymbol{\theta}_L^T} \right], \quad (2.66)$$

$$\tilde{\mathbf{D}}_2 = \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_2, \quad (2.67)$$

$$\bar{\mathbf{D}}_2 = \mathbf{R}^{-\frac{1}{2}} \mathbf{D}_2, \quad (2.68)$$

$$\bar{\mathbf{A}}(\boldsymbol{\theta}) = \mathbf{R}^{-\frac{1}{2}} \mathbf{A}(\boldsymbol{\theta}), \quad (2.69)$$

$$\boldsymbol{\Delta} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad (2.70)$$

$$\hat{\mathbf{R}}_{\mathbf{x}\mathbf{x}} = \frac{1}{N} \sum_{t=1}^N \mathbf{x}(t) \mathbf{x}^H(t), \quad (2.71)$$

$$\mathbf{R}_{\mathbf{x}\mathbf{x}} = \mathbb{E} \{ \mathbf{x}(t) \mathbf{x}(t)^H \}, \quad (2.72)$$

$$\tilde{\mathbf{R}} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{R} \mathbf{Q}^{-\frac{1}{2}}, \quad (2.73)$$

$$\mathbf{Q}'_k = \frac{d\mathbf{Q}}{d\sigma_k}, \quad k = 1, \dots, p, \quad (2.74)$$

$$\tilde{\mathbf{Q}}'_k = \mathbf{Q}^{-\frac{1}{2}} \mathbf{Q}'_k \mathbf{Q}^{-\frac{1}{2}}, \quad (2.75)$$

$$\mathbf{P}_2 = \mathbf{P} \otimes [1, 1], \quad (2.76)$$

$$\boldsymbol{\Lambda} = [\text{vec} \{ \tilde{\mathbf{Q}}'_1 \}, \dots, \text{vec} \{ \tilde{\mathbf{Q}}'_p \}], \quad (2.77)$$

$$\boldsymbol{\Xi} = [\text{vec} \{ \mathbf{e}_1 \mathbf{e}_1^T \}, \dots, \text{vec} \{ \mathbf{e}_{2L} \mathbf{e}_{2L}^T \}], \quad (2.78)$$

where $\sigma_1, \sigma_2, \dots, \sigma_p$ are the real unknown parameters from \mathbf{Q} , p is the number of real unknown parameters in \mathbf{Q} , $\text{vec}\{\cdot\}$ denotes the vectorization operator stacking all the columns of a matrix, one below another, into a vector, and \mathbf{e}_k is a $2L \times 1$ vector with the k -th element 1 and all the other elements 0. For simplicity, in the remainder of this chapter, we omit $\boldsymbol{\theta}$ and use \mathbf{A} to represent $\mathbf{A}(\boldsymbol{\theta})$. Also we let $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \dots, \sigma_p]^T$ be the real vector containing all the real unknown parameters from \mathbf{Q} .

2.4.1 Cramér-Rao Bounds and Asymptotic Errors

We first present the CRB on DOA estimation of mixed signals in the following theorem.

Theorem 1. For DOA estimation of mixed signals, the mixed CRB matrix $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ can be written as

$$\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}} = [\mathbf{CRB}_{\mathbf{d},\boldsymbol{\theta}}^{-1} + \mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}^{-1}]^{-1}, \quad (2.79)$$

where $\mathbf{CRB}_{\mathbf{d},\boldsymbol{\theta}}$ and $\mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}$ are the deterministic and stochastic CRB matrices on DOA estimation with measurements from $\mathcal{CN}(\mathbf{A}\mathbf{b}, \mathbf{Q})$ and $\mathcal{CN}(\mathbf{0}, \mathbf{R})$, respectively, in which the unknown parameters are $\boldsymbol{\theta}$, \mathbf{b} , \mathbf{P} , and $\boldsymbol{\sigma}$.

Proof: See Appendix A.

The CRBs on 1D DOA estimation of deterministic and stochastic signals have been well addressed in [19] and [43]. Herein, we extend these results to 2D DOA estimation.

Proposition 1. The CRB matrix on 2D DOA estimation of deterministic signals is

$$\mathbf{CRB}_{\mathbf{D},\boldsymbol{\theta}} = \frac{1}{2N} \text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \Pi_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \left(\hat{\mathbf{R}}_{\mathbf{x}\mathbf{x}}^T \otimes \boldsymbol{\Delta} \right) \right\}^{-1}, \quad (2.80)$$

and the resulting asymptotic CRB matrix is

$$\mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}} = \frac{1}{2N} \text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \Pi_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \left(\mathbf{R}_{\mathbf{x}\mathbf{x}}^T \otimes \boldsymbol{\Delta} \right) \right\}^{-1}. \quad (2.81)$$

Proposition 2. The CRB matrix on 2D DOA estimation of stochastic signals is

$$\mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}} = \frac{1}{N} \left(\boldsymbol{\Omega} - \mathbf{M}\mathbf{T}^{-1}\mathbf{M}^T \right)^{-1}, \quad (2.82)$$

where

$$\mathbf{\Omega} = 2\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \left[\left(\mathbf{P} \tilde{\mathbf{A}}^H \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{A}} \mathbf{P} \right)^T \otimes \mathbf{\Delta} \right] \right\}, \quad (2.83)$$

$$\mathbf{M}_{kl} = 2\text{Re} \left\{ \tilde{\mathbf{d}}_{2,k}^H \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{Q}}_l' \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{A}} \mathbf{p}_{2,k} \right\}, \quad (2.84)$$

$$\mathbf{T}_{kl} = 2\text{Re} \left\{ \text{trace} \left\{ \tilde{\mathbf{Q}}_k' \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{Q}}_l' \tilde{\mathbf{R}}^{-1} \right\} \right\} - \text{trace} \left\{ \tilde{\mathbf{Q}}_k' \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{Q}}_l' \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \right\}, \quad (2.85)$$

$$\mathbf{M} = 2\text{Re} \left\{ \mathbf{\Xi}^T \left[\left(\tilde{\mathbf{D}}_2^H \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \right) \otimes \left(\mathbf{P}_2^T \tilde{\mathbf{A}}^T \tilde{\mathbf{R}}^{-T} \right) \right] \mathbf{\Lambda}^* \right\}, \quad (2.86)$$

$$\mathbf{T} = 2\text{Re} \left\{ \mathbf{\Lambda}^H \left(\tilde{\mathbf{R}}^{-T} \otimes \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \right) \mathbf{\Lambda} \right\} - \mathbf{\Lambda}^H \left(\left(\mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \right)^T \otimes \mathbf{\Pi}_{\tilde{\mathbf{A}}}^\perp \right) \mathbf{\Lambda}, \quad (2.87)$$

where $\{\cdot\}^*$ denotes the complex conjugate, \mathbf{M}_{kl} and \mathbf{T}_{kl} are the (k,l) -th elements of \mathbf{M} and \mathbf{T} respectively, and $\tilde{\mathbf{d}}_{2,k}$ and $\mathbf{p}_{2,k}$ are the k -th columns of $\tilde{\mathbf{D}}_2$ and \mathbf{P}_2 respectively.

Note that if $\mathbf{Q} = \sigma^2 \mathbf{I}$, we have $\mathbf{M} = \mathbf{0}$, and (2.82) simplifies to

$$\text{CRB}_{\text{S},\theta} = \frac{1}{N} \mathbf{\Omega}^{-1}. \quad (2.88)$$

When $\mathbf{Q} = \sigma^2 \mathbf{I}$, the asymptotic error covariance matrix for deterministic ML 1D DOA estimation is given in [9] as

$$\begin{aligned} \text{AC}_{\text{D},\theta} &= \frac{\sigma^2}{2N} \left[\text{Re} \left\{ \left(\mathbf{D}^H \mathbf{\Pi}_{\mathbf{A}}^\perp \mathbf{D} \right) \odot \mathbf{R}_{\mathbf{x}\mathbf{x}}^T \right\} \right]^{-1} \text{Re} \left\{ \left(\mathbf{D}^H \mathbf{\Pi}_{\mathbf{A}}^\perp \mathbf{D} \right) \right. \\ &\quad \left. \odot \left[\mathbf{R}_{\mathbf{x}\mathbf{x}} + \sigma^2 \left(\mathbf{A}^H \mathbf{A} \right)^{-1} \right]^T \right\} \left[\text{Re} \left\{ \left(\mathbf{D}^H \mathbf{\Pi}_{\mathbf{A}}^\perp \mathbf{D} \right) \odot \mathbf{R}_{\mathbf{x}\mathbf{x}}^T \right\} \right]^{-1}. \end{aligned} \quad (2.89)$$

For 2D DOA estimation, we modify this result as follows.

Proposition 3. *For 2D DOA estimation under spatially white noise $\mathbf{Q} = \sigma^2 \mathbf{I}$, the asymptotic error covariance matrix of the deterministic ML DOA estimator is*

$$\begin{aligned}
\mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}} &= \frac{\sigma^2}{2N} \left[\text{Re} \left\{ \left(\mathbf{D}_2^H \boldsymbol{\Pi}_A^\perp \mathbf{D}_2 \right) \odot \left(\mathbf{R}_{xx}^T \otimes \boldsymbol{\Delta} \right) \right\} \right]^{-1} \\
&\times \text{Re} \left\{ \left(\mathbf{D}_2^H \boldsymbol{\Pi}_A^\perp \mathbf{D}_2 \right) \odot \left[\left(\mathbf{R}_{xx} + \sigma^2 \left(\mathbf{A}^H \mathbf{A} \right)^{-1} \right)^T \otimes \boldsymbol{\Delta} \right] \right\} \\
&\times \left[\text{Re} \left\{ \left(\mathbf{D}_2^H \boldsymbol{\Pi}_A^\perp \mathbf{D}_2 \right) \odot \left(\mathbf{R}_{xx}^T \otimes \boldsymbol{\Delta} \right) \right\} \right]^{-1}. \tag{2.90}
\end{aligned}$$

Propositions 1 to 3 can be proved following the same procedures as in [19], [43], and [9]. Details of the proofs are omitted here.

Now consider the asymptotic error covariance matrix of applying a stochastic ML estimator on DOA estimation of mixed signals.

Theorem 2. *If a stochastic ML estimator is applied on DOA estimation of mixed signals with correlation matrix \mathbf{R}_{xx} , then the asymptotic error covariance matrix $\mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}$ is equal to the stochastic CRB matrix on DOA estimation with measurements from $\mathcal{CN}(\mathbf{0}, \bar{\mathbf{R}})$, where*

$$\bar{\mathbf{R}} = \mathbf{A} \mathbf{R}_{xx} \mathbf{A}^H + \mathbf{Q} \tag{2.91}$$

and the unknown parameters are $\boldsymbol{\theta}$, \mathbf{R}_{xx} , and σ .

Proof: See Appendix B, in which the following lemma is used in deriving the result.

Lemma 2. *Suppose $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$, \mathbf{B} and \mathbf{D} are two square matrices with the same size as \mathbf{C} . Then we have*

$$\begin{aligned}
\mathbb{E} \left\{ \mathbf{x}^H \mathbf{B} \mathbf{x} \mathbf{x}^H \mathbf{D} \mathbf{x} \right\} &= \text{trace} \left\{ \mathbf{B} \left(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H \right) \right\} \text{trace} \left\{ \mathbf{D} \left(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H \right) \right\} \\
&+ \text{trace} \left\{ \mathbf{B} \left(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H \right) \mathbf{D} \left(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H \right) \right\} - \boldsymbol{\mu}^H \mathbf{B} \boldsymbol{\mu} \boldsymbol{\mu}^H \mathbf{D} \boldsymbol{\mu}. \tag{2.92}
\end{aligned}$$

Proof: See Appendix C.

2.4.2 Analytical Performance Comparisons

In this section, we present some analytical comparison results on the performances of the three types of estimators. We first examine the performance of the mixed ML estimator on DOA estimation of stochastic signals.

Proposition 4. *For DOA estimation of stochastic signals,*

$$\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}} = \mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}. \quad (2.93)$$

Proof: According to the results in (A.7), (A.8), (A.11), and (A.15) in Appendix A, we can see that $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}^{-1} = \mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}^{-1}$ if $\mathbf{b} = \mathbf{0}$, from which we obtain (2.93).

Proposition 4 shows that for DOA estimation of stochastic signals, the asymptotic accuracy of the mixed ML estimator is equal to that of the stochastic one.

For DOA estimation of stochastic signals, it was shown [43] that the asymptotic deterministic CRB is not larger than the stochastic CRB. We compare the asymptotic deterministic CRB with the mixed CRB and obtain the similar result as follows.

Proposition 5. *For DOA estimation of mixed signals,*

$$\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}} \geq \mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}}. \quad (2.94)$$

Proof: See Appendix D.

For DOA estimation of mixed signals, it seems difficult to make analytical comparisons for $\mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}}$, $\mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}$, and $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ under arbitrary proper \mathbf{Q} , since a closed-form $\mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}}$

is still not available and an analytical comparison between $\mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}$ and $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ seems difficult. However, we have the following proposition holds under the special case $\mathbf{Q} = \sigma^2 \mathbf{I}$.

Proposition 6. *For ML DOA estimation of mixed signals under spatially white noise $\mathbf{Q} = \sigma^2 \mathbf{I}$, we have*

$$\mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}} \geq \mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}} \geq \mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}. \quad (2.95)$$

Proof: See Appendix E. The following lemma is used in deriving $\mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}} \geq \mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}$ for 2D DOA estimation.

Lemma 3. *Let \mathbf{A} and \mathbf{C} be two nonnegative definite matrices of the same size as \mathbf{B} , a Hermitian matrix, and let \mathbf{C}^\dagger be the Moore-Penrose pseudoinverse of \mathbf{C} . Suppose $\mathcal{N}\{\mathbf{C}\} \subseteq \mathcal{N}\{\mathbf{B}\}$, where $\mathcal{N}\{\cdot\}$ denotes the null space of a matrix. Then,*

$$\{\text{Re}\{\mathbf{A} \odot \mathbf{B}\}\}^{-1} \text{Re}\{\mathbf{A} \odot \mathbf{C}\} \{\text{Re}\{\mathbf{A} \odot \mathbf{B}\}\}^{-1} \geq \{\text{Re}\{\mathbf{A} \odot (\mathbf{B}\mathbf{C}^\dagger\mathbf{B})\}\}^{-1} \quad (2.96)$$

if $\{\text{Re}\{\mathbf{A} \odot \mathbf{C}\}\}^{-1}$ and all the matrix inverses in (2.96) exist.

Proof: See Appendix F.

When the noise is spatially white, the second inequality in (2.95) shows that with the same signal correlation matrix and noise power, the mixed signals improve the DOA estimation accuracy compared with the stochastic ones, and the mixed ML estimator provides better performance than the stochastic ML estimator.

We have not shown (2.95) analytically for arbitrary proper noise covariance matrices. However, since $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ is the CRB on mixed signal DOA estimation, we should

always have $\mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}} \geq \mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ and $\mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}} \geq \mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$, where the equalities do not always hold due to the result in Proposition 6.

For the more special case of 1D DOA estimation of a single source under spatially white noise, we have the following proposition.

Proposition 7. *For 1D DOA estimation of a single stochastic or mixed signal under $\mathbf{Q} = \sigma^2 \mathbf{I}$, the asymptotic mean-square error of the deterministic ML estimator is equal to that of the stochastic estimator, i.e., $\mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}} = \mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}$.*

Proof: For 1D DOA estimation of a single source, $\mathbf{D}^H \boldsymbol{\Pi}_{\mathbf{A}}^\perp \mathbf{D}$, \mathbf{R}_{xx} , and $(\mathbf{A}^H \mathbf{A})^{-1}$ are all real scalars. According to the result in (2.89), we have

$$\begin{aligned} \mathbf{AC}_{\mathbf{D},\boldsymbol{\theta}} &= \frac{\sigma^2}{2N} \left[(\mathbf{D}^H \boldsymbol{\Pi}_{\mathbf{A}}^\perp \mathbf{D})^{-1} \mathbf{R}_{xx}^{-T} \right] \left(\mathbf{D}^H \boldsymbol{\Pi}_{\mathbf{A}}^\perp \mathbf{D} \right) \\ &\quad \times \left[\mathbf{R}_{xx} + \sigma^2 (\mathbf{A}^H \mathbf{A})^{-1} \right]^T \left[(\mathbf{D}^H \boldsymbol{\Pi}_{\mathbf{A}}^\perp \mathbf{D})^{-1} \mathbf{R}_{xx}^{-T} \right] \\ &= \frac{\sigma^2}{2N} \left\{ (\mathbf{D}^H \boldsymbol{\Pi}_{\mathbf{A}}^\perp \mathbf{D}) \left[\mathbf{R}_{xx}^{-1} + \sigma^2 \mathbf{R}_{xx}^{-1} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{R}_{xx}^{-1} \right]^{-T} \right\}^{-1} = \mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}. \end{aligned} \quad (2.97)$$

The last equality holds from the result in [8].

2.5 Numerical Examples

In this section, we compare the performances of the deterministic, stochastic, and mixed ML DOA estimators through two numerical examples.

In the first example, we consider a 1D DOA estimation problem of two sources using a linear scalar-sensor array consisting of four separated linear sub-arrays, which are composed of 3, 4, 3, and 4 scalar sensors, respectively. The distance between any

two adjacent sub-arrays is 3λ , where λ is the wavelength of the narrow-band signal. Within each sub-array, the spacing between scalar sensors is 0.5λ . The linear sensor array lies on the z axis. The elevation angles of the two sources are 15 and 30 degrees. The elements of the noise covariance matrix are generated with the following widely used noise model (see [24] and the references therein):

$$[\mathbf{Q}_i]_{kl} = \sigma_i^2 \exp \{ -(k-l)^2 \zeta_i \}, \quad (2.98)$$

where $\sigma_1^2 = 2$, $\sigma_2^2 = 3$, $\sigma_3^2 = 4$, $\sigma_4^2 = 5$, $\zeta_1 = 0.6$, $\zeta_2 = 0.7$, $\zeta_3 = 0.8$, and $\zeta_4 = 0.9$. We consider both stochastic and mixed signals and apply the deterministic, stochastic, and mixed ML DOA estimators to each of them. For the stochastic signal case, we assume the signal covariance matrix

$$\mathbf{P} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}. \quad (2.99)$$

For the mixed signal case, we assume the signal mean $\mathbf{b} = [0.7, 0]^T$ and the signal covariance matrix

$$\mathbf{P} = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.8 \end{bmatrix}. \quad (2.100)$$

We consider the estimation algorithm achieves convergence at the n -th iteration if the Euclidean norm of $\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_{n-1}$ is smaller than an error tolerance. The root-mean-square errors (RMSEs) from three types of estimators are illustrated in Figs. 2.1 and 2.2 for the source with $\theta = 15$ degrees. The mean square errors are calculated with 200 Monte Carlo runs.

We first examine the performances of the deterministic ML DOA estimator (Algorithm 1). From Figs. 2.1 and 2.2, we observe that the accuracy of the deterministic estimator does not always increase with iterations going forward. Furthermore, in the running of our MATLAB program, the deterministic estimator in Algorithm 1 diverged the value of the deterministic log-likelihood function to infinity and introduced errors in MATLAB usually within 10 iterations. Examining the deterministic log-likelihood function in (2.15), we can see that the deterministic ML DOA estimator finds the DOA estimate by maximizing $-\log|\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E}|$ over $\boldsymbol{\theta}$ and \mathbf{X} , where $\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) = [\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{X}][\mathbf{Y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{X}]^H$. Due to the large number of nuisance parameters in \mathbf{X} , no matter what value $\boldsymbol{\theta}$ takes, there always exist \mathbf{X} values that make $\mathbf{C}(\boldsymbol{\theta}, \mathbf{X}) \odot \mathbf{E}$ singular and the value of the log-likelihood function in (2.15) infinite. This introduces severe instability in DOA estimation and explains the phenomena we observed in the figures and program running. However, the first few iterations of Algorithm 1 are still able to provide close DOA estimates, as shown in Figs. 2.1 and 2.2, as well as in [24].

We now examine the performances of the stochastic estimator (Algorithm 2) and the mixed estimator (Algorithm 3). Both algorithms are initialized using the results from one iteration of Algorithm 1. Specifically, suppose $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{Q}}$, and $\hat{\mathbf{X}}$ are the estimates from one iteration of Algorithm 1. Then $\boldsymbol{\theta}$ and \mathbf{Q} are initialized as $\boldsymbol{\theta}_{\text{init}} = \hat{\boldsymbol{\theta}}$ and $\mathbf{Q}_{\text{init}} = \hat{\mathbf{Q}}$ for both estimators. For the stochastic estimator, \mathbf{P} is initialized as $\mathbf{P}_{\text{init}} = \hat{\mathbf{X}}\hat{\mathbf{X}}^H/N$. For the mixed estimator, \mathbf{b} and \mathbf{Q} are initialized as

$$\mathbf{b}_{\text{init}} = \bar{\hat{\mathbf{x}}} = \frac{1}{N} \sum_{t=1}^N \hat{\mathbf{x}}(t), \quad (2.101)$$

$$\mathbf{P}_{\text{init}} = \frac{1}{N} \sum_{t=1}^N [\hat{\mathbf{x}}(t) - \bar{\hat{\mathbf{x}}}] [\hat{\mathbf{x}}(t) - \bar{\hat{\mathbf{x}}}]^H. \quad (2.102)$$

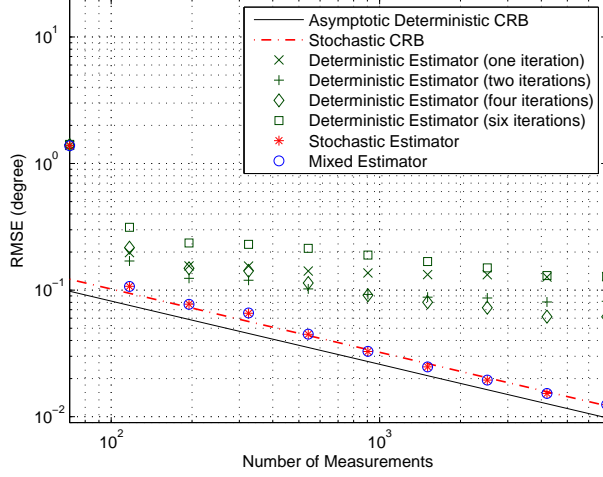


Figure 2.1: Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from stochastic signal one-dimensional DOA estimation.

The optimizations over θ in Steps 3 of both algorithms are implemented iteratively, starting from θ_{init} or the $\hat{\theta}$ value from the previous iteration. The error tolerance used is 10^{-2} degree for both Figs. 1 and 2. Algorithms 2 and 3 both converge with an average number of 3 iterations or so in our simulations for either Fig. 1 or 2.

From Fig. 2.1, we see that for DOA estimation of stochastic signals, the performances of the stochastic and the mixed estimators are similar to each other and are normally better than that of the deterministic estimator. From Fig. 2.2, we observe that for DOA estimation of the mixed signals, the mixed estimator provides higher accuracy than either the deterministic or the stochastic estimator.

In our second example, we consider a 2D DOA estimation problem of a single source using a uniform linear acoustic vector-sensor (AVS) array [11]. The array consists of six vector sensors and lies on the z axis. Each vector sensor consists of four sensors measuring acoustic pressure and three acoustic particle velocity components, respectively. The two adjacent vector sensors are 0.5λ apart from each other. The

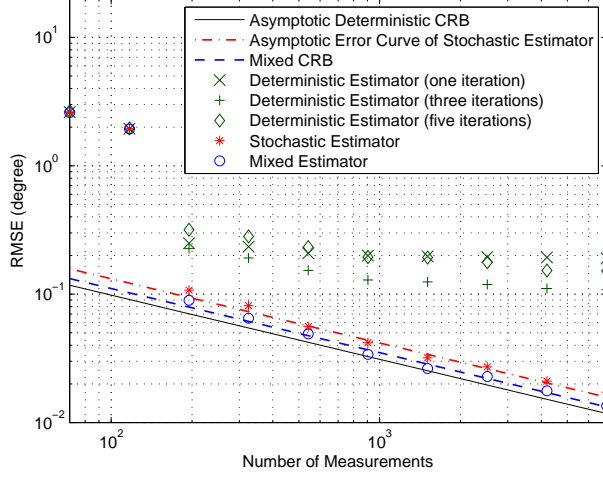


Figure 2.2: Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from mixed signal one-dimensional DOA estimation.

elevation and azimuth angles of the source are 15 and 30 degrees, respectively. We assume the noise is correlated among the four sensors within one vector sensor but uncorrelated between different vector sensors. Furthermore, we assume the 4×4 noise covariance matrices from different vector sensors are equal. This noise scenario matches the situation of an underwater fast-towed linear AVS array, where the flow noise is the dominant noise on the array. When the distance between vector sensors is not smaller than 0.5λ , the covariance matrix of the noise can be well approximated by a block-diagonal matrix with all blocks equal [45]. The realistic modeling of flow noise is not the topic of this dissertation. Herein, we assume the 4×4 noise covariance matrix on each vector sensor follows equation (2.98) with $\sigma^2 = 2$ and $\zeta = 0.5$. We consider both stochastic and mixed signals, and apply the deterministic, stochastic, and mixed ML DOA estimators to each of them. For the stochastic signal case, we assume the noise power is 0.5. For the mixed signal case, we assume the signal mean is 0.3 and the signal variance is 0.1. The estimation errors from three types of estimators are illustrated in Figs. 2.3 and 2.4 for the elevation angle. The mean-square errors are

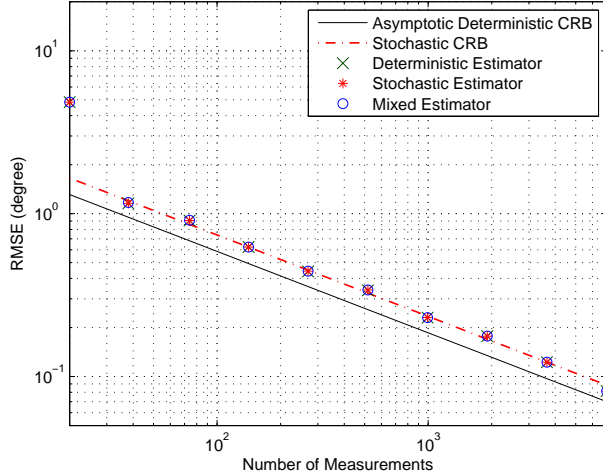


Figure 2.3: Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from stochastic signal two-dimensional DOA estimation with the diagonal blocks equal in the noise covariance matrix.

also computed with 200 Monte Carlo runs. The error tolerance used is 10^{-2} degree for both Figs. 3 and 4. In our simulations for either Fig. 3 or 4, Algorithm 1 converges with an average number of 4 iterations or so, and both Algorithms 2 and 3 converge with an average number of about 3 iterations.

From Figs. 2.3 and 2.4, we observe that the mixed estimator performs as well as the stochastic estimator in the stochastic signal case and better in mixed signal case. These results are similar to those in our first example. The phenomenon distinguishing this example from the first one is that with all diagonal blocks equal in the noise covariance matrix, the deterministic ML DOA estimator converges with ongoing iterations and provides estimation accuracy close to that of the stochastic estimator, as shown in Figs. 2.3 and 2.4. The conditions that guarantee the convergence of the deterministic estimator are not within the scope of this dissertation.

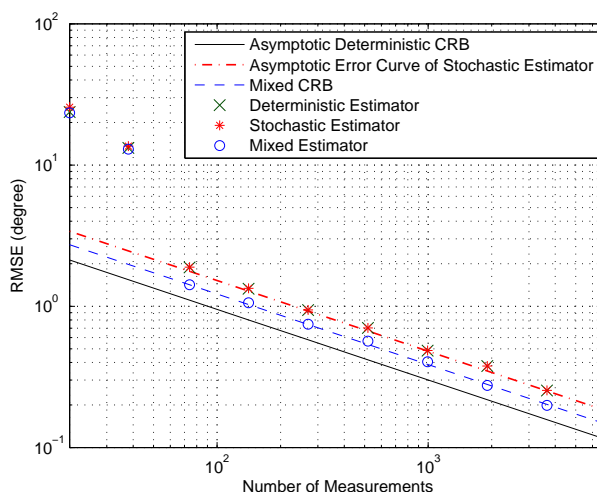


Figure 2.4: Cramér-Rao bounds and root-mean-square errors as functions of the number of measurements from mixed signal two-dimensional DOA estimation with the diagonal blocks equal in the noise covariance matrix.

So far, we only gave the numerical results produced by Algorithm 2 for the stochastic ML estimator. We also ran the simulations using the pure-EM algorithm based on (2.43)-(2.46) as the stochastic ML estimator. The results showed that the pure-EM algorithm is able to provide similar performance to that of Algorithm 2, but it normally requires smaller error tolerances and more iterations and running time to do so. As expected, the pure-EM algorithm is less efficient than Algorithm 2. The estimation results from the pure-EM algorithm are omitted here since they are just similar to those of Algorithm 2. The detailed comparisons for these two algorithms are not within the interest of this dissertation.

2.6 Summary

We considered the problem of narrow-band DOA estimation under spatially colored noise using sparse sensor arrays and proposed new methods for ML DOA estimation of stochastic and mixed signals based on an EM framework. We gave the CRB on DOA estimation of mixed signals. We also derived the asymptotic error covariance matrix of applying the stochastic DOA estimator on mixed signal DOA estimation. We presented both analytical and numerical comparisons for the performances of the deterministic, stochastic, and mixed estimators. Our results showed that: (i) the performance of the deterministic estimator is normally inferior to those of the stochastic and mixed estimators; (ii) for DOA estimation of stochastic signals, the mixed estimator provides similar performance to that of the stochastic estimator; (iii) for DOA estimation of mixed signals, the mixed estimator yields better performance than the stochastic estimator.

Chapter 3

Narrow-Band Direction-of-Arrival Estimation of Hydroacoustic Signals From Marine Vessels ²

As a topic in Chapter 2, we considered the direction-of-arrival (DOA) estimation of mixed signals, which are mixtures of non-zero means with the typical zero-mean Gaussian signals assumed by the many existing DOA estimators. In this chapter, we consider the DOA estimation problem of another type of mixed signals – the underwater acoustic signals (hydroacoustic signals) from marine vessels like ships, submarines, or torpedoes, which contain both sinusoidal and random components. We model this type of signals as the sum of deterministic sinusoidal signals and zero-mean Gaussian signals, and derive the maximum likelihood (ML) DOA estimator for them under spatially white noise. We compute the asymptotic error covariance matrix of the proposed ML estimator, as well as that of the typical ML estimator assuming zero-mean Gaussian signals, for DOA estimation of this type of signals. Our

²Based on T. Li and A. Nehorai, “Maximum Likelihood Direction-of-Arrival Estimation of Underwater Acoustic Signals Containing Sinusoidal Components,” *IEEE Trans. Signal Process.*, vol. 59, pp. 5302-5314, Nov. 2011. ©[2011] IEEE.

analytical comparison and numerical examples show that the proposed ML estimator improves the DOA estimation accuracy for the hydroacoustic signals from marine vessels compared with the typical ML estimator assuming zero-mean Gaussian signals.

3.1 Introduction

Direction-of-arrival (DOA) estimation plays an important role in underwater sonar applications such as the localization and tracking of ships and submarines. Among diverse parameter estimation techniques, maximum likelihood (ML) estimation is distinguished by its excellent asymptotic estimation performance often able to achieve the Cramér-Rao bound (CRB) [42], which is highly desirable in many underwater DOA estimation scenarios where the signal-to-noise ratios (SNRs) are usually low whereas high DOA estimation accuracy is required.

ML methods estimate the DOA by maximizing the likelihood functions, which vary with the models describing the signals. Among existing signal models, the deterministic and the stochastic models are the most widely exploited in ML DOA estimation (see [7], [8], [10], [19], [20], and [24] for examples). The deterministic signal model (see [7], [8], [19], and [24]) considers the signal values at all snapshots as deterministic unknown parameters. As a result, the number of unknown parameters in the likelihood function increases with the number of measurements, such that the ML DOA estimator based on this model turns out not to be able to achieve the deterministic CRB [8]. In contrast, the stochastic signal model (see [8], [10], and [20]) considers the signals to be random processes following specific probability density functions (pdfs). As a consequence, the unknown signal parameters in the likelihood function are fixed in size with respect to the change in the number of measurements. The ML DOA

estimator based on the stochastic signal model achieves the stochastic CRB and is shown to be able to provide better performance than the ML estimator based on the deterministic signal model [8].

In most existing ML DOA estimation works using stochastic signal models, the signals are assumed to follow Gaussian distributions with zero means (see [8], [10], and [20] for examples). Though the zero-mean Gaussian distribution is able to well describe a wide range of signals in practice, there exist applications where the signals cannot be accurately characterized in this way (see [14] and [25]), and better estimation accuracy may be achieved by using more precise signal models. In this chapter, we consider the DOA estimation problem of hydroacoustic signals containing sinusoidal components, which cannot be well described by the zero-mean Gaussian distribution.

The hydroacoustic signals from ships, submarines, or torpedoes are known to consist of two parts ([31], [48], [49]): the noise-like part with continuous spectra, and the sinusoidal part with discrete frequencies. In this chapter, we call this type of signals “mixed signals”, and model them as the mixture of a zero-mean Gaussian part with unknown covariance matrix and a sinusoidal part with unknown coefficients and frequencies, which correspond to the noise-like part with continuous spectra and the sinusoidal part with discrete frequencies, respectively. For simplicity of notation and presentation, in the rest of this chapter, we use “stochastic signals” to represent zero-mean Gaussian signals. We use “stochastic” and “mixed” estimators or CRBs to represent the ML estimators or CRBs derived under the assumptions of “stochastic” and “mixed” signals, respectively.

In this chapter, we derive the ML estimator and its asymptotic error covariance matrix for DOA estimation of mixed signals. In addition, we compute the asymptotic error

covariance matrix of applying the stochastic estimator to DOA estimation of mixed signals. We provide both analytical and numerical comparisons for the stochastic and the proposed mixed estimators. The results show that the proposed mixed signal model and estimator improve the DOA estimation accuracy for mixed signals.

The remainder of this chapter is organized as follows. We first give our measurement model in Section 3.2, and derive the ML estimator in Section 3.3. Section 3.4 provides the results of analytical performance analysis. Numerical examples and conclusions appear in Sections 3.5 and 3.6, respectively.

3.2 Measurement Model

In this chapter, we limit our problem to narrow-band DOA estimation of mixed signals. Recall that the signal can be considered as narrow-band if $D/c \ll 1/B$, where D , c , and B are the array length, the signal propagation speed, and the signal bandwidth, respectively. Consider the mixed signals from L far-field sources impinging on an array of M sensors. We write the narrow-band array output as

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta}) [\mathbf{C}\boldsymbol{\varphi}(\boldsymbol{\omega}, t) + \mathbf{x}(t)] + \boldsymbol{\epsilon}(t), \quad t = 1, \dots, N, \quad (3.1)$$

where $\mathbf{y}(t)$ is the $M \times 1$ measurement vector at the t -th snapshot,

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_L)] \quad (3.2)$$

is the array steering matrix, $\mathbf{a}(\theta_l)$ is the steering vector corresponding to the l -th source, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_L]^T$ is the DOA vector with θ_l the DOA of the l -th source,

$\{\cdot\}^T$ denotes the matrix transpose, $\boldsymbol{\omega} = [\omega_1, \dots, \omega_J]^T$ with J the total number of sinusoidal components with different frequencies in the received signals and ω_m the radian frequency of the m -th sinusoidal component, $\boldsymbol{\varphi}(\boldsymbol{\omega}, t) = [e^{j\omega_1 t}, \dots, e^{j\omega_J t}]^T$ with $j^2 = -1$, \mathbf{C} is the $L \times J$ matrix containing the coefficients of $\boldsymbol{\varphi}(\boldsymbol{\omega}, t)$ for all sources, $\mathbf{x}(t)$ is the $L \times 1$ Gaussian signal vector, $\boldsymbol{\epsilon}(t)$ is the $M \times 1$ Gaussian noise vector, and N is the total number of measurements.

In equation (3.1), the mixed signals are represented by the sum of $\mathbf{C}\boldsymbol{\varphi}(\boldsymbol{\omega}, t)$ and $\mathbf{x}(t)$, which correspond to the parts with discrete frequencies and continuous spectra, respectively. We assume $\mathbf{x}(t)$ and $\boldsymbol{\epsilon}(t)$ follow zero-mean circularly complex Gaussian distributions with unknown covariance matrices \mathbf{P} and \mathbf{Q} , respectively. We further assume $\mathbf{x}(t)$ and $\boldsymbol{\epsilon}(t)$ are both temporally white and uncorrelated with each other. Additionally, we assume L is known, which is a quite common assumption in existing DOA estimation research (see [7], [8], [10], [19], [20], and [24] for examples). Also, we assume J is known, which is a quite common assumption as well in existing sinusoidal frequency estimation research (see [33], [34], [36], [50], [51] for examples). The unknown parameters in (3.1) are those from $\boldsymbol{\theta}$, \mathbf{C} , $\boldsymbol{\omega}$, \mathbf{P} , and \mathbf{Q} .

3.3 Maximum Likelihood Estimation

In this section, we present the ML estimator for DOA finding of mixed signals, which is called the mixed estimator in this chapter. For simplicity, we omit $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ in notations, and use \mathbf{A} and $\boldsymbol{\varphi}(t)$ to represent $\mathbf{A}(\boldsymbol{\theta})$ and $\boldsymbol{\varphi}(\boldsymbol{\omega}, t)$, respectively.

Combining measurements from all snapshots, we rewrite the narrow-band measurement model in (3.1) as

$$\mathbf{Y} = \mathbf{A}\mathbf{C}\boldsymbol{\phi} + \mathbf{A}\mathbf{X} + \mathbf{E}, \quad (3.3)$$

where $\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(N)]$, $\boldsymbol{\phi} = [\boldsymbol{\varphi}(1), \dots, \boldsymbol{\varphi}(N)]$, $\mathbf{X} = [\mathbf{x}(1), \dots, \mathbf{x}(N)]$, and $\mathbf{E} = [\boldsymbol{\epsilon}(1), \dots, \boldsymbol{\epsilon}(N)]$. Vectorizing both sides of (3.3), we obtain

$$\mathbf{y} = [\boldsymbol{\phi}^T \otimes \mathbf{A}] \text{vec}\{\mathbf{C}\} + \text{vec}\{\mathbf{A}\mathbf{X} + \mathbf{E}\}, \quad (3.4)$$

where $\mathbf{y} = \text{vec}\{\mathbf{Y}\} = [\mathbf{y}^T(1), \dots, \mathbf{y}^T(N)]^T$, \otimes denotes the Kronecker product, and $\text{vec}\{\cdot\}$ denotes the vectorization operator stacking all the columns of a matrix, one below another, into a vector. Note that in the derivation of (3.4), we use the property that

$$\text{vec}\{\mathbf{A}\mathbf{B}\mathbf{C}\} = [\mathbf{C}^T \otimes \mathbf{A}] \text{vec}\{\mathbf{B}\} \quad (3.5)$$

for any matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} that can make $\mathbf{A}\mathbf{B}\mathbf{C}$ [44].

From (3.4), we have

$$\mathbf{y} \sim \mathcal{CN}\left([\boldsymbol{\phi}^T \otimes \mathbf{A}] \text{vec}\{\mathbf{C}\}, \mathbf{I}_N \otimes (\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})\right), \quad (3.6)$$

where $\mathcal{CN}(\cdot)$ denotes the complex Gaussian distribution, \mathbf{I}_N is the $N \times N$ identity matrix, and $\{\cdot\}^H$ denotes the conjugate transpose. From (3.6), we obtain the ML estimate for $\text{vec}\{\mathbf{C}\}$ as

$$\begin{aligned} \text{vec}\{\hat{\mathbf{C}}\} &= \left\{ [\boldsymbol{\phi}^T \otimes \mathbf{A}]^H \left[\mathbf{I}_N \otimes (\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1} \right] \right. \\ &\quad \left. \times [\boldsymbol{\phi}^T \otimes \mathbf{A}] \right\}^{-1} [\boldsymbol{\phi}^T \otimes \mathbf{A}]^H \left[\mathbf{I}_N \otimes (\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1} \right] \mathbf{y} \\ &= \left\{ [(\boldsymbol{\phi}^* \boldsymbol{\phi}^T)^{-1} \boldsymbol{\phi}^*] \otimes \left\{ \left[\mathbf{A}^H (\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1} \mathbf{A} \right]^{-1} \right. \right. \\ &\quad \left. \left. \times \mathbf{A}^H (\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1} \right\} \right\} \mathbf{y}, \end{aligned} \quad (3.7)$$

where $\{\cdot\}^*$ denotes the complex conjugate.

In Appendix G, we show that

$$\left[\mathbf{A}^H(\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^H(\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1} = (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}. \quad (3.8)$$

As a result, equation (3.7) becomes

$$\text{vec}\{\hat{\mathbf{C}}\} = \left\{ \left[(\boldsymbol{\phi}^* \boldsymbol{\phi}^T)^{-1} \boldsymbol{\phi}^* \right] \otimes \left[(\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{Q}^{-1} \right] \right\} \mathbf{y}, \quad (3.9)$$

from which we obtain the ML estimate for \mathbf{C} as

$$\begin{aligned} \hat{\mathbf{C}} &= (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{Q}^{-1} \mathbf{Y} \boldsymbol{\phi}^H (\boldsymbol{\phi} \boldsymbol{\phi}^H)^{-1} \\ &= (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H \tilde{\mathbf{Y}} \boldsymbol{\phi}^H (\boldsymbol{\phi} \boldsymbol{\phi}^H)^{-1}, \end{aligned} \quad (3.10)$$

where $\tilde{\mathbf{A}} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A}$ and $\tilde{\mathbf{Y}} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{Y}$. We can see that the ML estimate of \mathbf{C} is a function of $\boldsymbol{\theta}$, $\boldsymbol{\omega}$, and \mathbf{Q} but independent of the signal covariance matrix \mathbf{P} . This result is important for the further reduction of \mathbf{P} from the likelihood function.

By omitting constant terms, the log-likelihood (LL) function can be written as

$$L(\boldsymbol{\theta}, \mathbf{C}, \boldsymbol{\omega}, \mathbf{P}, \mathbf{Q}) = -\log |\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}| - \text{trace} \left\{ [\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}]^{-1} \mathbf{R}_{yy} \right\}, \quad (3.11)$$

where $|\cdot|$ denotes the matrix determinant, and

$$\mathbf{R}_{yy} = \frac{1}{N} (\mathbf{Y} - \mathbf{A}\mathbf{C}\boldsymbol{\phi}) (\mathbf{Y} - \mathbf{A}\mathbf{C}\boldsymbol{\phi})^H. \quad (3.12)$$

Substituting (3.10) into (3.11), we rewrite the LL function as

$$L(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{P}, \mathbf{Q}) = -\log |\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}| - \text{trace} \left\{ [\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}]^{-1} \hat{\mathbf{R}}_{yy} \right\} \quad (3.13)$$

$$= -\log |\mathbf{Q}| - \log |\tilde{\mathbf{A}}\mathbf{P}\tilde{\mathbf{A}}^H + \mathbf{I}_M| - \text{trace} \left\{ [\tilde{\mathbf{A}}\mathbf{P}\tilde{\mathbf{A}}^H + \mathbf{I}_M]^{-1} \tilde{\mathbf{R}}_{yy} \right\}, \quad (3.14)$$

where

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N} (\mathbf{Y} - \mathbf{A}\hat{\mathbf{C}}\boldsymbol{\phi}) (\mathbf{Y} - \mathbf{A}\hat{\mathbf{C}}\boldsymbol{\phi})^H, \quad (3.15)$$

and

$$\begin{aligned} \tilde{\mathbf{R}}_{yy} &= \mathbf{Q}^{-\frac{1}{2}} \hat{\mathbf{R}}_{yy} \mathbf{Q}^{-\frac{1}{2}} \\ &= \frac{1}{N} (\tilde{\mathbf{Y}} - \boldsymbol{\Pi}_{\tilde{\mathbf{A}}} \tilde{\mathbf{Y}} \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^H) (\tilde{\mathbf{Y}} - \boldsymbol{\Pi}_{\tilde{\mathbf{A}}} \tilde{\mathbf{Y}} \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^H)^H, \end{aligned} \quad (3.16)$$

in which $\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} = \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H$.

In (3.14), the last two terms on the right-hand side of the equality constitute a function similar to the Gaussian LL function for stochastic signals under spatially white noise [10]. Following procedures similar to those in [10], we have the ML estimate for \mathbf{P} as

$$\hat{\mathbf{P}} = [\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}]^{-1} \tilde{\mathbf{A}}^H \tilde{\mathbf{R}}_{yy} \tilde{\mathbf{A}} [\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}]^{-1} - [\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}]^{-1}, \quad (3.17)$$

which is a function of $\boldsymbol{\theta}$, $\boldsymbol{\omega}$, and \mathbf{Q} .

Replacing \mathbf{P} in (3.14) with the estimate in (3.17), we can obtain a reduced LL function $L(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{Q})$, which depends on $\boldsymbol{\theta}$, $\boldsymbol{\omega}$, and \mathbf{Q} only. $L(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{Q})$ can hardly be further reduced, except in some special cases (see [20] and [52], and [53] for examples from DOA estimation of stochastic signals). One such special case is DOA estimation under spatially white noise, which has been widely addressed for deterministic and stochastic signals (see [7], [8], [10], [52], and [54] for examples). In the following, we

give the expression of the further reduced LL function under spatially white noise $\mathbf{Q} = \sigma^2 \mathbf{I}$ for DOA estimation of mixed signals.

The LL function in (3.13) has a form similar to the Gaussian LL function for stochastic signals. Thus, following procedures and results similar to those for stochastic signal DOA estimation under spatially white noise [8], [52], we can obtain the ML estimates of σ^2 and \mathbf{P} , along with the further reduced LL function under spatially white noise as

$$\hat{\sigma}^2 = \frac{1}{M-L} \text{trace} \left\{ \mathbf{\Pi}_A^\perp \hat{\mathbf{R}}_{yy} \right\}, \quad (3.18)$$

$$\hat{\mathbf{P}} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \hat{\mathbf{R}}_{yy} \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} - \hat{\sigma}^2 (\mathbf{A}^H \mathbf{A})^{-1}, \quad (3.19)$$

$$L(\boldsymbol{\theta}, \boldsymbol{\omega}) = -\log |\mathbf{A} \hat{\mathbf{P}} \mathbf{A}^H + \hat{\sigma}^2 \mathbf{I}_M|, \quad (3.20)$$

where $\mathbf{\Pi}_A^\perp = \mathbf{I}_M - \mathbf{\Pi}_A$. Note that when $\mathbf{Q} = \sigma^2 \mathbf{I}$, we have

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N} (\mathbf{Y} - \mathbf{\Pi}_A \mathbf{Y} \mathbf{\Pi}_{\phi^H}) (\mathbf{Y} - \mathbf{\Pi}_A \mathbf{Y} \mathbf{\Pi}_{\phi^H})^H. \quad (3.21)$$

We thus are able to simplify (3.18) as

$$\hat{\sigma}^2 = \frac{1}{M-L} \text{trace} \left\{ \mathbf{\Pi}_A^\perp \frac{\mathbf{Y} \mathbf{Y}^H}{N} \right\}. \quad (3.22)$$

Substituting (3.19), (3.21), and (3.22) into (3.20), we have

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\omega}) &= -\log |\mathbf{A} \hat{\mathbf{P}} \mathbf{A}^H + \hat{\sigma}^2 \mathbf{I}_M| \\ &= -\log |\mathbf{\Pi}_A \hat{\mathbf{R}}_{yy} \mathbf{\Pi}_A + \hat{\sigma}^2 \mathbf{\Pi}_A^\perp| \\ &= -\log \left| \frac{\mathbf{\Pi}_A \mathbf{Y} \mathbf{\Pi}_{\phi^H}^\perp \mathbf{Y}^H \mathbf{\Pi}_A}{N} + \hat{\sigma}^2 \mathbf{\Pi}_A^\perp \right|. \end{aligned} \quad (3.23)$$

As a result, the mixed DOA estimator under spatially white noise can be implemented by maximizing (3.23) with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$. When the noise is not spatially white and \mathbf{Q} cannot be reduced, the implementation of the mixed estimator requires the maximization of $L(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{Q})$ over \mathbf{Q} as well as $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ (see [20] and [67] for examples of the maximization over \mathbf{Q} for stochastic signal DOA estimation).

3.4 Analytical Performance Analysis

In this section, we present analytical results on the performances of the stochastic and the mixed estimators. For convenience of formulation, we define the following notations:

$$\mathbf{D} = \left[\frac{d\mathbf{a}(\theta_1)}{d\theta_1}, \frac{d\mathbf{a}(\theta_2)}{d\theta_2}, \dots, \frac{d\mathbf{a}(\theta_L)}{d\theta_L} \right], \quad (3.24)$$

$$\check{\mathbf{P}} = \mathbf{C}\mathbf{C}^H + \mathbf{P}, \quad (3.25)$$

$$\mathbf{Q}'_i = \frac{d\mathbf{Q}}{d\sigma_i}, \quad i = 1, \dots, l_\sigma, \quad (3.26)$$

$$\tilde{\mathbf{Q}}'_i = \mathbf{Q}^{-\frac{1}{2}} \mathbf{Q}'_i \mathbf{Q}^{-\frac{1}{2}}, \quad i = 1, \dots, l_\sigma, \quad (3.27)$$

$$\boldsymbol{\Lambda} = \left[\text{vec} \left\{ \tilde{\mathbf{Q}}'_1 \right\}, \dots, \text{vec} \left\{ \tilde{\mathbf{Q}}'_{l_\sigma} \right\} \right], \quad (3.28)$$

$$\boldsymbol{\Xi} = \left[\text{vec} \left\{ \mathbf{e}_1 \mathbf{e}_1^T \right\}, \dots, \text{vec} \left\{ \mathbf{e}_L \mathbf{e}_L^T \right\} \right], \quad (3.29)$$

where $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \dots, \sigma_{l_\sigma}]^T$ is the vector of length l_σ containing all the real unknown parameters from \mathbf{Q} , and \mathbf{e}_i is a $L \times 1$ vector with the i -th element 1 and all the other elements 0. We also let \mathbf{c} , \mathbf{p} , and $\check{\mathbf{p}}$ be the real column vectors containing the unknown parameters from \mathbf{C} , \mathbf{P} , and $\check{\mathbf{P}}$, respectively. Note that \mathbf{p} consists of the diagonal elements of \mathbf{P} , and the real and imaginary parts of all the upper-triangular

elements above the diagonal in \mathbf{P} . The vector $\check{\mathbf{p}}$ is composed similarly of those from $\check{\mathbf{P}}$.

In the following Proposition 1, we give the CRB on DOA estimation of mixed signals.

Proposition 8. *For DOA estimation of mixed signals, the mixed CRB matrix $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ on DOA can be written as*

$$\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}} = [\mathbf{CRB}_{\mathbf{D},\boldsymbol{\theta}}^{-1} + \mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}^{-1}]^{-1}, \quad (3.30)$$

where $\mathbf{CRB}_{\mathbf{D},\boldsymbol{\theta}}$ is the deterministic CRB matrix on DOA estimation with the measurement at snapshot t , $t = 1, \dots, N$, following $\mathcal{CN}(\mathbf{A}\mathbf{C}\boldsymbol{\varphi}(t), \mathbf{Q})$, in which $\boldsymbol{\theta}$, $\boldsymbol{\omega}$, \mathbf{c} , and $\boldsymbol{\sigma}$ are the unknown parameters, and $\mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}$ is the stochastic CRB matrix on DOA estimation with each of the N measurements following $\mathcal{CN}(\mathbf{0}, \mathbf{R})$, in which $\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}$ and the unknown parameters are $\boldsymbol{\theta}$, \mathbf{p} , and $\boldsymbol{\sigma}$.

Proof: See Appendix H.

The equations for $\mathbf{CRB}_{\mathbf{D},\boldsymbol{\theta}}$ and $\mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}$ have been addressed in [56], [57], and [43].

Though we now have the CRB for DOA estimation of mixed signals, it is not appropriate to simply assume that the mixed ML estimator achieves the CRB asymptotically. Since the measurements are not temporally stationary due to the sinusoidal components, the conclusion that the error covariance matrix of the ML estimator asymptotically achieves the CRB (see [42]), which is typically based on the assumption of independent and identically distributed measurements, can not be directly applied here. Note that a similar concern has been addressed in [50] for ML estimation of sinusoidal frequencies, in which the measurements are not temporally stationary either

and a proof was given to show that the ML frequency estimator achieves the CRB asymptotically.

In the following Proposition 9 and its proof in Appendix I, we confirm the asymptotic efficiency of the mixed estimator.

Proposition 9. *For DOA estimation of mixed signals, the error covariance matrix of the mixed estimator on $\boldsymbol{\alpha} = [\boldsymbol{\theta}^T, \mathbf{p}^T, \boldsymbol{\sigma}^T, \mathbf{c}^T, \boldsymbol{\omega}^T]^T$ achieves the mixed CRB matrix asymptotically.*

The following Theorem 1 (Lindeberg-Feller central limit theorem) [58] is used in our proof of Proposition 2. Note that the setup of our problem and thus the proof of Proposition 2 are absolutely different from those in [50].

Theorem 3. *Let $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_N$ be independent random vectors, each of which are dependent on N , such that as $N \rightarrow \infty$,*

$$\sum_{n=1}^N \int_{\|\boldsymbol{\gamma}_n\| > \epsilon} \|\boldsymbol{\gamma}_n\|^2 f_{\boldsymbol{\gamma}_n}(\boldsymbol{\gamma}_n) d\boldsymbol{\gamma}_n \rightarrow \mathbf{0} \quad (3.31)$$

for arbitrary $\epsilon > 0$ and $\sum_{n=1}^N \text{cov}\{\boldsymbol{\gamma}_n\} \rightarrow \mathbf{Q}$, where $\|\cdot\|$ denotes the Euclidean norm of a vector, $f_{\boldsymbol{\gamma}_n}(\boldsymbol{\gamma}_n)$ is the pdf of $\boldsymbol{\gamma}_n$, and $\text{cov}\{\cdot\}$ denotes the covariance matrix of a random vector. Then $\sum_{n=1}^N (\boldsymbol{\gamma}_n - \mathbb{E}\{\boldsymbol{\gamma}_n\})$, where $\mathbb{E}\{\boldsymbol{\gamma}_n\}$ is the expectation of $\boldsymbol{\gamma}_n$, converges in distribution to a random vector following the normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{Q})$ as $N \rightarrow \infty$.

Now we examine the performance of applying the stochastic estimator to DOA estimation of mixed signals. In Appendix J, we show that the stochastic estimator provides consistent estimate for $\boldsymbol{\xi} = [\boldsymbol{\theta}^T, \boldsymbol{\sigma}^T, \check{\mathbf{p}}^T]^T$ if it is used for mixed signal DOA

estimation, and we give the expression for the asymptotic error covariance matrix of the stochastic estimator on ξ . We also show the following proposition in Appendix J.

Proposition 10. *If the stochastic estimator is applied to DOA estimation of mixed signals, the asymptotic error covariance matrix on $\boldsymbol{\rho} = [\boldsymbol{\theta}^T, \boldsymbol{\sigma}^T]^T$ equals $\mathbf{CRB}_{\check{\mathbf{S}}_\rho}$, which is the stochastic CRB matrix on the estimation of $\boldsymbol{\rho}$ with each of the N measurements following $\mathcal{CN}(\mathbf{0}, \check{\mathbf{R}})$, where $\check{\mathbf{R}} = \mathbf{A}\check{\mathbf{P}}\mathbf{A}^H + \mathbf{Q}$ and the unknown parameters are $\boldsymbol{\theta}$, $\check{\mathbf{p}}$, and $\boldsymbol{\sigma}$. As a result, the asymptotic error covariance matrix on $\boldsymbol{\theta}$ equals $\mathbf{CRB}_{\check{\mathbf{S}}_\theta}$.*

We can see from Proposition 3 that the stochastic and the mixed signals provide similar asymptotic accuracy for the estimation of $\boldsymbol{\rho}$ using the stochastic estimator.

We now compare the performances of the stochastic and the mixed estimators. Let $\mathbf{C}_{\mathbf{M}_\rho}^{\text{as}}$ be the asymptotic error covariance matrix on $\boldsymbol{\rho}$ of the mixed estimator. We show the following proposition in Appendix K.

Proposition 11. *For DOA estimation of mixed signals, $\mathbf{C}_{\mathbf{M}_\rho}^{\text{as}} \leq \mathbf{CRB}_{\check{\mathbf{S}}_\rho}$. As a result, $\mathbf{C}_{\mathbf{M}_\theta}^{\text{as}} \leq \mathbf{CRB}_{\check{\mathbf{S}}_\theta}$.*

Note that $\boldsymbol{\rho}$ consists of all the common unknown parameters that appear in both the mixed and the stochastic LL functions, and can be estimated by both the mixed and the stochastic estimators. We cannot make comparisons of the two estimators for the parameters not included in $\boldsymbol{\rho}$. Compared with the stochastic estimator, the mixed estimator has two extra sets of parameters to estimate: the coefficients and the frequencies of the sinusoidal components. However, from Proposition 11, we see that with the same correlation matrices of signals and noise, the mixed signals and estimator provide higher accuracy than the stochastic ones for common parameter estimation as well as DOA estimation.

3.5 Numerical Examples

In this section, we compare the performances of the stochastic and the mixed estimators through two numerical examples.

In the first example, we consider a DOA estimation problem of two sources under spatially white noise using a uniform linear array of 10 sensors with the inter-sensor distance equal to 0.5λ , where λ is the wavelength of the narrow-band signals. The linear sensor array lies on the z axis, and the elevation angles of the two sources are 15 and 20 degrees, respectively. We assume there are two sinusoidal waves incident on the array, one from each source, with radian frequencies -0.4π and 0.3π and coefficients 0.8 and 0.7, respectively. For the random signal part, we assume the covariance matrix is

$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.2 \\ 0.2 & 0.5 \end{bmatrix}. \quad (3.32)$$

We assume the noise power is $\sigma^2 = 10$, and apply both the stochastic and the mixed estimators to DOA estimation. The optimization for the proposed mixed estimator is implemented iteratively, starting from selected initial values of $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$. The DOA vector $\boldsymbol{\theta}$ is initialized using the result from the stochastic estimator. The radian frequency vector $\boldsymbol{\omega}$ can be initialized using many well-developed techniques for frequency estimation of sinusoidal signals (see [33]- [51] for examples). For instance, we can obtain the initial value for $\boldsymbol{\omega}$ by maximizing the likelihood function based on the following measurement model:

$$\mathbf{y}(t) = \mathbf{G}\boldsymbol{\varphi}(t) + \tilde{\boldsymbol{\epsilon}}(t), \quad (3.33)$$

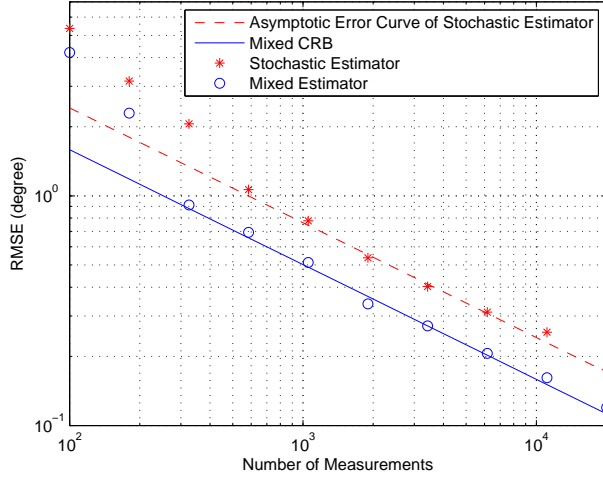


Figure 3.1: Cramér-Rao bound and root-mean-square errors as functions of the number of measurements from mixed signal DOA estimation under spatially white noise using uniform linear array.

where $\mathbf{y}(t)$ is still the measurement at time t , \mathbf{G} is a matrix of the same size as \mathbf{AC} , and $\tilde{\boldsymbol{\epsilon}}(t)$ is the noise assumed to be spatially white. The unknown parameters in (3.33) consist of $\boldsymbol{\omega}$, \mathbf{G} , and the noise power. However, the matrix \mathbf{G} and the noise power can be easily reduced in the likelihood function, which, as a result, is able to be fully concentrated on $\boldsymbol{\omega}$, and can be maximized accurately using ML frequency estimation methods (see [50] and [51]). Efficient and accurate frequency estimation is not within the scope or interest of this dissertation. For simplicity, in this example, we maximize the reduced likelihood function from (3.33) using a uniform grid search to obtain the initial estimate of $\boldsymbol{\omega}$.

The root-mean-square errors (RMSEs) of the two DOA estimators, which were computed with 200 Monte Carlo runs, are illustrated in Fig. 3.1 for the source with $\theta = 15$ degrees.

In our second example, we consider a DOA estimation problem under spatially colored noise using a linear array consisting of separated sub-arrays. In this type of DOA estimation problems (see [21], [24], [67], and [59] for examples), the noise is considered as uncorrelated between the widely separated sub-arrays such that the noise covariance matrix \mathbf{Q} of the whole array has a block-diagonal structure, in which each diagonal block is assumed to be arbitrarily positive definite and unknown. In this example, we examine a DOA estimation problem of two sources using a linear array consisting of four separated sub-arrays, which are composed of 3, 2, 3, and 2 sensors, respectively. The distance between any two adjacent sub-arrays is 3λ . Within each sub-array, the spacing between adjacent sensors is 0.5λ . The linear sensor array lies on the z axis. The elevation angles of the two sources are 15 and 30 degrees. The (k, l) -th element of \mathbf{Q}_i , which is the noise covariance matrix of the i -th sub-array, is generated by [24]

$$[\mathbf{Q}_i]_{kl} = \sigma_i^2 \exp \{ -(k-l)^2 \zeta_i \}, \quad (3.34)$$

with $\sigma_1^2 = 7$, $\sigma_2^2 = 6$, $\sigma_3^2 = 5$, $\sigma_4^2 = 8$, $\zeta_1 = 0.6$, $\zeta_2 = 0.7$, $\zeta_3 = 0.7$, and $\zeta_4 = 0.6$. We assume there are two sinusoidal waves incident on the array, one from each source, with radian frequencies -0.3π and 0.4π and coefficients 0.5 and 0.6, respectively. For the random signal part, we assume the covariance matrix is

$$\mathbf{P} = \begin{bmatrix} 0.2 & 0.0 \\ 0.0 & 0.5 \end{bmatrix}. \quad (3.35)$$

Unlike the case of spatially white noise, the block-diagonal array noise covariance matrix \mathbf{Q} in this example cannot be reduced from $\mathbf{L}(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{Q})$. To obtain the DOA estimate, the iterative optimization needs to be implemented over \mathbf{Q} as well as $\boldsymbol{\theta}$ and

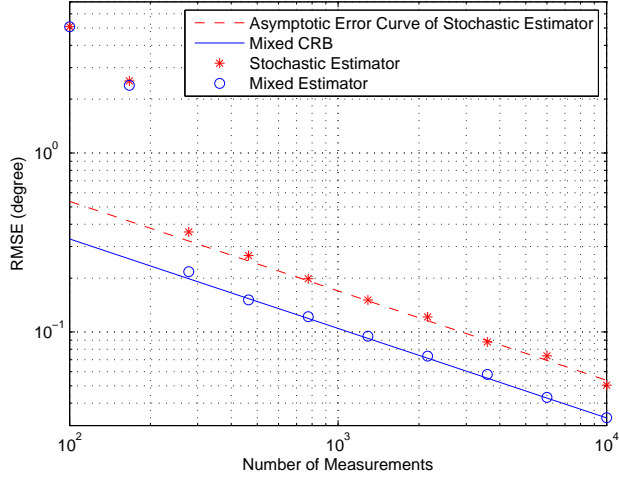


Figure 3.2: Cramér-Rao bound and root-mean-square errors as functions of the number of measurements from mixed signal DOA estimation under spatially colored noise using linear array consisting of separated sub-arrays.

ω . In one iteration of such optimization, we first maximize $L(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{Q})$ over $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ starting from their values obtained in the previous iteration, with \mathbf{Q} fixed at its value from the previous iteration. Then with $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ fixed at their newly updated values, we obtain the improved estimate for \mathbf{Q} , which can be achieved using the Expectation-Maximization method proposed in [67]. We implement such iterations until the Euclidean norm of the change in $\boldsymbol{\theta}$ from two consecutive iterations is smaller than 10^{-2} degree. The DOA estimation errors of the mixed estimator computed from 200 Monte Carlo runs are shown in Fig. 3.2 for the source with $\theta = 15$ degrees. We obtain the initial values of $\boldsymbol{\theta}$ and \mathbf{Q} for the mixed estimator using the stochastic estimator proposed in [67], whose errors calculated from 200 Monte Carlo runs are also shown in Fig. 3.2. We use the same rule as for the mixed estimator to stop the optimization iterations of the stochastic estimator. We initialize $\boldsymbol{\omega}$ for the mixed estimator using the same method as in the first example.

From the results in Figs. 3.1 and 3.2, we can see that since the mixed model describes the signals more accurately, the mixed estimator evidently provides better DOA estimation performance than the stochastic estimator. This matches the result from our analytical performance analysis.

3.6 Summary

We considered the narrow-band DOA estimation problem of hydroacoustic signals from marine vessels containing both sinusoidal and random components, for which we presented a mixed signal model and gave ML estimation results. We derived the asymptotic error covariance matrix on ML DOA estimation of the mixed signals, as well as that of applying the typical stochastic estimator on DOA estimation of the mixed signals. We presented both analytical and numerical comparisons for the performances of the typical stochastic and the proposed mixed estimators for mixed signal DOA estimation. The results showed that the proposed mixed signal model and estimator improve the mixed signal DOA estimation accuracy compared with the typical stochastic ones. In practice, the hydroacoustic signals from marine vessels are usually wide-band. We will address the wide-band DOA estimation of such signals in the following two chapters.

Chapter 4

Direction-of-Arrival Finding of Wide-band Hydroacoustic Signals From Marine Vessels: An Extension of the Narrow-Band Case

In this chapter, we consider the problem of wide-band maximum likelihood (ML) direction-of-arrival (DOA) finding of underwater sources like ships, submarines, or torpedoes, which emit hydroacoustic signals containing sinusoidal waves. These signals, which we call mixed signals, are modeled as the mixture of sinusoidal waves and stochastic Gaussian signals in Chapter 3 for narrow-band DOA estimation. In this chapter, we generalize the narrow-band results in Chapter 3 to the wide-band case, and give the wide-band ML DOA estimator for the mixed signals under spatially white noise. We derive the asymptotic error covariance matrix of the ML estimator,

as well as that of the typical stochastic estimator assuming zero-mean Gaussian signals, for DOA estimation of wide-band mixed signals. Our results demonstrate that compared with the typical stochastic estimator, the proposed ML estimator provides better DOA estimation accuracy for wide-band mixed signals.

4.1 Introduction

Direction-of-arrival (DOA) estimation has important applications in the localization and tracking of underwater acoustic sources like ships and submarines. This chapter develops a wide-band ML method for DOA estimation of hydroacoustic signals from ships, submarines, or torpedoes, which are found to consist of two parts (see [31], [48], and [49]): the noise-like part with continuous spectra, and the tonal or sinusoidal part with discrete frequencies.

In Chapter 3, these signals are called “mixed signals”, and are modeled as the mixture of a zero-mean Gaussian part with unknown covariance matrix and a sinusoidal part with unknown coefficients and frequencies. The work in Chapter 3 only considered the narrow-band DOA estimation of the mixed signals, which, however, are usually wide-band in practice. Therefore, the results developed in Chapter 3 cannot be directly applied to mixed signal DOA estimation in practice. In this chapter, we extend the narrow-band results in Chapter 3 to the wide-band case. We give the wide-band ML DOA estimator for mixed signals and derive its asymptotic error covariance matrix. Additionally, we derive the asymptotic error covariance matrix of applying the typical stochastic estimator, which assumes zero-mean Gaussian signals, to wide-band mixed signal DOA estimation. Our analytical comparison and numerical example show that the proposed ML estimator improves the DOA estimation accuracy for mixed signals.

The remainder of this chapter is organized as follows. We give the wide-band measurement model and ML estimator in Section 4.2. Section 4.3 presents the results of analytical performance analysis. The numerical example and conclusions appear at last in Sections 4.4 and 4.5, respectively. For simplicity of presentation, in the rest of this chapter, we use “stochastic signals” to represent zero-mean Gaussian signals. We use “stochastic” and “mixed” estimators or CRBs (Cramér-Rao bounds) to represent the ML estimators or CRBs derived under the assumptions of “stochastic” and “mixed” signals, respectively.

4.2 Measurement Model And Maximum Likelihood Estimation

In this section, we give the wide-band measurement model and ML estimator for DOA finding of mixed signals.

4.2.1 Measurement Model

Consider the wide-band mixed signals from L far-field sources impinging on an array of M sensors. We decompose the frequency band into a set of non-overlapping (or roughly non-overlapping) sub-bands, each of which is a narrow band. We make measurements from each sub-band following the narrow-band measurement model (see (1) in [67]), and formulate the wide-band measurement model as

$$\mathbf{y}_k(t) = \mathbf{A}_k(\boldsymbol{\theta}) [\mathbf{C}_k \boldsymbol{\varphi}_k(\boldsymbol{\omega}_k, t) + \mathbf{x}_k(t)] + \boldsymbol{\epsilon}_k(t), \quad t = 1, \dots, N_k, \quad k = 1, \dots, K, \quad (4.1)$$

where $\mathbf{y}_k(t)$ is the $M \times 1$ measurement vector at the t -th snapshot from the k -th sub-band, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_L]^T$ is the DOA vector with θ_l the DOA of the l -th source and $\{\cdot\}^T$ denoting the matrix transpose, $\mathbf{A}_k(\boldsymbol{\theta}) = [\mathbf{a}_k(\theta_1), \dots, \mathbf{a}_k(\theta_L)]$ is the array steering matrix with steering vector $\mathbf{a}_k(\theta_l)$ focused on the carrier frequency of the k -th sub-band, $\boldsymbol{\omega}_k = [\omega_{k,1}, \dots, \omega_{k,J_k}]^T$ with J_k the total number of sinusoidal components and $\omega_{k,m}$ the radian frequency of the m -th sinusoidal component in the k -th sub-band, $\boldsymbol{\varphi}_k(\boldsymbol{\omega}_k, t) = [e^{j\omega_{k,1}t}, \dots, e^{j\omega_{k,J_k}t}]^T$ with $j^2 = -1$, \mathbf{C}_k is the $L \times J_k$ coefficient matrix of $\boldsymbol{\varphi}_k(\boldsymbol{\omega}_k, t)$, $\mathbf{x}_k(t)$ is the $L \times 1$ Gaussian signal vector from the k -th sub-band, $\boldsymbol{\epsilon}_k(t)$ is the $M \times 1$ Gaussian noise vector from the k -th sub-band, N_k is the number of measurements from the k -th sub-band, and K is the total number of sub-bands. We assume $\mathbf{x}_k(t)$ and $\boldsymbol{\epsilon}_k(t)$ are temporally white and uncorrelated with each other, and that they follow circular complex Gaussian distributions $\mathcal{CN}(\mathbf{0}, \mathbf{P}_k)$ and $\mathcal{CN}(\mathbf{0}, \mathbf{Q}_k)$, respectively, with unknown covariance matrices \mathbf{P}_k and \mathbf{Q}_k .

In (4.1), the mixed signals are represented in the k -th sub-band by the sum of $\mathbf{C}_k \boldsymbol{\varphi}_k(\boldsymbol{\omega}_k, t)$ and $\mathbf{x}_k(t)$, which correspond to the parts with discrete frequencies and continuous spectra, respectively. We assume the number of sources is known. Also we assume the number of sinusoidal components in each sub-band has been estimated and known, which is a common assumption in existing sinusoidal frequency estimation research (see [33], [50], [51], and [68]). We further assume the signals or noise from different sub-bands are uncorrelated since different sub-bands do not overlap in frequency. The value of N_k may vary for different sub-bands. This is often true when sub-bands with different band widths are sampled at Nyquist rates. However, the ratio N_k/N_m should roughly remain constant at τ_m/τ_k as N_1, \dots, N_K increase, where τ_k is the sampling period of the k -th sub-band. The unknown parameters in the wide-band model (4.1) are those from $\boldsymbol{\theta}$, \mathbf{C}_k , $\boldsymbol{\omega}_k$, \mathbf{P}_k , and \mathbf{Q}_k , $k = 1, \dots, K$.

4.2.2 Maximum Likelihood Estimation

In this sub-section, we present the mixed DOA estimator for wide-band mixed signals. For simplicity, we omit $\boldsymbol{\theta}$ and $\boldsymbol{\omega}_k$ in notations, and use \mathbf{A}_k and $\boldsymbol{\varphi}_k(t)$ to represent $\mathbf{A}_k(\boldsymbol{\theta})$ and $\boldsymbol{\varphi}_k(\boldsymbol{\omega}_k, t)$, respectively.

Recall that in our wide-band model, the measurements are captured separately from each sub-band following the narrow-band measurement model, and recall that the measurements from different sub-bands are uncorrelated. As a consequence, the wide-band log-likelihood (LL) function is the sum of the narrow-band LL functions from all sub-bands, which, by omitting constant factors, can be written as

$$\mathcal{L}(\boldsymbol{\alpha}) = \sum_{k=1}^K N_k \left\{ -\log |\mathbf{A}_k \mathbf{P}_k \mathbf{A}_k^H + \mathbf{Q}_k| - \text{trace} \left\{ [\mathbf{A}_k \mathbf{P}_k \mathbf{A}_k^H + \mathbf{Q}_k]^{-1} \mathbf{R}_{\mathbf{y}_k \mathbf{y}_k} \right\} \right\}, \quad (4.2)$$

where $\boldsymbol{\alpha}$ is the real vector containing all the unknown parameters in the wide-band model, $|\cdot|$ denotes the matrix determinant, and

$$\mathbf{R}_{\mathbf{y}_k \mathbf{y}_k} = \frac{1}{N_k} \sum_{t=1}^{N_k} [\mathbf{y}_k(t) - \mathbf{A}_k \mathbf{C}_k \boldsymbol{\varphi}_k(t)][\mathbf{y}_k(t) - \mathbf{A}_k \mathbf{C}_k \boldsymbol{\varphi}_k(t)]^H. \quad (4.3)$$

Applying the same narrow-band analysis as from equations (2) to (22) in [67] to each sub-band, we can obtain the wide-band ML estimation results as in the following paragraph.

For arbitrary proper noise covariance matrices \mathbf{Q}_k , $k = 1, \dots, K$, the ML estimates for \mathbf{C}_k and \mathbf{P}_k as functions of $\boldsymbol{\theta}$, $\boldsymbol{\omega}_k$, and \mathbf{Q}_k can be found as

$$\hat{\mathbf{C}}_k = \left(\tilde{\mathbf{A}}_k^H \tilde{\mathbf{A}}_k \right)^{-1} \tilde{\mathbf{A}}_k^H \tilde{\mathbf{Y}}_k \boldsymbol{\phi}_k^H (\boldsymbol{\phi}_k \boldsymbol{\phi}_k^H)^{-1}, \quad (4.4)$$

$$\hat{\mathbf{P}}_k = \left[\tilde{\mathbf{A}}_k^H \tilde{\mathbf{A}}_k \right]^{-1} \tilde{\mathbf{A}}_k^H \tilde{\mathbf{R}}_{\mathbf{y}_k \mathbf{y}_k} \tilde{\mathbf{A}}_k \left[\tilde{\mathbf{A}}_k^H \tilde{\mathbf{A}}_k \right]^{-1} - \left[\tilde{\mathbf{A}}_k^H \tilde{\mathbf{A}}_k \right]^{-1}, \quad (4.5)$$

where $\tilde{\mathbf{A}}_k = \mathbf{Q}_k^{-\frac{1}{2}} \mathbf{A}_k$, $\tilde{\mathbf{Y}}_k = \mathbf{Q}_k^{-\frac{1}{2}} \mathbf{Y}_k$ with

$$\mathbf{Y}_k = [\mathbf{y}_k(1), \dots, \mathbf{y}_k(N_k)], \quad (4.6)$$

$$\boldsymbol{\phi}_k = [\boldsymbol{\varphi}_k(1), \dots, \boldsymbol{\varphi}_k(N_k)], \quad (4.7)$$

and

$$\tilde{\mathbf{R}}_{\mathbf{y}_k \mathbf{y}_k} = \frac{1}{N_k} \left(\tilde{\mathbf{Y}}_k - \boldsymbol{\Pi}_{\tilde{\mathbf{A}}_k} \tilde{\mathbf{Y}}_k \boldsymbol{\Pi}_{\boldsymbol{\phi}_k^H} \right) \left(\tilde{\mathbf{Y}}_k - \boldsymbol{\Pi}_{\tilde{\mathbf{A}}_k} \tilde{\mathbf{Y}}_k \boldsymbol{\Pi}_{\boldsymbol{\phi}_k^H} \right)^H \quad (4.8)$$

with $\boldsymbol{\Pi}_{\tilde{\mathbf{A}}_k} = \tilde{\mathbf{A}}_k (\tilde{\mathbf{A}}_k^H \tilde{\mathbf{A}}_k)^{-1} \tilde{\mathbf{A}}_k^H$. Substituting (4.4) and (4.5) into (4.2), we can obtain the reduced LL function dependent on $\boldsymbol{\theta}$, $\boldsymbol{\omega}_k$, and \mathbf{Q}_k , $k = 1, \dots, K$. For the special case of spatially white noise $\mathbf{Q}_k = \sigma_k^2 \mathbf{I}_M$, where \mathbf{I}_M is the $M \times M$ identity matrix, the noise parameters can be further reduced and the LL function becomes

$$\mathcal{L}(\boldsymbol{\theta}, \bar{\boldsymbol{\omega}}) = - \sum_{k=1}^K N_k \log \left| \boldsymbol{\Pi}_{\mathbf{A}_k} \frac{\mathbf{Y}_k \boldsymbol{\Pi}_{\boldsymbol{\phi}_k^H} \mathbf{Y}_k^H}{N_k} \boldsymbol{\Pi}_{\mathbf{A}_k} + \hat{\sigma}_k^2 \boldsymbol{\Pi}_{\mathbf{A}_k}^\perp \right| \quad (4.9)$$

with $\bar{\boldsymbol{\omega}} = [\boldsymbol{\omega}_1^T, \dots, \boldsymbol{\omega}_K^T]^T$, $\boldsymbol{\Pi}_{\mathbf{A}_k}^\perp = \mathbf{I}_M - \boldsymbol{\Pi}_{\mathbf{A}_k}$, and

$$\hat{\sigma}_k^2 = \frac{1}{M-L} \text{trace} \left\{ \boldsymbol{\Pi}_{\mathbf{A}_k}^\perp \frac{\mathbf{Y}_k \mathbf{Y}_k^H}{N_k} \right\}. \quad (4.10)$$

The wide-band mixed DOA estimator under spatially white noise can be achieved by maximizing (4.9) with respect to $\boldsymbol{\theta}$ and $\bar{\boldsymbol{\omega}}$. If noise parameters cannot be reduced from the LL function, the wide-band mixed estimator requires maximizing the LL function over \mathbf{Q}_k , $k = 1, \dots, K$, as well as $\boldsymbol{\theta}$ and $\bar{\boldsymbol{\omega}}$.

4.3 Analytical Performance Analysis

In this section, we present analytical results on the performances of the stochastic and the mixed estimators. For convenience of formulation, we let

$$\mathbf{D}_k = \left[\frac{d\mathbf{a}_k(\theta_1)}{d\theta_1}, \frac{d\mathbf{a}_k(\theta_2)}{d\theta_2}, \dots, \frac{d\mathbf{a}_k(\theta_L)}{d\theta_L} \right], \quad (4.11)$$

$$\check{\mathbf{P}}_k = \mathbf{C}_k \mathbf{C}_k^H + \mathbf{P}_k, \quad (4.12)$$

and \mathbf{c}_k , \mathbf{p}_k , $\check{\mathbf{p}}_k$, and $\boldsymbol{\sigma}_k$ be the real column vectors containing the unknown parameters from \mathbf{C}_k , \mathbf{P}_k , $\check{\mathbf{P}}_k$, and \mathbf{Q}_k , respectively. Note that \mathbf{p}_k consists of the diagonal elements of \mathbf{P}_k , and the real and imaginary parts of all the upper-triangular elements above the diagonal in \mathbf{P}_k . The vector $\check{\mathbf{p}}_k$ is composed similarly of those from $\check{\mathbf{P}}_k$.

In the following Proposition 1, we give the CRB on DOA estimation of mixed signals.

Proposition 12. *For DOA estimation of wide-band mixed signals, the mixed CRB matrix $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ on DOA can be written as*

$$\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}} = \left[\sum_{k=1}^K (\mathbf{CRB}_{\mathbf{D}_k-\boldsymbol{\theta}}^{-1} + \mathbf{CRB}_{\mathbf{S}_k-\boldsymbol{\theta}}^{-1}) \right]^{-1}, \quad (4.13)$$

where $\mathbf{CRB}_{\mathbf{D}_k-\boldsymbol{\theta}}$ is the deterministic CRB matrix on DOA estimation with the measurement at snapshot t , $t = 1, \dots, N_k$, following $\mathcal{CN}(\mathbf{A}_k \mathbf{C}_k \boldsymbol{\varphi}_k(t), \mathbf{Q}_k)$, in which $\boldsymbol{\theta}$, \mathbf{c}_k , $\boldsymbol{\omega}_k$, and $\boldsymbol{\sigma}_k$ are the unknown parameters, and $\mathbf{CRB}_{\mathbf{S}_k-\boldsymbol{\theta}}$ is the stochastic CRB matrix on DOA estimation with each of the N_k measurements from the sensor array following $\mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$, in which $\mathbf{R}_k = \mathbf{A}_k \mathbf{P}_k \mathbf{A}_k^H + \mathbf{Q}_k$ and the unknown parameters are $\boldsymbol{\theta}$, \mathbf{p}_k , and $\boldsymbol{\sigma}_k$.

Proof: See Appendix L.

The expressions of $\mathbf{CRB}_{D_k-\boldsymbol{\theta}}$ and $\mathbf{CRB}_{S_k-\boldsymbol{\theta}}$ can be formulated using the results in [56], [57], and [43]. As stated in [67], it is not appropriate to simply assume the mixed ML estimator achieves the CRB asymptotically, since the measurements are not temporally stationary due to the sinusoidal components. In the following Proposition 13 and its proof in Appendix M, we verify the asymptotic efficiency of the wide-band mixed estimator.

Proposition 13. *For DOA estimation of wide-band mixed signals, the error covariance matrix of the mixed estimator on $\boldsymbol{\alpha}$ achieves the mixed CRB matrix asymptotically.*

We also examine the performance of applying the stochastic estimator to DOA estimation of wide-band mixed signals.

Proposition 14. *If the stochastic estimator is applied to DOA estimation of wide-band mixed signals, the asymptotic error covariance matrix on $\boldsymbol{\rho} = [\boldsymbol{\theta}^T, \boldsymbol{\sigma}_1^T, \dots, \boldsymbol{\sigma}_K^T]^T$ equals $\mathbf{CRB}_{\check{S}-\boldsymbol{\rho}}$, which is the stochastic CRB matrix on the estimation of $\boldsymbol{\rho}$ with each of the N_k measurements from the k -th sub-band, $k = 1, \dots, K$, following $\mathcal{CN}(\mathbf{0}, \check{\mathbf{R}}_k)$, where $\check{\mathbf{R}}_k = \mathbf{A}_k \check{\mathbf{P}}_k \mathbf{A}_k^H + \mathbf{Q}_k$ and the unknown parameters are $\boldsymbol{\theta}$, $\check{\mathbf{p}}_k$, and $\boldsymbol{\sigma}_k$. As a result, the asymptotic error covariance matrix on $\boldsymbol{\theta}$ equals $\mathbf{CRB}_{\check{S}-\boldsymbol{\theta}}$.*

We now compare the performances of the mixed and the stochastic estimators. Let $\mathbf{C}_{M-\boldsymbol{\theta}}^{\text{as}}$ be the asymptotic error covariance matrix on $\boldsymbol{\theta}$ of the mixed estimator. We have the following proposition proved.

Proposition 15. *For DOA estimation of wide-band mixed signals, $\mathbf{C}_{M-\boldsymbol{\theta}}^{\text{as}} \leq \mathbf{CRB}_{\check{S}-\boldsymbol{\theta}}$.*

Proof: According to Propositions 1 and 2, and the results in their proofs, we have

$$\mathbf{C}_{M-\boldsymbol{\theta}}^{\text{as}} = \left(\sum_{k=1}^K \mathbf{C}_{M_k-\boldsymbol{\theta}}^{\text{as}-1} \right)^{-1}, \text{ where } \mathbf{C}_{M_k-\boldsymbol{\theta}}^{\text{as}-1} = \mathbf{CRB}_{D_k-\boldsymbol{\theta}}^{\text{as}-1} + \mathbf{CRB}_{S_k-\boldsymbol{\theta}}^{-1} \text{ and } \mathbf{CRB}_{D_k-\boldsymbol{\theta}}^{\text{as}}$$

the asymptotic value of $\mathbf{CRB}_{D_k-\boldsymbol{\theta}}$ that can be obtained by replacing $\hat{\mathbf{R}}_{\phi_k\phi_k}$ in (L.11) with its asymptotic value, which is an identity matrix. From Proposition 4 in [67], we have $\mathbf{C}_{M_k-\boldsymbol{\theta}}^{\text{as}} \leq \mathbf{CRB}_{\xi_k-\boldsymbol{\theta}}$, $k = 1, \dots, K$. We thus have the proposition holds.

From Proposition 15, we see that with the same correlation matrices of signals and noise, the mixed signals and estimator provide better DOA estimation than the stochastic ones.

4.4 Numerical Example

In this numerical example, we consider a wide-band DOA estimation problem of two sources under spatially white noise using a uniform linear array consisting of 10 sensors. The linear sensor array lies on the z axis, and the elevation angles of the two sources are 15 and 20 degrees, respectively. The wide frequency band of the signals is composed of three narrow sub-bands, and the inter-sensor distance $d = 0.5\lambda_1 = 0.4\lambda_2 = 0.3\lambda_3$, where λ_1 , λ_2 , and λ_3 are the wavelengths of the narrow-band signals from three sub-bands. In each sub-band, there are two sinusoidal waves incident on the array, one from each source, with radian frequencies -0.4π and 0.3π and coefficients 0.8 and 0.7, respectively. For the random signal part and noise, we assume

$$\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \begin{bmatrix} 0.4 & 0.2 \\ 0.2 & 0.5 \end{bmatrix} \quad (4.14)$$

and $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 10$. In addition, we sample the same number of measurements from each sub-band.

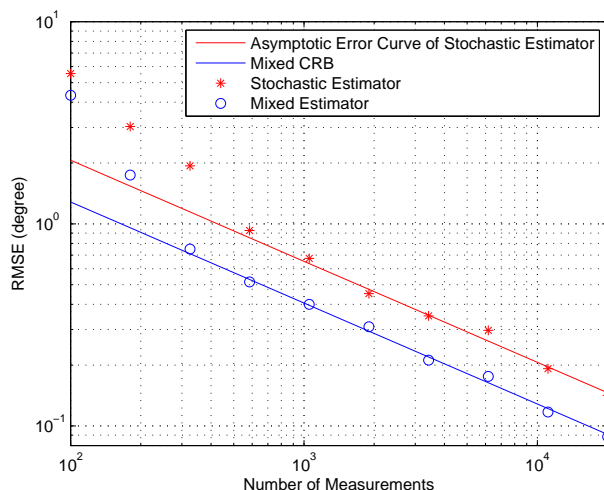


Figure 4.1: Cramér-Rao bound and root-mean-square errors as functions of the number of measurements from wide-band mixed signal DOA estimation under spatially white noise.

We apply both the stochastic and the mixed estimators on DOA estimation. The optimization for the mixed estimator is implemented iteratively starting from selected initial values of θ and $\bar{\omega}$. The DOA is initialized using the result from the stochastic estimator. The initial radian frequency estimates are obtained in each sub-bands using the method given in the first numerical example in [67]. The performances of two estimators are illustrated in Fig. 4.1 for the source with $\theta = 15$ degrees. In Fig. 4.1, the root-mean-square errors (RMSEs) are calculated with 200 Monte Carlo runs.

From Fig. 4.1, we observe that the mixed estimator provides better DOA estimation performance than the stochastic one. This matches the result of our analytical performance analysis.

4.5 Summary

We considered the wide-band DOA estimation problem of the hydroacoustic signals from marine vessel sources. We gave the ML DOA estimator. We computed the asymptotic error covariance matrices of the proposed ML and the typical stochastic estimators for DOA finding of such signals. Our analytical and numerical comparisons showed that the proposed ML estimator improves the DOA estimation accuracy of such signals compared with the typical stochastic estimator.

Chapter 5

Direction-of-Arrival Estimation of Hydroacoustic Signals From Marine Vessels: An Approach From Fourier Transform

In Chapter 4, we considered the wide-band direction-of-arrival (DOA) estimation problem for the hydroacoustic signals from marine vessels. In this chapter, we reconsider this problem following a Fourier-transform-based approach, which is a quite useful and common technique widely adopted in wide-band DOA estimation (see [54], [69], and [70] for examples). We set up new wide-band measurement models from the Fourier transform (FT) of the hydroacoustic signals from marine vessels, based on which we give the maximum likelihood (ML) DOA estimator. We compute and compare the asymptotic error covariance matrices of the proposed estimator, and that of the typical stochastic estimator assuming zero-mean Gaussian signals. The results show that the proposed estimator provides better DOA estimation than the typical stochastic estimator.

The remainder of this chapter is presented as follows. The wide-band measurement models and estimator are presented in Section 5.1. Section 5.2 gives analytical performance analysis. The numerical example and conclusions appear in Sections 5.3 and 5.4, respectively. For simplicity, we use “mixed signals” to represent the wide-band hydroacoustic signals containing both random and sinusoidal components.

5.1 Measurement Models And DOA Estimation

In this section, we present the measurement models and DOA estimator based on the Fourier decomposition of the mixed signals.

5.1.1 Measurement Model

Assume L mixed signals from the far field incident on an array of M sensors. We write the L incident signals at the t -th snapshot as

$$\bar{\mathbf{s}}(t) = \mathbf{C}\boldsymbol{\varphi}(\boldsymbol{\nu}, t) + \bar{\mathbf{x}}(t), \quad t = 1, \dots, N, \quad (5.1)$$

where $\bar{\mathbf{s}}(t) = [\bar{s}_1(t), \dots, \bar{s}_L(t)]^T$ with $\bar{s}_l(t)$ the signal value of the l -th source and $\{\cdot\}^T$ denoting the matrix transpose, $\boldsymbol{\varphi}(\boldsymbol{\nu}, t) = [e^{j\nu_1 t}, \dots, e^{j\nu_K t}]^T$ contains the K sinusoids from all sources with $j^2 = -1$ and $\boldsymbol{\nu} = [\nu_1, \dots, \nu_K]^T$, ν_k is the frequency of the k -th sinusoid, \mathbf{C} is an $L \times K$ matrix containing the coefficients of $\boldsymbol{\varphi}(\boldsymbol{\nu}, t)$, and $\bar{\mathbf{x}}(t) = [\bar{x}_1(t), \dots, \bar{x}_L(t)]^T$ is the random part of the signals, which is assumed to zero-mean complex Gaussian. Here the mixed signals are represented in (5.1) as the sum of the sinusoidal part $\mathbf{C}\boldsymbol{\varphi}(\boldsymbol{\nu}, t)$ and the random part $\bar{\mathbf{x}}(t)$.

To apply Fourier decomposition to the mixed signals, we divide the total N snapshots of signals into T_d subintervals, each of which consists of N_d snapshots, i.e., $N = N_d T_d$. For each subinterval, we implement a discrete Fourier transform (DFT) and obtain

$$\mathbf{s}_n(i) = \mathbf{C}\boldsymbol{\phi}_n(\boldsymbol{\nu}, i) + \mathbf{x}_n(i), \quad i = 1, \dots, T_d, \quad n = 1, \dots, N_d, \quad (5.2)$$

where $\mathbf{s}_n(i)$ and $\mathbf{x}_n(i)$ are $L \times 1$ vectors containing the DFT coefficients of $\bar{\mathbf{s}}(t)$ and $\bar{\mathbf{x}}(t)$, respectively, from the i -th subinterval and n -th Fourier basis function, and $\boldsymbol{\phi}_n(\boldsymbol{\nu}, i) = [\phi_n(\nu_1, i), \dots, \phi_n(\nu_K, i)]^T$ with

$$\phi_n(\nu_k, i) = \frac{1}{\sqrt{N_d}} \sum_{t=1}^{N_d} e^{j\nu_k[(i-1)N_d+t]} e^{-j\zeta_n(t-1)} = h_{k,n} e^{jN_d\nu_k i}, \quad (5.3)$$

where

$$h_{k,n} = \frac{e^{j(\zeta_n - N_d\nu_k)} - e^{j(1-N_d)\zeta_n}}{\sqrt{N_d}(e^{-j(\nu_k - \zeta_n)} - 1)}, \quad (5.4)$$

and ζ_n is the frequency of the n -th Fourier basis function. Therefore, we can write the DFT-based measurement model as

$$\mathbf{y}_n(i) = \mathbf{A}_n(\boldsymbol{\theta}) [\mathbf{C}\boldsymbol{\phi}_n(\boldsymbol{\nu}, i) + \mathbf{x}_n(i)] + \boldsymbol{\epsilon}_n(i), \quad i = 1, \dots, T_d, \quad n = 1, \dots, N_d, \quad (5.5)$$

where $\mathbf{y}_n(i)$ and $\boldsymbol{\epsilon}_n(i)$ are $M \times 1$ vectors, which respectively contain the array output and array noise DFT coefficients from the i -th subinterval and n -th Fourier basis function, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_L]^T$ with θ_l the DOA of the l -th source, $\mathbf{A}_n(\boldsymbol{\theta}) = [\mathbf{a}_n(\theta_1), \dots, \mathbf{a}_n(\theta_L)]$ is the array steering matrix with the $M \times 1$ steering vector $\mathbf{a}_n(\theta_l)$ focused on frequency ζ_n , $l = 1, \dots, L$. We assume the number of sources is known. Also we assume the number of sinusoidal components in each sub-band is known (see [33] and [68] for similar assumptions). We assume $\mathbf{x}_n(i)$ and $\boldsymbol{\epsilon}_n(i)$ are temporally

white and uncorrelated with each other, and follow circular complex Gaussian distributions $\mathcal{CN}(\mathbf{0}, \mathbf{P}_n)$ and $\mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_M)$, respectively, where \mathbf{I}_M is an $M \times M$ identity matrix. We further assume the DFT coefficients from different Fourier basis functions are independent [54]. The unknown parameters in (5.5) are $\boldsymbol{\theta}$, \mathbf{C} , $\boldsymbol{\nu}$, \mathbf{P}_n , and σ_n , $n = 1, \dots, N_d$. For simplicity, we omit $\boldsymbol{\theta}$ and $\boldsymbol{\nu}$ is the following notations.

5.1.2 DOA Estimation From A Modified Model

We first consider the ML estimation based on the model in (5.5), for which we rewrite (5.5) as

$$\mathbf{Y}_n = \mathbf{A}_n \mathbf{C} \boldsymbol{\phi}_n + \mathbf{A}_n \mathbf{X}_n + \mathbf{E}_n, \quad n = 1, \dots, N_d, \quad (5.6)$$

where $\mathbf{Y}_n = [\mathbf{y}_n(1), \dots, \mathbf{y}_n(T_d)]$, $\boldsymbol{\phi}_n = [\phi_n(1), \dots, \phi_n(T_d)]$, $\mathbf{X}_n = [\mathbf{x}_n(1), \dots, \mathbf{x}_n(T_d)]$, and $\mathbf{E}_n = [\boldsymbol{\epsilon}_n(1), \dots, \boldsymbol{\epsilon}_n(T_d)]$. Noting that $\text{vec}\{\mathbf{ABC}\} = [\mathbf{C}^T \otimes \mathbf{A}] \text{vec}\{\mathbf{B}\}$ for any matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} that can produce \mathbf{ABC} , we have

$$\mathbf{y}_n = [\boldsymbol{\phi}_n^T \otimes \mathbf{A}_n] \text{vec}\{\mathbf{C}\} + \text{vec}\{\mathbf{A}_n \mathbf{X}_n + \mathbf{E}_n\}, \quad n = 1, \dots, N_d, \quad (5.7)$$

where $\mathbf{y}_n = \text{vec}\{\mathbf{Y}_n\} = [\mathbf{y}_n^T(1), \dots, \mathbf{y}_n^T(T_d)]^T$, \otimes denotes the Kronecker product, and $\text{vec}\{\cdot\}$ is the vectorization operator stacking all the columns of a matrix, one below another, into a vector. Stacking the coefficients from all Fourier basis functions together, we have

$$\mathbf{y} = \mathbf{A}_\phi \mathbf{c} + \mathbf{n}, \quad (5.8)$$

where $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_{N_d}^T]^T$, $\mathbf{A}_\phi = [(\boldsymbol{\phi}_1^T \otimes \mathbf{A}_1)^T, \dots, (\boldsymbol{\phi}_{N_d}^T \otimes \mathbf{A}_{N_d})^T]^T$, $\mathbf{c} = \text{vec}\{\mathbf{C}\}$, and $\mathbf{n} = [\text{vec}\{\mathbf{A}_1 \mathbf{X}_1 + \mathbf{E}_1\}^T, \dots, \text{vec}\{\mathbf{A}_{N_d} \mathbf{X}_{N_d} + \mathbf{E}_{N_d}\}^T]^T$.

Therefore, we have $\mathbf{y} \sim \mathcal{CN}(\mathbf{A}_\phi \mathbf{c}, \Sigma)$, where $\Sigma = \text{blkdiag}\{\mathbf{I}_{T_d} \otimes (\mathbf{A}_n \mathbf{P}_n \mathbf{A}_n^H + \mathbf{Q}_n)\}_{N_d}$, which is a block-diagonal matrix with the n -th diagonal block equal to $\mathbf{I}_{T_d} \otimes (\mathbf{A}_n \mathbf{P}_n \mathbf{A}_n^H + \mathbf{Q}_n)$, $n = 1, \dots, N_d$. Thus, the ML estimate of \mathbf{c} as a function of $\boldsymbol{\theta}$, $\boldsymbol{\nu}$, σ_n , and \mathbf{P}_n , $n = 1, \dots, N_d$ is

$$\begin{aligned} \hat{\mathbf{c}} &= (\mathbf{A}_\phi^H \Sigma^{-1} \mathbf{A}_\phi)^{-1} \mathbf{A}_\phi^H \Sigma^{-1} \mathbf{y} \\ &= \left\{ \sum_{n=1}^{N_d} (\boldsymbol{\phi}_n^* \boldsymbol{\phi}_n^T) \otimes [\mathbf{A}_n^H (\mathbf{A}_n \mathbf{P}_n \mathbf{A}_n^H + \mathbf{Q}_n)^{-1} \mathbf{A}_n] \right\}^{-1} \\ &\quad \times \sum_{n=1}^{N_d} \left\{ \boldsymbol{\phi}_n^* \otimes [\mathbf{A}_n^H (\mathbf{A}_n \mathbf{P}_n \mathbf{A}_n^H + \mathbf{Q}_n)^{-1}] \right\} \mathbf{y}_n, \end{aligned} \quad (5.9)$$

where $\{\cdot\}^*$ denotes the complex conjugate.

In addition, we have the log-likelihood (LL) function from (5.6) as

$$\mathcal{L}(\boldsymbol{\alpha}) = \sum_{n=1}^{N_d} T_d \left\{ -\log |\mathbf{A}_n \mathbf{P}_n \mathbf{A}_n^H + \mathbf{Q}_n| - \text{trace} \left\{ [\mathbf{A}_n \mathbf{P}_n \mathbf{A}_n^H + \mathbf{Q}_n]^{-1} \mathbf{R}_{\mathbf{y}_n \mathbf{y}_n} \right\} \right\}, \quad (5.10)$$

where $\boldsymbol{\alpha}$ is the real vector containing all the unknown parameters in (5.5), $|\cdot|$ denotes the matrix determinant, $\text{trace}\{\cdot\}$ denotes the matrix trace, and

$$\mathbf{R}_{\mathbf{y}_n \mathbf{y}_n} = \frac{1}{T_d} \sum_{t=1}^{T_d} [\mathbf{y}_n(t) - \mathbf{A}_n \mathbf{C} \boldsymbol{\phi}_n(t)] [\mathbf{y}_n(t) - \mathbf{A}_n \mathbf{C} \boldsymbol{\phi}_n(t)]^H. \quad (5.11)$$

Substituting (5.9) back into From (5.10), we obtain a reduce LL function dependent on $\boldsymbol{\theta}$, $\boldsymbol{\nu}$, as well as σ_n and \mathbf{P}_n , $n = 1, \dots, N_d$. To the best of our knowledge, the nuisance parameters \mathbf{P}_n , $n = 1, \dots, N_d$, cannot be further reduced from the LL function. Alternatively, we can first find the ML estimates of \mathbf{P}_n as a function of $\boldsymbol{\theta}$, $\boldsymbol{\nu}$, \mathbf{C} , and σ_n , $n = 1, \dots, N_d$, following the narrow-band estimation result in [67]. Substituting these estimates back into the LL function, we obtain another reduced

LL function dependent on $\boldsymbol{\theta}$, $\boldsymbol{\nu}$, \mathbf{C} , and σ_n , $n = 1, \dots, N_d$. Again, the nuisance parameters in \mathbf{C} cannot be further reduced from the LL function.

To seek further reduction of nuisance parameters, we proposed the modified model based on that in (5.5) as follows.

$$\mathbf{y}_n(i) = \mathbf{A}_n(\boldsymbol{\theta}) [\mathbf{C}_n \boldsymbol{\varphi}(\boldsymbol{\omega}, i) + \mathbf{x}_n(i)] + \boldsymbol{\epsilon}_n(i), \quad i = 1, \dots, T_d, \quad n = 1, \dots, N_d, \quad (5.12)$$

where $\mathbf{C}_n = \mathbf{C} \mathbf{H}_n$, $\mathbf{H}_n = \text{diag}\{h_{1,n}, \dots, h_{K,n}\}$, which is a diagonal matrix with diagonal elements $h_{1,n}, \dots, h_{K,n}$, and $\boldsymbol{\varphi}(\boldsymbol{\omega}, i) = [\varphi(\omega_1, i), \dots, \varphi(\omega_K, i)]^T$ with $\varphi(\omega_k, i) = e^{j\omega_k i}$ and ω_k being the remainder of $N_d \nu_k / (2\pi)$. The unknown parameters in (5.12) consist of $\boldsymbol{\theta}$, $\boldsymbol{\omega}$, \mathbf{C}_n , \mathbf{P}_n , and σ_n , $n = 1, \dots, N_d$. Note that $\mathbf{C}_1, \dots, \mathbf{C}_{N_d}$ are considered as arbitrarily unknown here. So the modified model in (5.12) has more unknowns than the original model in (5.5). For simplicity of formulation, we omit $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ in the following notations.

Note that the LL function from (5.12), which can be readily obtained using (5.10), is the sum of N_d narrow-band LL functions, each of which is of the same form as the narrow-band LL function in [67]. Thus, the ML estimates of \mathbf{C}_n , \mathbf{P}_n , and σ_n for (5.12) can be found as functions of $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ following the similar procedures in [67] as

$$\hat{\mathbf{C}}_n = (\mathbf{A}_n^H \mathbf{A}_n)^{-1} \mathbf{A}_n^H \mathbf{Y}_n \boldsymbol{\varphi}^H (\boldsymbol{\varphi} \boldsymbol{\varphi}^H)^{-1}, \quad (5.13)$$

$$\hat{\mathbf{P}}_n = (\mathbf{A}_n^H \mathbf{A}_n)^{-1} \mathbf{A}_n^H \hat{\mathbf{R}}_{\mathbf{y}_n \mathbf{y}_n} \mathbf{A}_n (\mathbf{A}_n^H \mathbf{A}_n)^{-1} \hat{\sigma}_n^2 (\mathbf{A}_n^H \mathbf{A}_n)^{-1}, \quad (5.14)$$

$$\hat{\sigma}_n^2 = \frac{1}{M - L} \text{trace} \left\{ \boldsymbol{\Pi}_{\mathbf{A}_n}^\perp \frac{\mathbf{Y}_n \mathbf{Y}_n^H}{T_d} \right\}, \quad (5.15)$$

where

$$\hat{\mathbf{R}}_{\mathbf{y}_n \mathbf{y}_n} = \frac{1}{T_d} (\mathbf{Y}_n - \mathbf{\Pi}_{\mathbf{A}_n} \mathbf{Y}_n \mathbf{\Pi}_{\varphi^H}) (\mathbf{Y}_n - \mathbf{\Pi}_{\mathbf{A}_n} \mathbf{Y}_n \mathbf{\Pi}_{\varphi^H})^H \quad (5.16)$$

with $\mathbf{\Pi}_{\mathbf{A}_n}^\perp = \mathbf{I}_M - \mathbf{\Pi}_{\mathbf{A}_n}$ and $\mathbf{\Pi}_{\mathbf{A}_n} = \mathbf{A}_n (\mathbf{A}_n^H \mathbf{A}_n)^{-1} \mathbf{A}_n^H$. Substituting the ML estimates in (5.13)-(5.15) into the LL function from (5.12) and omitting constant terms, we obtain the reduced LL function dependent on $\boldsymbol{\theta}$ and $\boldsymbol{\omega} = [\omega_1^T, \dots, \omega_K^T]^T$ as

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega}) = - \sum_{n=1}^{N_d} T_d \log \left| \mathbf{\Pi}_{\mathbf{A}_n} \frac{\mathbf{Y}_n \mathbf{\Pi}_{\varphi^H}^\perp \mathbf{Y}_n^H}{T_d} \mathbf{\Pi}_{\mathbf{A}_n} + \hat{\sigma}_n^2 \mathbf{\Pi}_{\mathbf{A}_n}^\perp \right|. \quad (5.17)$$

Remark: The maximization of the LL functions are usually implemented through an iterative optimization method. Compared with the LL function in (5.17), the reduced LL functions from the original model in (5.5) always have nuisance signal parameters irreducible. This definitely introduces more optimization burden and less estimation efficiency. Moreover, to guarantee a correct convergence, the iterative optimization ought to start from initial parameter estimates close to the true parameter values. However, we have not found an efficient method at present to initialize the frequencies $\boldsymbol{\nu}$ and the nuisance signal parameter \mathbf{C} or \mathbf{P}_n . Due to the multiplication factor N_d in the exponent in (5.3), the initialization errors in $\boldsymbol{\nu}$ can be significantly enlarged in the LL function. This introduces additional challenge in initializing the nuisance signal parameters \mathbf{C} or \mathbf{P}_n . In contrast, the LL function in (5.17) from the modified model does not contain or require optimization over the nuisance signal parameters, and it only needs the initialization of $\boldsymbol{\theta}$ and the remainder of $N_d \nu_k / (2\pi)$ for iterative optimization, which can be readily obtained from existing techniques. So the ML DOA estimator based on the modified model is definitely more computationally efficient.

Though the ML estimator based on the original model in (5.5) may be able to give better performance than that based on the modified model in (5.12), its efficient implementation remains an unsolved problem to us mainly due to the initialization problem of the frequencies and nuisance signal parameters. Moreover, the derivations of the Cramér-Rao bound (CRB) and the asymptotic error covariance matrix of the ML estimator based on the original model seem quite challenging, and we have not been able to make the analytical performance analysis for the ML estimator based on the original model. So the ML estimator based on the original model will be left open as a topic of our future work, and will not be examined in this correspondence. Instead, we will concentrate on the performance analysis of the ML estimator from the modified model in (5.12) in the following sections. We will show that the ML DOA estimator from the modified model still provides much better estimation performance than the typical ML estimator assuming zero-mean Gaussian signals.

5.2 Analytical Performance Analysis

In this section, we present analytical results on the performances of different estimators. For simplicity, we use “stochastic” estimator and CRB to represent the ML estimator and CRB derived under the assumption of zero-mean Gaussian signals, respectively (see [8] for examples). We use “mixed” estimator and CRB to represent the ML estimator and CRB derived from the modified model in (5.12). Also, we define $\check{\mathbf{P}}_n = \mathbf{C}_n \mathbf{C}_n^H + \mathbf{P}_n$, and \mathbf{c}_n , \mathbf{p}_n , and $\check{\mathbf{p}}_n$ to be the real column vectors containing the unknown parameters from \mathbf{C}_n , \mathbf{P}_n , and $\check{\mathbf{P}}_n$ respectively. Note that \mathbf{p}_n consists of the diagonal elements of \mathbf{P}_n , and the real and imaginary parts of all the upper-triangular

elements above the diagonal in \mathbf{P}_n . The vector $\check{\mathbf{p}}_n$ is composed similarly of those from $\check{\mathbf{P}}_n$. In Appendix O, we prove the following proposition.

Proposition 16. *For DOA estimation based on the model in (5.12), the mixed CRB matrix $\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}$ on DOA can be expressed as*

$$\mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}} = \left[\sum_{n=1}^{N_d} (\mathbf{CRB}_{\mathbf{D}_n,\boldsymbol{\theta}}^{-1} + \mathbf{CRB}_{\mathbf{S}_n,\boldsymbol{\theta}}^{-1}) \right]^{-1}, \quad (5.18)$$

where $\mathbf{CRB}_{\mathbf{D}_n,\boldsymbol{\theta}}$ is the CRB matrix on DOA estimation with the measurement at snapshot t , $t = 1, \dots, T_d$, following $\mathcal{CN}(\mathbf{A}_n \mathbf{C}_n \boldsymbol{\varphi}(t), \mathbf{Q}_n)$, in which $\boldsymbol{\theta}$, \mathbf{c}_n , $\boldsymbol{\omega}$, and σ_n are the unknown parameters, and $\mathbf{CRB}_{\mathbf{S}_n,\boldsymbol{\theta}}$ is the stochastic CRB matrix on DOA estimation with each of the T_d measurements following $\mathcal{CN}(\mathbf{0}, \mathbf{R}_n)$, in which $\mathbf{R}_n = \mathbf{A}_n \mathbf{P}_n \mathbf{A}_n^H + \mathbf{Q}_n$ and the unknown parameters are $\boldsymbol{\theta}$, \mathbf{p}_n , and σ_n .

The expressions for $\mathbf{CRB}_{\mathbf{S}_n,\boldsymbol{\theta}}$ and $\mathbf{CRB}_{\mathbf{D}_n,\boldsymbol{\theta}}$ can be found in [8]-[57]. As explained in [67], since the measurements are not temporally stationary, we cannot take it for granted that the mixed estimator achieves the mixed CRB asymptotically. However, it is shown in [67] for the narrow-band case that the mixed estimator asymptotically achieves the CRB. Since our LL function from the modified model in (5.12) is simply the sum of a set of narrow-band LL functions, each of which is in the same form as the narrow-band LL function in [67], we can easily show following the narrow-band results in [67] that the error covariance matrix of the mixed estimator achieves the mixed CRB matrix asymptotically for the wide-band case in (5.12). Similarly, following the narrow-band results in [67] (see Propositions 3 and 4 in [67] and their proofs), we can show the following two corollaries.

Corollary 1. *If the stochastic estimator is applied to mixed signal DOA estimation, the asymptotic error covariance matrix on $\boldsymbol{\theta}$ is equal to $\mathbf{CRB}_{\check{\mathbf{s}},\boldsymbol{\theta}}$, which is the stochastic CRB matrix on DOA estimation with each of the T_d measurements from the n -th sub-band, $n = 1, \dots, N_d$, following $\mathcal{CN}(\mathbf{0}, \check{\mathbf{R}}_n)$, where $\check{\mathbf{R}}_n = \mathbf{A}_n \check{\mathbf{P}}_n \mathbf{A}_n^H + \mathbf{Q}_n$ and the unknown parameters are $\boldsymbol{\theta}$, $\check{\mathbf{p}}_n$, and σ_n .*

Corollary 2. *For DOA estimation of the mixed signals, $\mathbf{C}_{\mathbf{M},\boldsymbol{\theta}}^{\text{as}} \leq \mathbf{CRB}_{\check{\mathbf{s}},\boldsymbol{\theta}}$, where $\mathbf{C}_{\mathbf{M},\boldsymbol{\theta}}^{\text{as}}$ denotes the asymptotic error covariance matrix on $\boldsymbol{\theta}$ of the mixed estimator.*

Corollary 1 examines the performance of the stochastic estimator when it is used for DOA estimation of the mixed signals. Corollary 2 shows that the mixed estimator provides higher accuracy than the stochastic estimator for mixed signal DOA estimation.

5.3 Numerical Example

In this numerical example, we consider a DOA estimation problem under spatially white noise using a uniform linear array of 10 sensors. The array lies on the z axis, and the elevation angles of the two sources are 15 and 20 degrees, respectively. The wide-band measurement model is constructed with 8 Fourier basis functions ($N_d = 8$). The inter-sensor distance $d = 0.5\lambda_1 = 0.45\lambda_2 = 0.4\lambda_3 = 0.35\lambda_4 = 0.3\lambda_5 = 0.25\lambda_6 = 0.2\lambda_7 = 0.15\lambda_8$, where $\lambda_1, \dots, \lambda_8$ are the wavelengths of eight Fourier basis functions. There are two sinusoidal waves incident on the array, one from each source, with frequencies 0.6π and 1.2π and coefficients 0.6 and 0.7, respectively. For the random

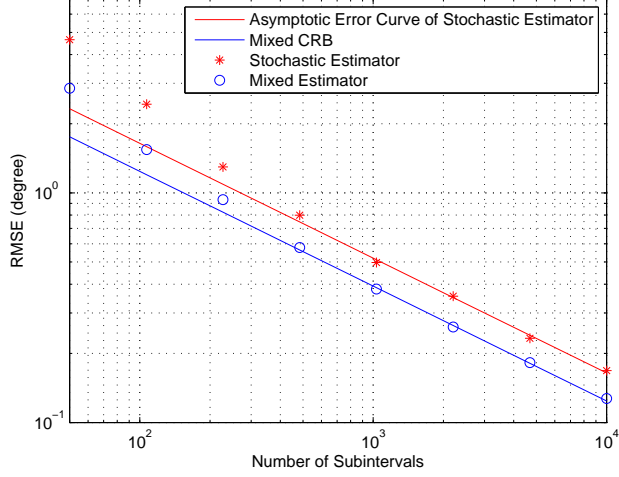


Figure 5.1: Cramér-Rao bound and root-mean-square errors as functions of the number of subintervals N_d for mixed signal DOA estimation.

signal part and noise, we assume

$$\mathbf{P}_1 = \mathbf{P}_2 = \dots = \mathbf{P}_8 = \begin{bmatrix} 0.4 & 0.2 \\ 0.2 & 0.5 \end{bmatrix} \quad (5.19)$$

and $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_8^2 = 10$.

Both the stochastic and the mixed estimators are applied to DOA estimation. The maximization of the LL function in (5.17) is implemented iteratively starting from initial estimates of $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$. The initial DOA estimate is selected using the estimate from the stochastic estimator. The initial frequency estimates are obtained using the method given in the first numerical example in [67]. The performances of two estimators are illustrated in Fig. 5.1 for the source with $\theta = 15$ degrees, with the root-mean-square errors (RMSEs) calculated with 200 Monte Carlo runs.

We can see in Fig. 5.1 that the mixed estimator provides higher estimation accuracy than the stochastic estimator. This matches our analytical performance analysis in Section III.

5.4 Summary

We considered the DOA estimation problem of the hydroacoustic signals from marine vessels. We derived the wide-band measurement model based on Fourier decomposition, from which we further proposed a modified measurement model and its ML DOA estimator that greatly simplify the estimation procedure. Both our analytical and numerical results demonstrated that in comparison with the typical stochastic estimator, the proposed estimator improves the DOA estimation accuracy for the hydroacoustic signals considered.

Chapter 6

A Barankin-Type Bound on Direction-of-Arrival Estimation ³

From the numerical examples and figures in the previous chapters, we can see that the the Cramér-Rao bound (CRB) is able to predict the performance of the maximum likelihood (ML) estimator only when the number of measurements is large. The similar phenomenon occurs to the performance of the ML estimator with respect to the signal-to-noise ratio (SNR). The CRB can predict the performance of the ML estimator only at high SNR values. When the SNR is low, the mean-square errors of the ML estimator deviate from the CRB. The SNR region below which the performance of the ML estimator deviates rapidly from the CRB is called the SNR threshold region. The CRB is not able to provide any information about this threshold region.

In this chapter, we derive a Barankin-type bound (BTB) on the mean-square error (MSE) in estimating the directions of arrivals (DOAs) of far-field sources using acoustic sensor arrays. We consider narrow-band and wide-band deterministic source

³Based on T. Li, J. Tabrikian, and A. Nehorai, "A Barankin-Type Bound on Direction Estimation Using Acoustic Sensor Arrays," *IEEE Trans. Signal Process.*, vol. 59, pp. 431-435, Jan. 2011. ©[2011] IEEE.

signals, and scalar or vector sensors. Our results provide an approximation to the threshold of the SNR below which the performance of the maximum likelihood estimation (MLE) degrades rapidly. For narrow-band DOA estimation using uniform linear vector-sensor arrays, we show that this threshold increases with the distance between the sensors. As a result, for medium SNR values the performance does not necessarily improve with this distance.

6.1 Introduction

The Barankin bound is the tightest lower bound on the mean square errors of any unbiased estimator [71]–[75], and it is a useful tool in estimation problems for predicting the threshold region of SNR [76]–[79], below which the accuracy of the MLE degrades rapidly. Identification of the threshold region enables to determine the operation conditions, such as observation time and transmission power, to obtain a desired performance.

In the recent years many works have been carried out for identification of the threshold region of the MLE. One approach is based on the method of interval estimation (MIE) [80] in which the performance of the MLE in the threshold region is approximated. However, it does not provide a performance lower bound. Another well investigated approach is the use of the non-Bayesian bounds which can predict the threshold region and provide a performance lower bound. In this chapter, we derive a Barankin-type bound (BTB) on the DOA estimation of acoustic signals based on measurement models with deterministic unknown parameters. The derived bound provides an approximation for the threshold SNR even in the case of a high-dimensional unknown parameter vector, for which computing the Barankin bound is usually formidable. We

examine the performance of the derived bound for both wide-band and narrow-band DOA estimation problems.

In addition, we examine the effect of the distance between the acoustic vector sensors (AVSs) in a uniform linear array on the narrow-band DOA estimation error. Unlike scalar sensors, AVSs measure not only the sound pressure but also the components of particle velocity [11]. One of the advantages of the AVS arrays over the scalar ones is that the distance between adjacent AVSs can be larger than half of the signal wavelength without introducing spatial ambiguities in narrow-band DOA estimation. Thus, we may design the distance between the AVSs to achieve better estimation results. Our examination of the derived BTB and actual MLE error shows that the threshold SNR increases with the distance between the AVSs, and that enlarging this distance does not always improve the estimation accuracy. Compared with the CRB, which consistently decreases with the increase of inter-AVS distance, the derived BTB provides more accurate reference for the optimal distance determination, as it incorporates the effect of large estimation errors.

The remainder of this chapter is organized as follows. In Section 6.2, we formulate the model for data collected by acoustic sensor arrays. In Section 6.3, we derive the BTB for source localization. Section 6.4 presents the numerical examples. The conclusion of this paper appears in Section 6.5.

6.2 Measurement Model

In this section, we present the measurement models for both wide-band and narrow-band sources for uniform linear AVS arrays. The measurement models for other types

of sensor arrays can be formulated in the same way using the corresponding steering vectors.

We first present the measurement model for wide-band sources. Consider signal waves from K distant wide-band acoustic sources impinging on a uniform linear array of M AVSs, each of which consists of m components. In similar to [69], [70], using the wide-band harmonic signal model, we approximate the incident signal from the k -th source as

$$x_k(t) = \sum_{j=1}^J c_{kj} \psi_j(t), \quad (6.1)$$

where $\psi_1(t), \dots, \psi_J(t)$ are the monochromatic waves representing Fourier series kernels and c_{k1}, \dots, c_{kJ} are their coefficients. Assuming the frequency of $\psi_j(t)$ is f_j , we write the measurement model as

$$\begin{aligned} \mathbf{y}(t) &= \sum_{j=1}^J \sum_{k=1}^K \mathbf{a}_j(\boldsymbol{\theta}_k) c_{kj} \psi_j(t) + \mathbf{e}(t) \\ &= \sum_{j=1}^J \mathbf{A}_j(\boldsymbol{\theta}) \mathbf{c}_j \psi_j(t) + \mathbf{e}(t), \quad t = 1, \dots, N, \end{aligned} \quad (6.2)$$

where $\mathbf{y}(t)$ is the $Mm \times 1$ measurement vector, $\boldsymbol{\theta}_k$ is the one- or two-dimensional DOA of the k -th source, $\mathbf{e}(t)$ is the $Mm \times 1$ noise vector, $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \dots, \boldsymbol{\theta}_K^T]^T$, $\mathbf{c}_j = [c_{1j}, \dots, c_{Kj}]^T$, “ T ” denotes the transpose operation, N is the number of temporal measurements, $\mathbf{A}_j(\boldsymbol{\theta}) = [\mathbf{a}_j(\boldsymbol{\theta}_1) \cdots \mathbf{a}_j(\boldsymbol{\theta}_K)]$, and $\mathbf{a}_j(\boldsymbol{\theta}_k)$ is the steering vector for the k -th source given by

$$\mathbf{a}_j(\boldsymbol{\theta}_k) = \mathbf{h}_j(\boldsymbol{\theta}_k) \otimes \begin{bmatrix} 1 \\ \mathbf{u}(\boldsymbol{\theta}_k) \end{bmatrix}, \quad (6.3)$$

where “ \otimes ” denotes the Kronecker product, $\mathbf{h}_j(\boldsymbol{\theta}_k) = [1, e^{i2\pi f_j \tau(\boldsymbol{\theta}_k)}, \dots, e^{i2\pi f_j (M-1)\tau(\boldsymbol{\theta}_k)}]^T$ with $\tau(\boldsymbol{\theta}_k)$ the differential time delay of the k -th source signal between two adjacent

AVSs, and $\mathbf{u}(\boldsymbol{\theta}_k)$ is the $(m-1) \times 1$ unit-norm direction vector from the array to the k -th source. We assume the noise is zero-mean circular Gaussian with known spatial covariance matrix $\boldsymbol{\Sigma}$ and is temporally white.

For narrow-band sources, the steering vectors can be assumed to be constant over the signal bandwidth with frequency fixed on the carrier frequency of the narrow-band sources, i.e., $\mathbf{a}_j(\boldsymbol{\theta}) = \mathbf{a}(\boldsymbol{\theta})$, and the measurement model (6.2) then becomes

$$\begin{aligned} \mathbf{y}(t) &= \sum_{k=1}^K \mathbf{a}(\boldsymbol{\theta}_k) \left(\sum_{j=1}^J c_{kj} \psi_j(t) \right) + \mathbf{e}(t) \\ &= \mathbf{A}(\boldsymbol{\theta}) \mathbf{C} \boldsymbol{\psi}(t) + \mathbf{e}(t), \quad t = 1, \dots, N, \end{aligned} \quad (6.4)$$

where \mathbf{C} is a $K \times J$ matrix whose (k, j) th entry is c_{kj} , $\boldsymbol{\psi}(t) = [\psi_1(t), \dots, \psi_J(t)]^T$, $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\boldsymbol{\theta}_1) \cdots \mathbf{a}(\boldsymbol{\theta}_K)]$, and

$$\mathbf{a}(\boldsymbol{\theta}_k) = \mathbf{h}(\boldsymbol{\theta}_k) \otimes \begin{bmatrix} 1 \\ \mathbf{u}(\boldsymbol{\theta}_k) \end{bmatrix}, \quad (6.5)$$

where $\mathbf{h}(\boldsymbol{\theta}_k) = [1, e^{i2\pi f\tau(\boldsymbol{\theta}_k)}, \dots, e^{i2\pi f(M-1)\tau(\boldsymbol{\theta}_k)}]^T$ and f is the carrier frequency of the narrow-band signals. Note that in (6.4), $\sum_{j=1}^J c_{kj} \psi_j(t)$ is the complex amplitude of the k -th incident signal approximated by Fourier series expansion.

6.3 Barankin-Type Bound

In this section, we derive a BTB on the DOA vector $\boldsymbol{\theta}$ based on the measurement models in Section II.

An analytical form of the Barankin bound is generally not available. We use the Barankin-type lower bound developed in [72]–[74] to approximate the actual Barankin Bound. For an unbiased estimator $\hat{\boldsymbol{\phi}}$ of an unknown deterministic parameter vector $\boldsymbol{\phi}$, this lower bound can be formulated as

$$\text{cov}\{\hat{\boldsymbol{\phi}}\} \geq \mathbf{T}(\mathbf{B} - \mathbf{1}\mathbf{1}^T)^{-1}\mathbf{T}^T, \quad (6.6)$$

where $\mathbf{T} = [\bar{\boldsymbol{\phi}}_1 - \boldsymbol{\phi} \cdots \bar{\boldsymbol{\phi}}_L - \boldsymbol{\phi}]$, $\bar{\boldsymbol{\phi}}_1, \dots, \bar{\boldsymbol{\phi}}_L$ are the test points selected from the parameter space of $\boldsymbol{\phi}$, $\mathbf{1}$ is a column vector of dimension L with all entries equal to 1, and \mathbf{B} is the $L \times L$ Barankin matrix whose (l, n) th entry is

$$B_{ln} = \text{E}\{r(\mathbf{y}, \bar{\boldsymbol{\phi}}_l, \boldsymbol{\phi})r(\mathbf{y}, \bar{\boldsymbol{\phi}}_n, \boldsymbol{\phi})\}, \quad (6.7)$$

where $\text{E}\{\cdot\}$ denotes the expectation operation, $r(\mathbf{y}, \bar{\boldsymbol{\phi}}_l, \boldsymbol{\phi}) = f(\mathbf{y}; \bar{\boldsymbol{\phi}}_l)/f(\mathbf{y}; \boldsymbol{\phi})$, and $f(\mathbf{y}; \boldsymbol{\phi})$ is the probability density function (pdf) of \mathbf{y} with parameters $\boldsymbol{\phi}$. The supremum of the BTB given in (6.6) approaches the actual Barankin bound when the number of test points tends to infinity [81].

In the following, we first derive the BTB on $\boldsymbol{\theta}$ based on the wide-band measurement model (6.2).

Stacking all measurements together, we rewrite (6.2) as

$$\mathbf{Y} = \sum_{j=1}^J \mathbf{A}_j(\boldsymbol{\theta}) \mathbf{c}_j \boldsymbol{\psi}_j^T + \mathbf{E}, \quad (6.8)$$

where $\mathbf{Y} = [\mathbf{y}(1) \cdots \mathbf{y}(N)]$, $\mathbf{E} = [\mathbf{e}(1) \cdots \mathbf{e}(N)]$, and $\boldsymbol{\psi}_j = [\psi_j(1), \dots, \psi_j(N)]^T$. Vectorizing both sides of (6.8), we obtain

$$\mathbf{y} = \sum_{j=1}^J (\boldsymbol{\psi}_j \otimes \mathbf{A}_j(\boldsymbol{\theta})) \mathbf{c}_j + \mathbf{e} = \mathbf{R}(\boldsymbol{\theta})\mathbf{c} + \mathbf{e}, \quad (6.9)$$

where $\mathbf{y} = [\mathbf{y}^T(1), \dots, \mathbf{y}^T(N)]^T$, $\mathbf{e} = [\mathbf{e}^T(1), \dots, \mathbf{e}^T(N)]^T$, $\mathbf{c} = [\mathbf{c}_1^T, \dots, \mathbf{c}_J^T]^T$, and $\mathbf{R}(\boldsymbol{\theta}) = [\boldsymbol{\psi}_1 \otimes \mathbf{A}_1(\boldsymbol{\theta}) \cdots \boldsymbol{\psi}_J \otimes \mathbf{A}_J(\boldsymbol{\theta})]$.

The unknown parameter vector in (6.9) is $\boldsymbol{\phi} = [\boldsymbol{\theta}^T, \text{Re}\{\mathbf{c}^T\}, \text{Im}\{\mathbf{c}^T\}]^T$, in which $\text{Re}\{\mathbf{c}^T\}$ and $\text{Im}\{\mathbf{c}^T\}$ are the real and imaginary parts of \mathbf{c}^T respectively. The probability density function of \mathbf{y} with parameters $\boldsymbol{\phi}$ is

$$f(\mathbf{y}; \boldsymbol{\phi}) = \frac{1}{\det(\pi \mathbf{I}_N \otimes \boldsymbol{\Sigma})} \exp \left\{ -(\mathbf{y} - \mathbf{R}(\boldsymbol{\theta})\mathbf{c})^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) (\mathbf{y} - \mathbf{R}(\boldsymbol{\theta})\mathbf{c}) \right\} \quad (6.10)$$

where $\det(\cdot)$ denotes the determinant of a matrix and \mathbf{I}_N is the unit matrix of size N .

Substituting (6.10) into (6.7), we obtain

$$\begin{aligned} B_{ln} &= \frac{1}{\det(\pi \mathbf{I}_N \otimes \boldsymbol{\Sigma})} \int \exp \left\{ (\mathbf{y} - \mathbf{R}(\boldsymbol{\theta})\mathbf{c})^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) (\mathbf{y} - \mathbf{R}(\boldsymbol{\theta})\mathbf{c}) - \right. \\ &\quad \left. (\mathbf{y} - \mathbf{R}(\bar{\boldsymbol{\theta}}_l)\bar{\mathbf{c}}_l)^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) (\mathbf{y} - \mathbf{R}(\bar{\boldsymbol{\theta}}_l)\bar{\mathbf{c}}_l) - \right. \\ &\quad \left. (\mathbf{y} - \mathbf{R}(\bar{\boldsymbol{\theta}}_n)\bar{\mathbf{c}}_n)^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) (\mathbf{y} - \mathbf{R}(\bar{\boldsymbol{\theta}}_n)\bar{\mathbf{c}}_n) \right\} d\mathbf{y} \\ &= A \exp \left\{ 2\text{Re} \left\{ (\mathbf{R}(\boldsymbol{\theta})\mathbf{c} - \mathbf{R}(\bar{\boldsymbol{\theta}}_l)\bar{\mathbf{c}}_l)^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) (\mathbf{R}(\boldsymbol{\theta})\mathbf{c} - \mathbf{R}(\bar{\boldsymbol{\theta}}_n)\bar{\mathbf{c}}_n) \right\} \right\}, \quad (6.11) \end{aligned}$$

where $[\bar{\boldsymbol{\theta}}_l^T, \text{Re}\{\bar{\mathbf{c}}_l^T\}, \text{Im}\{\bar{\mathbf{c}}_l^T\}]^T$ is the l th selected test point for $\boldsymbol{\phi}$ and

$$\begin{aligned} A &= \frac{1}{\det(\pi \mathbf{I}_N \otimes \boldsymbol{\Sigma})} \int \exp \left\{ (\mathbf{y} - \mathbf{R}(\bar{\boldsymbol{\theta}}_l)\bar{\mathbf{c}}_l - \mathbf{R}(\bar{\boldsymbol{\theta}}_n)\bar{\mathbf{c}}_n + \mathbf{R}(\boldsymbol{\theta})\mathbf{c})^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) \right. \\ &\quad \left. \times (\mathbf{y} - \mathbf{R}(\bar{\boldsymbol{\theta}}_l)\bar{\mathbf{c}}_l - \mathbf{R}(\bar{\boldsymbol{\theta}}_n)\bar{\mathbf{c}}_n + \mathbf{R}(\boldsymbol{\theta})\mathbf{c}) \right\} d\mathbf{y} = 1. \quad (6.12) \end{aligned}$$

Thus, we have

$$B_{ln} = \exp \left\{ 2\text{Re} \left\{ (\mathbf{R}(\boldsymbol{\theta})\mathbf{c} - \mathbf{R}(\bar{\boldsymbol{\theta}}_l)\bar{\mathbf{c}}_l)^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) (\mathbf{R}(\boldsymbol{\theta})\mathbf{c} - \mathbf{R}(\bar{\boldsymbol{\theta}}_n)\bar{\mathbf{c}}_n) \right\} \right\}. \quad (6.13)$$

To obtain a tight bound, we ought to choose the test points that maximize the right side of (6.6). Observe that in many cases $\boldsymbol{\phi}$ is a high-dimensional vector, which generally requires a huge amount of test points to produce a tight bound. This would inevitably augment the burden in computing the inverse of $\mathbf{B} - \mathbf{1}\mathbf{1}^T$.

We employ the approximation method proposed in [78] to reduce the required number of test points. This method is adapted as follows for DOA estimation:

- Select \bar{L} candidate test points $\bar{\boldsymbol{\theta}}_1, \dots, \bar{\boldsymbol{\theta}}_{\bar{L}}$ for the DOA parameter $\boldsymbol{\theta}$.
- For each selected $\bar{\boldsymbol{\theta}}_l$, $l = 1, \dots, \bar{L}$, select test point $\bar{\mathbf{c}}_{jl}$ for \mathbf{c}_j , $j = 1, \dots, J$ such that B_{ll} is minimized at $\bar{\boldsymbol{\phi}}_l = [\bar{\boldsymbol{\theta}}_l^T, \text{Re} \{ \bar{\mathbf{c}}_l^T \}, \text{Im} \{ \bar{\mathbf{c}}_l^T \}]^T = [\bar{\boldsymbol{\theta}}_l^T, \text{Re} \{ [\bar{\mathbf{c}}_{1l}^T, \dots, \bar{\mathbf{c}}_{Jl}^T] \}, \text{Im} \{ [\bar{\mathbf{c}}_{1l}^T, \dots, \bar{\mathbf{c}}_{Jl}^T] \}]^T$.
- From $\bar{\boldsymbol{\phi}}_1, \dots, \bar{\boldsymbol{\phi}}_{\bar{L}}$, choose the L test points producing the L smallest values in $B_{11}, \dots, B_{\bar{L}\bar{L}}$ as the test points to compute the bound [78]. Alternatively, choose the test points at the L highest lobe peaks of $B_{11}^{-1}, \dots, B_{\bar{L}\bar{L}}^{-1}$ provided that the values of $B_{11}^{-1}, \dots, B_{\bar{L}\bar{L}}^{-1}$ exhibit evident sidelobes.

To obtain the BTB according to the above method, we first find the test point $\bar{\mathbf{c}}_l$ that minimizes B_{ll} for a given $\bar{\boldsymbol{\theta}}_l$. Taking the derivative of B_{ll} with respect to $\bar{\mathbf{c}}_l$, we have the test point $\bar{\mathbf{c}}_l$ minimizing B_{ll} is

$$\bar{\mathbf{c}}_l = \mathbf{G}^{-1}(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_l) \mathbf{G}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) \mathbf{c}, \quad (6.14)$$

where $\mathbf{G}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) = \mathbf{R}^H(\bar{\boldsymbol{\theta}}_l) (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{R}(\boldsymbol{\theta})$.

The general result provided by (6.14) can be further simplified if we have

$$\boldsymbol{\psi}_k^H \boldsymbol{\psi}_j = \delta_{kj} \quad (6.15)$$

with δ_{kj} the Kronecker delta function, which is a general property of the problem at hand since $\psi_1(t), \dots, \psi_J(t)$ can be chosen using the kernels of the Fourier series or by Karhunen-Loève transform. Using (6.14) and (6.15), we can further obtain (see Appendix P) the test point $\bar{\mathbf{c}}_{jl}$ that minimizes B_{ll}

$$\bar{\mathbf{c}}_{jl} = \mathbf{P}_j^{-1}(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_l) \mathbf{P}_j(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) \mathbf{c}_j, \quad j = 1, \dots, J, \quad (6.16)$$

where $\mathbf{P}_j(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) = \mathbf{A}_j^H(\bar{\boldsymbol{\theta}}_l) \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta})$.

Substituting (6.16) into (6.13), we obtain the following expression for the (l, n) th entry of \mathbf{B} :

$$B_{ln} = \exp \left\{ 2 \operatorname{tr} \left\{ \operatorname{Re} \left\{ \mathbf{Q}^H(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) \boldsymbol{\Sigma}^{-1} \mathbf{Q}(\bar{\boldsymbol{\theta}}_n, \boldsymbol{\theta}) \right\} \right\} \right\}, \quad (6.17)$$

where $\operatorname{tr}\{\cdot\}$ denotes the trace of a matrix, $\mathbf{Q}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) = [\mathbf{q}_1(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) \cdots \mathbf{q}_J(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta})]$, and $\mathbf{q}_j(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) = (\mathbf{A}_j(\boldsymbol{\theta}) - \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l) \mathbf{P}_j^{-1}(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_l) \mathbf{P}_j(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta})) \mathbf{c}_j$.

Assuming the noise covariance matrix $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$ and there is only one source, we are able to simplify the formula in (6.17) as

$$B_{ln} = \exp \left\{ \frac{2}{\sigma^2} \sum_{j=1}^J \operatorname{Re} \left\{ (\mathbf{A}_j(\boldsymbol{\theta}) \mathbf{c}_j)^H \boldsymbol{\Lambda}_j(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_n) (\mathbf{A}_j(\boldsymbol{\theta}) \mathbf{c}_j) \right\} \right\}, \quad (6.18)$$

where

$$\boldsymbol{\Lambda}_j(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_n) = (\mathbf{I} - \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l) \mathbf{A}_j^H(\bar{\boldsymbol{\theta}}_l) / \alpha) (\mathbf{I} - \mathbf{A}_j(\bar{\boldsymbol{\theta}}_n) \mathbf{A}_j^H(\bar{\boldsymbol{\theta}}_n) / \alpha). \quad (6.19)$$

In (6.19), α is the square of the Euclidean norm of the steering vector. For scalar-sensor arrays, we have $\alpha = M$. For vector-sensor arrays, we have $\alpha = 2M$ (note that the Euclidean norm of $\mathbf{u}(\boldsymbol{\theta})$ is 1). The diagonal entries of \mathbf{B} can be further simplified as

$$B_{ll} = \exp \left\{ \frac{2}{\sigma^2} \sum_{j=1}^J \mathbf{c}_j^H \mathbf{c}_j (\alpha - |\mathbf{A}_j^H(\boldsymbol{\theta}) \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)|^2 / \alpha) \right\}, \quad (6.20)$$

where $|\cdot|$ denotes the amplitude of a complex number.

Using the derived expressions of \mathbf{B} , we are able to compute the BTB on an unbiased estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, which is given by

$$\text{cov}\{\hat{\boldsymbol{\theta}}\} \geq \mathbf{T}_\theta (\mathbf{B} - \mathbf{1}\mathbf{1}^T)^{-1} \mathbf{T}_\theta^T, \quad (6.21)$$

where $\mathbf{T}_\theta = [\bar{\boldsymbol{\theta}}_1 - \boldsymbol{\theta} \cdots \bar{\boldsymbol{\theta}}_L - \boldsymbol{\theta}]$.

The bound based on the narrow-band model (6.4) can be derived similarly and herein we only present the results. For the narrow-band model in (6.4), the test point $\bar{\mathbf{C}}_l$ minimizing B_{ll} is

$$\bar{\mathbf{C}}_l = \mathbf{P}^{-1}(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_l) \mathbf{P}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) \mathbf{C}. \quad (6.22)$$

where $\mathbf{P}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) = \mathbf{A}^H(\bar{\boldsymbol{\theta}}_l) \boldsymbol{\Sigma}^{-1} \mathbf{A}(\boldsymbol{\theta})$. Under the condition in (6.15), the formula of \mathbf{B} for the narrow-band model can be written as

$$B_{ln} = \exp \left\{ 2 \text{tr} \left\{ \text{Re} \left\{ \mathbf{Q}^H(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) \boldsymbol{\Sigma}^{-1} \mathbf{Q}(\bar{\boldsymbol{\theta}}_n, \boldsymbol{\theta}) \right\} \right\} \right\}, \quad (6.23)$$

where $\mathbf{Q}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) = (\mathbf{A}(\boldsymbol{\theta}) - \mathbf{A}(\bar{\boldsymbol{\theta}}_l) \mathbf{P}^{-1}(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_l) \mathbf{P}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta})) \mathbf{C}$.

Note that though our test-point selection method is based on the idea in [78], our models and derived results are more general than their deterministic counterparts

in [78] and thus have broader applications. Our derived bounds can be used as well for some other array-processing problems such as DOA estimation for signals of known waveforms [15], [16], in which $\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_J$ denote the known signal waveforms.

6.4 Numerical Examples

To demonstrate the performance of the derived bound, we present numerical examples of DOA estimation for both scalar- and vector-sensor arrays. For simplicity, we assume the noise covariance matrix $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$ in all examples.

6.4.1 Examples for Scalar-Sensor Array

We present results of two examples for scalar-sensor array, one for narrow-band and the other for wide-band signal. For the narrow-band example, we consider a sinusoidal signal with elevation angle of 0 degree arriving at a uniform linear scalar-sensor array of 40 sensors. The distance d between the sensors equals half of the signal's wavelength λ . We took 10 samples of measurement and uniformly selected 30 test points from $[-\pi/2, \pi/2]$ to compute the bound. Note that uniformly selecting test points is a special case of the proposed test point selection method with $\bar{L} = L$. Also for a uniform linear scalar-sensor array, the distance between adjacent sensors should not be larger than half of the signal wavelength to avoid ambiguity in the DOA estimation of a narrow-band signal.

For the wide-band example, we assume the distance between the sensors is 1 meter and the signal consists of 10 harmonic components, with elevation angle of 0 degree.

The frequencies of the 10 harmonic components of the signal are 20, 40, \dots , 180, and 200 Hz, respectively. The linear array consists of 40 sensors. The sound propagation speed is $c = 1500$ meters per second. We took 10 measurements at a sampling period of 5 ms and uniformly selected 60 test points to compute the bound. In this problem, we have 10 unknown signal parameters, which would make approximating the Barankin bound challenging.

We present the computed square roots (SRs) of the derived BTB and actual MLE error in Figs. 6.1 and 6.2. It can be seen that the derived bounds predict the true threshold SNRs with differences about 2 to 3 dB in these two examples.

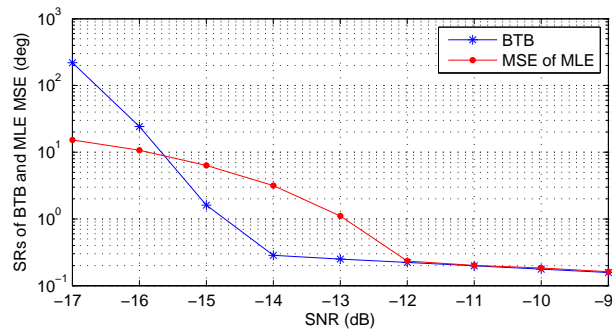


Figure 6.1: Square roots of the derived Barankin-type bound and maximum likelihood estimation mean-square error versus signal-to-noise ratio for narrow-band direction of arrival estimation using uniform linear scalar-sensor array.

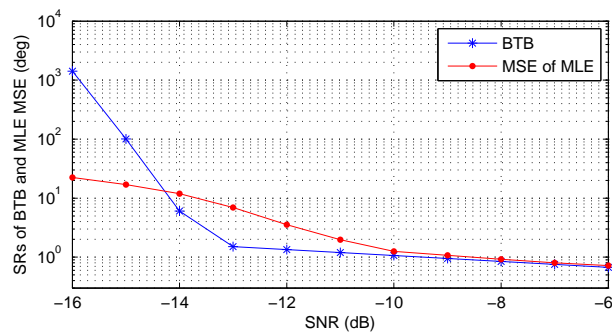


Figure 6.2: Same as in Fig.1, but for the wide-band direction of arrival estimation.

6.4.2 Examples for Vector-Sensor Array

AVS arrays outperform the standard scalar-sensor arrays, as they utilize measurements from the components of the acoustic particle velocity as well as the pressure [11]. They resolve end-fire and conical ambiguities [82] and a single AVS can identify up to two sources in three-dimensional space [83]. Clearly, since a single AVS is able to provide DOA information, the distance d between the AVSs in a uniform linear array is not limited to be smaller than half of the wavelength of the narrow-band impinging signal.

Since d can be larger than half the signal wavelength, it is possible to improve the estimation accuracy through adjusting the distance between the AVSs. The lower bounds on the mean-square estimation error should provide useful information about the choice of the optimal distance. We examined the effect of the distance between the AVSs on the derived BTB, CRB, and actual MLE error for a narrow-band one-dimensional DOA estimation problem (see Fig. 6.3). In our example, we consider a narrow-band signal with DOA = 0 degree impinges on a uniform linear array consisting of 10 AVSs. The vector sensors measure the pressure and only two velocity components such that $\mathbf{h}(\theta) = [1, e^{i\frac{2\pi}{\lambda}d\sin\theta}, \dots, e^{i\frac{2\pi}{\lambda}(M-1)d\sin\theta}]^T$ and $\mathbf{u}(\theta) = [\cos\theta, \sin\theta]^T$ in (6.5), where θ is the one-dimensional DOA and λ is the wavelength of the narrow-band signal. We selected the test points at the peaks of the B_{ll}^{-1} values over 200 uniformly sampled candidate points, and computed the bounds and MLE errors using one measurement for $d = \lambda, 2\lambda$ and 3λ , respectively. According to [84], selecting test points at the peaks of B_{ll}^{-1} is equivalent to selecting test points at the lobes of the maximum likelihood ambiguity function (MLAF).

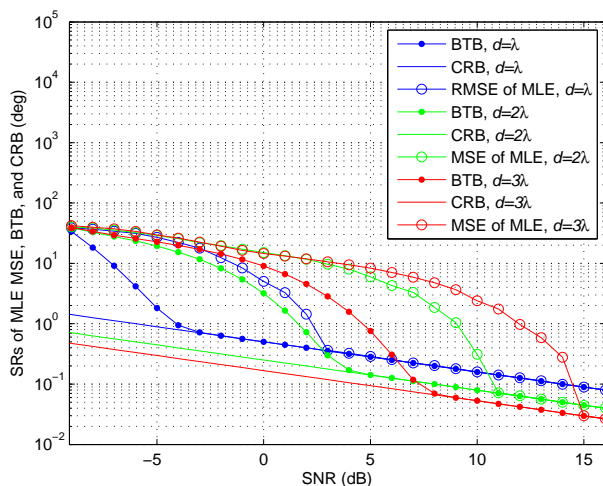


Figure 6.3: Square roots of the derived Barankin-type bound, Cramér-Rao bound, and maximum likelihood estimation mean-square error versus signal-to-noise ratio for narrow-band direction of arrival estimation using uniform linear acoustic vector-sensor arrays with $d = \lambda$, 2λ , and 3λ , respectively.

Observe from the results of the BTB and MLE error in Fig. 6.3 that the SNR threshold increases with the distance between the AVSs, and that increasing the distance between the AVSs does not always improve the estimation accuracy. From the MLAF plots in Fig. 6.4, we can see that with the increase of d , the main lobe width of the MLAF decreases whereas the number and level of side lobes increase. Therefore, when the d value increases, the asymptotic ML estimation error decreases while the threshold SNR increases. This explains the phenomena observed in Fig. 6.3. Since the MLE search includes ambiguities from side lobes, we must use a global error bound like the BTB to capture the threshold effect impact of these ambiguities often observed at medium to low SNRs. From Fig. 3, we can see that the derived BTB provides better estimation of the optimal distance than the CRB, which is a local error bound and consistently decreases with the increase of inter-AVS distance.

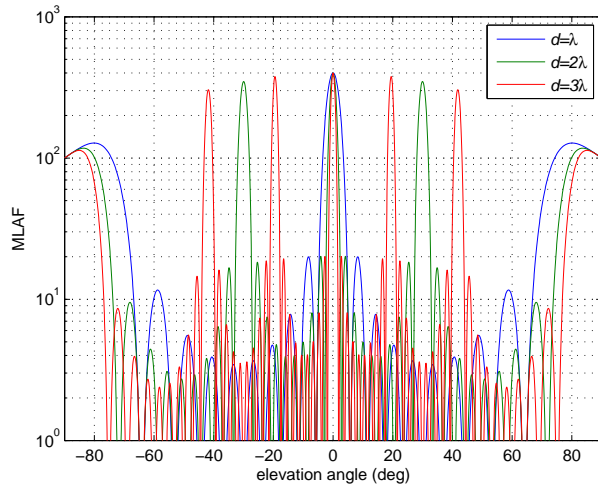


Figure 6.4: Variation of maximum likelihood ambiguity functions with respect to the inter-sensor distance d .

Note that although the derived bound correctly predicts that the threshold SNR increases with the distance between the sensors, the true threshold SNRs are about 6 to 7 dB larger than the predicted ones due to the approximation in computing the bound. We found similar differences between the true and predicted threshold SNRs in the wide-band DOA estimation with an AVS array of 10 sensors and the same setup as above for the wide-band signal.

6.5 Summary

We derived a BTB on DOA estimation using acoustic sensor arrays. Numerical examples show that this bound closely predicts the threshold SNR for DOA estimation using scalar-sensor arrays. For the narrow-band DOA estimation using AVS arrays, the distance between the sensors is not required to be lower than half a wavelength to avoid ambiguity. Numerical results from the derived BTB demonstrate that increasing

the distance between the AVSs improves the DOA estimation accuracy consistently only at high SNRs, and it increases the threshold SNR as well. As a result, at medium SNR values the estimation accuracy does not necessarily improve with this distance. This predicts the behavior of the actual MLE error. The derived BTB exhibits evident advantage over the CRB in determining the optimal distance between the AVSs.

Chapter 7

Conclusions

In this dissertation, we proposed new signal models, maximum likelihood (ML) estimators, and performance analysis results for some selected topics in underwater direction-of-arrival (DOA) estimation. In the following, we first summarize the key contributions of this dissertation, and then discuss possible topics for future research.

7.1 Key Contributions

We first considered the DOA estimation problem in spatially colored noise using sparse sensor arrays, for which we developed new ML DOA estimators for both zero-mean and non-zero-mean Gaussian signals based on an Expectation-Maximization framework. We derived the CRB on DOA estimation of non-zero-mean Gaussian signals, and the asymptotic error covariance matrix of using the typical stochastic estimator, which assumes zero-mean Gaussian signals, for DOA estimation of non-zero-mean Gaussian signals. Our analytical comparison and numerical examples show that the non-zero means in the signals improves the accuracy of DOA estimation.

Next, we considered the narrow-band DOA estimation of hydroacoustic signals from marine vessels containing both sinusoidal and random components, for which we presented a new signal model and derived the ML estimator for DOA finding of such hydroacoustic signals. We computed the asymptotic error covariance matrices of the proposed ML estimator and the typical stochastic estimator assuming zero-mean Gaussian signals on DOA finding of such hydroacoustic signals, which, together with our analytical and numerical comparisons, showed that the proposed signal model and estimator improve the DOA estimation accuracy for the hydroacoustic signals from marine vessels in comparison with the typical stochastic ones.

We then generalized the narrow-band DOA estimation results to the wide-band case for hydroacoustic signals from marine vessels through narrow-band decomposition and Fourier series expansion. We proposed the wide-band measurement models and ML estimators. We showed that compared with the typical stochastic estimators, the proposed ML estimators provide better performance for wide-band DOA finding of the hydroacoustic signals from marine vessels.

At last, to identify the SNR threshold region, we derived a Barankin-type bound on DOA estimation using acoustic sensor arrays. Numerical examples show that this bound closely predicts the threshold SNR for DOA estimation using scalar-sensor arrays. The derived Barankin-type bound also demonstrated that increasing the distance between the acoustic vector sensors improves the DOA estimation accuracy only at high SNRs, and it increases the SNR threshold as well, and the estimation accuracy does not necessarily improve with this distance at medium SNR values. This predicts the actual behavior of the ML estimation error. In addition, the derived Barankin-type bound exhibits evident advantage over the Cramér-Rao bound in determining the optimal distance between the acoustic vector sensors.

7.2 Future Work

In our future work, we will develop new schemes to improve the efficiency of DOA estimation in spatially colored noise using sparse sensor arrays. Sparse modeling techniques [85]-[89] have been shown to be quite efficient and accurate in DOA estimation under spatially white noise. However, concrete sparse modeling methods for DOA estimation using sparse sensor arrays under spatially colored noise have not been addressed yet to the best of our knowledge. We will develop novel and efficient sparse-modeling-based estimators for DOA finding under spatially colored noise using sparse sensor arrays.

Also in our future work, we will develop new methods to improve the accuracy and efficiency for DOA estimation of hydroacoustic signals from marine vessels. We will develop new DOA estimation methods based on the sparse modeling on both DOA and sinusoidal frequencies to improve the estimation efficiency. For the wide-band estimation based on Fourier transform, we will develop efficient estimation schemes based on the measurement model proposed in (5.5) to improve the estimation performance. This model is more accurate than the one in (5.12), based on which the current estimator in Chapter 5 is developed. We expect it to introduce a better estimator than the current one proposed in Chapter 5.

For the Barankin-type bound, we will develop more accurate computation of the Barankin bound for high-dimensional unknown parameter vectors as well as better signal models such as those assuming Gaussian signals. We will aim to improve the approximation of the Barankin bound on DOA estimation by developing new test point selection methods and considering alternative approximation approaches of the Barankin bound such as the one in [81].

Appendix A

Proof of Theorem 1

Let $\boldsymbol{\beta} = [\boldsymbol{\theta}^T, \boldsymbol{\delta}^T, \bar{\mathbf{b}}^T]^T$ be the real column vector containing all the unknown parameters, where $\boldsymbol{\delta}$ is the real column vector containing the unknown parameters from \mathbf{P} and \mathbf{Q} , and $\bar{\mathbf{b}}$ is the real column vector containing those from \mathbf{b} . According to the Fisher information matrix equation [42], we have

$$\begin{aligned} [\mathbf{FIM}_M]_{kl} = 2N\text{Re} \left\{ \frac{\partial[\mathbf{A}(\boldsymbol{\theta})\mathbf{b}]^H}{\partial\beta_k} \mathbf{R}^{-1} \frac{\partial[\mathbf{A}(\boldsymbol{\theta})\mathbf{b}]}{\partial\beta_l} \right\} \\ + N\text{trace} \left\{ \mathbf{R}^{-1} \frac{\partial\mathbf{R}}{\partial\beta_k} \mathbf{R}^{-1} \frac{\partial\mathbf{R}}{\partial\beta_l} \right\}, \end{aligned} \quad (\text{A.1})$$

where \mathbf{FIM}_M is the mixed Fisher information matrix, $[\mathbf{FIM}_M]_{kl}$ is the (k, l) -th element of \mathbf{FIM}_M , $\text{Re}\{\cdot\}$ denotes the real part of a complex number, and β_k is the k -th element of $\boldsymbol{\beta}$.

Let \mathbf{F}_d be the matrix whose entries are from the first term on the right side of (A.1).

If β_k is a parameter from $\boldsymbol{\delta}$, we have

$$\frac{\partial[\mathbf{A}(\boldsymbol{\theta})\mathbf{b}]}{\partial\beta_k} = \mathbf{0}. \quad (\text{A.2})$$

As a result, we obtain

$$\mathbf{F}_d = \begin{bmatrix} \mathbf{F}_{d,\theta\theta} & \mathbf{F}_{d,\theta\delta} & \mathbf{F}_{d,\theta\bar{b}} \\ \mathbf{F}_{d,\delta\theta} & \mathbf{F}_{d,\delta\delta} & \mathbf{F}_{d,\delta\bar{b}} \\ \mathbf{F}_{d,\bar{b}\theta} & \mathbf{F}_{d,\bar{b}\delta} & \mathbf{F}_{d,\bar{b}\bar{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{d,\theta\theta} & \mathbf{0} & \mathbf{F}_{d,\theta\bar{b}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{F}_{d,\bar{b}\theta} & \mathbf{0} & \mathbf{F}_{d,\bar{b}\bar{b}} \end{bmatrix}, \quad (\text{A.3})$$

where $\mathbf{F}_{d,\theta\theta}, \mathbf{F}_{d,\theta\delta}, \dots, \mathbf{F}_{d,\bar{b}\bar{b}}$ are the sub-matrices of \mathbf{F}_d containing the elements related to the parameters specified in their subscripts.

Similarly, let \mathbf{F}_S be the matrix whose entries are from the second term on the right side of (A.1). Since

$$\frac{\partial \mathbf{R}}{\partial \beta_k} = \mathbf{0} \quad (\text{A.4})$$

if β_k is a parameter from \mathbf{b} , we have

$$\mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{S,\theta\theta} & \mathbf{F}_{S,\theta\delta} & \mathbf{F}_{S,\theta\bar{b}} \\ \mathbf{F}_{S,\delta\theta} & \mathbf{F}_{S,\delta\delta} & \mathbf{F}_{S,\delta\bar{b}} \\ \mathbf{F}_{S,\bar{b}\theta} & \mathbf{F}_{S,\bar{b}\delta} & \mathbf{F}_{S,\bar{b}\bar{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{S,\theta\theta} & \mathbf{F}_{S,\theta\delta} & \mathbf{0} \\ \mathbf{F}_{S,\delta\theta} & \mathbf{F}_{S,\delta\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (\text{A.5})$$

Combining the results from (A.3) and (A.5), we obtain

$$\mathbf{FIM}_M = \mathbf{F}_d + \mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{d,\theta\theta} + \mathbf{F}_{S,\theta\theta} & \mathbf{F}_{S,\theta\delta} & \mathbf{F}_{d,\theta\bar{b}} \\ \mathbf{F}_{S,\delta\theta} & \mathbf{F}_{S,\delta\delta} & \mathbf{0} \\ \mathbf{F}_{d,\bar{b}\theta} & \mathbf{0} & \mathbf{F}_{d,\bar{b}\bar{b}} \end{bmatrix}. \quad (\text{A.6})$$

Therefore, we have

$$\begin{aligned} \mathbf{CRB}_{M,\theta}^{-1} &= \mathbf{F}_{d,\theta\theta} + \mathbf{F}_{S,\theta\theta} - [\mathbf{F}_{S,\theta\delta} \quad \mathbf{F}_{d,\theta\bar{b}}] \begin{bmatrix} \mathbf{F}_{S,\delta\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{d,\bar{b}\bar{b}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{F}_{S,\delta\theta} \\ \mathbf{F}_{d,\bar{b}\theta} \end{bmatrix} \\ &= \mathbf{F}_{d,\theta\theta} - \mathbf{F}_{d,\theta\bar{b}} \mathbf{F}_{d,\bar{b}\bar{b}}^{-1} \mathbf{F}_{d,\bar{b}\theta}^H + \mathbf{F}_{S,\theta\theta} - \mathbf{F}_{S,\theta\delta} \mathbf{F}_{S,\delta\delta}^{-1} \mathbf{F}_{S,\theta\delta}^H, \end{aligned} \quad (\text{A.7})$$

where

$$\mathbf{F}_{\mathbf{S},\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{S},\boldsymbol{\theta}\boldsymbol{\delta}}\mathbf{F}_{\mathbf{S},\boldsymbol{\delta}\boldsymbol{\delta}}^{-1}\mathbf{F}_{\mathbf{S},\boldsymbol{\delta}\boldsymbol{\theta}}^H = \mathbf{CRB}_{\mathbf{S},\boldsymbol{\theta}}^{-1}. \quad (\text{A.8})$$

For 1D DOA estimation, we have

$$\frac{\partial[\mathbf{A}(\boldsymbol{\theta})\mathbf{b}]}{\partial\theta_l} = b_l \frac{d\mathbf{a}(\theta_l)}{d\theta_l}, \quad (\text{A.9})$$

where b_l is the l -th element of \mathbf{b} . As a result,

$$\frac{\partial[\mathbf{A}(\boldsymbol{\theta})\mathbf{b}]}{\partial\boldsymbol{\theta}^T} = \mathbf{D}\mathbf{H}, \quad (\text{A.10})$$

where $\mathbf{H} = \text{diag}\{\mathbf{b}\}$ is a diagonal matrix with diagonal elements from vector \mathbf{b} .

According to the results in [19], we have

$$\begin{aligned} \mathbf{F}_{\mathbf{d},\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{d},\boldsymbol{\theta}\tilde{\mathbf{b}}}\mathbf{F}_{\mathbf{d},\tilde{\mathbf{b}}\tilde{\mathbf{b}}}^{-1}\mathbf{F}_{\mathbf{d},\tilde{\mathbf{b}}\boldsymbol{\theta}}^H &= 2N\text{Re} \left\{ \mathbf{H}^H \bar{\mathbf{D}}^H \Pi_{\bar{\mathbf{A}}}^\perp \bar{\mathbf{D}} \mathbf{H} \right\} \\ &= 2N\text{Re} \left\{ \left(\bar{\mathbf{D}}^H \Pi_{\bar{\mathbf{A}}}^\perp \bar{\mathbf{D}} \right) \odot (\mathbf{b}\mathbf{b}^H)^T \right\} \\ &= 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}^H \Pi_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}} \right) \odot (\mathbf{b}\mathbf{b}^H)^T \right\} \\ &= \mathbf{CRB}_{\mathbf{d},\boldsymbol{\theta}}^{-1}. \end{aligned} \quad (\text{A.11})$$

The second last equality in (A.11) is obtained from the following result given in [43]:

$$\mathbf{R}^{-\frac{1}{2}} \Pi_{\bar{\mathbf{A}}}^\perp \mathbf{R}^{-\frac{1}{2}} = \mathbf{Q}^{-\frac{1}{2}} \Pi_{\tilde{\mathbf{A}}}^\perp \mathbf{Q}^{-\frac{1}{2}}. \quad (\text{A.12})$$

For 2D DOA estimation, we have

$$\frac{\partial[\mathbf{A}(\boldsymbol{\theta})\mathbf{b}]}{\partial\boldsymbol{\theta}_l^T} = \frac{\partial\mathbf{a}(\boldsymbol{\theta}_l)}{\partial\boldsymbol{\theta}_l^T} \begin{bmatrix} b_l & 0 \\ 0 & b_l \end{bmatrix} = \begin{bmatrix} d\mathbf{a}(\boldsymbol{\theta}_l) & d\mathbf{a}(\boldsymbol{\theta}_l) \\ d\phi_l & d\psi_l \end{bmatrix} \begin{bmatrix} b_l & 0 \\ 0 & b_l \end{bmatrix}, \quad (\text{A.13})$$

and thus

$$\frac{\partial[\mathbf{A}(\boldsymbol{\theta})\mathbf{b}]}{\partial\boldsymbol{\theta}^T} = \mathbf{D}_2\mathbf{H}_2, \quad (\text{A.14})$$

where $\mathbf{H}_2 = \mathbf{H} \otimes \mathbf{I}_2$, and \mathbf{I}_2 is the 2×2 identity matrix.

Therefore, for 2D DOA estimation,

$$\begin{aligned} \mathbf{F}_{\text{d},\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\text{d},\boldsymbol{\theta}\bar{\mathbf{b}}}\mathbf{F}_{\text{d},\bar{\mathbf{b}}\bar{\mathbf{b}}}^{-1}\mathbf{F}_{\text{d},\bar{\mathbf{b}}\boldsymbol{\theta}}^H &= 2N\text{Re} \left\{ \mathbf{H}_2^H \bar{\mathbf{D}}_2^H \Pi_A^\perp \bar{\mathbf{D}}_2 \mathbf{H}_2 \right\} \\ &= 2N\text{Re} \left\{ \left(\bar{\mathbf{D}}_2^H \Pi_A^\perp \bar{\mathbf{D}}_2 \right) \odot [(\mathbf{b}\mathbf{b}^H) \otimes \boldsymbol{\Delta}]^T \right\} \\ &= 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \Pi_A^\perp \tilde{\mathbf{D}}_2 \right) \odot [(\mathbf{b}\mathbf{b}^H) \otimes \boldsymbol{\Delta}]^T \right\} \\ &= \mathbf{CRB}_{\text{d},\boldsymbol{\theta}}^{-1}. \end{aligned} \quad (\text{A.15})$$

We thus prove Theorem 1.

Appendix B

Proof of Theorem 2

If the stochastic ML estimator is used for DOA estimation of mixed signals, as the number of measurements N increases, the stochastic log-likelihood function tends to the asymptotic stochastic log-likelihood function with signal covariance matrix equal to $\mathbf{R}_{xx} = \mathbf{P} + \mathbf{b}\mathbf{b}^H$, and we are still able to obtain consistent estimators for $\boldsymbol{\theta}$, \mathbf{R}_{xx} , and \mathbf{Q} .

Let $f(\mathbf{Y}; \boldsymbol{\alpha})$ be the pdf based on the stochastic measurement model with $\boldsymbol{\alpha}$ the real column vector containing all the unknown parameters from $\boldsymbol{\theta}$, \mathbf{R}_{xx} , and \mathbf{Q} . Assuming $\hat{\boldsymbol{\alpha}}$ is the ML estimator of $\boldsymbol{\alpha}$, we apply a Taylor expansion as follows around the true value of $\boldsymbol{\alpha}$, which is denoted by $\boldsymbol{\alpha}_0$.

$$\left. \frac{\partial \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}} = \left. \frac{\partial \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} + \left. \frac{\partial^2 \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^2} \right|_{\boldsymbol{\alpha}=\tilde{\boldsymbol{\alpha}}} (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) = \mathbf{0}, \quad (\text{B.1})$$

where $\tilde{\boldsymbol{\alpha}}$ is a vector value between $\boldsymbol{\alpha}_0$ and $\hat{\boldsymbol{\alpha}}$. Thus we have

$$\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0 = - \left[\left. \frac{1}{N} \frac{\partial^2 \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} \right]^{-1} \right|_{\boldsymbol{\alpha}=\tilde{\boldsymbol{\alpha}}} \left[\left. \frac{1}{N} \frac{\partial \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right] \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}. \quad (\text{B.2})$$

Using equations (15.47) and (15.48) in [42], we can show that

$$\mathbb{E} \left\{ \left. \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} = \mathbf{0}, \quad (\text{B.3})$$

$$\mathbb{E} \left\{ \left. \frac{\partial^2 \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \alpha_k \partial \alpha_l} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} = \text{trace} \left\{ \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_k} \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_l} \right\}. \quad (\text{B.4})$$

Note that the expectations in (B.3) and (B.4) are computed under the pdf of the mixed measurement model. This also holds for the all the following expectations in this proof.

According to the central limit theorem, the second term on the right side of (B.2) follows a Gaussian distribution with mean

$$\mathbb{E} \left\{ \left. \frac{1}{N} \frac{\partial \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} = \mathbf{0} \quad (\text{B.5})$$

and covariance matrix

$$\begin{aligned} & \mathbb{E} \left\{ \left. \frac{1}{N^2} \frac{\partial \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \frac{\partial \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^T} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} \\ &= \frac{1}{N^2} \mathbb{E} \left\{ \left[\sum_{t=1}^N \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right] \left[\sum_{s=1}^N \frac{\partial \ln f(\mathbf{y}(s); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^T} \right] \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} \\ &= \frac{1}{N^2} \sum_{t=1}^N \mathbb{E} \left\{ \left. \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^T} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} \\ &= \frac{1}{N} \mathbb{E} \left\{ \left. \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^T} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\}. \end{aligned} \quad (\text{B.6})$$

Using equations (15.47) and (15.48) in [42], and after some algebraic manipulations, we obtain

$$\begin{aligned}
& \mathbb{E} \left\{ \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \alpha_k} \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \alpha_l} \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} \\
&= \mathbb{E} \left\{ \mathbf{y}^H(t) \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_k} \bar{\mathbf{R}}^{-1} \mathbf{y}(t) \mathbf{y}^H(t) \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_l} \bar{\mathbf{R}}^{-1} \mathbf{y}(t) \right\} \\
&\quad - \text{trace} \left\{ \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_k} \right\} \text{trace} \left\{ \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_l} \right\}. \tag{B.7}
\end{aligned}$$

Applying Lemma 2 to the first term on the right side of (B.7), we obtain

$$\begin{aligned}
& \mathbb{E} \left\{ \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \alpha_k} \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \alpha_l} \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} \\
&= \text{trace} \left\{ \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_k} \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_l} \right\} - \text{trace} \left\{ \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_k} \bar{\mathbf{R}}^{-1} \right. \\
&\quad \left. \times \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_l} \bar{\mathbf{R}}^{-1} \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \right\}. \tag{B.8}
\end{aligned}$$

Thus, we have

$$\mathbb{E} \left\{ \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \frac{\partial \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^T} \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\} = \mathbf{F}_{\text{SS}} - \mathbf{F}_{\mathbf{b}}, \tag{B.9}$$

where $[\mathbf{F}_{\text{SS}}]_{kl}$ and $[\mathbf{F}_{\mathbf{b}}]_{kl}$ are equal to the first and second terms on the right side of (B.8), respectively. Note that \mathbf{F}_{SS} is the Fisher information matrix for $\boldsymbol{\alpha}$ under the single measurement pdf $f(\mathbf{y}(t); \boldsymbol{\alpha})$.

For the first term on the right side of (B.2), according to the law of large numbers and the consistency of $\hat{\boldsymbol{\alpha}}$, we have the following result as $N \rightarrow \infty$:

$$\begin{aligned}
& \frac{1}{N} \frac{\partial^2 \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \alpha_k \partial \alpha_l} \Big|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}} \rightarrow \frac{1}{N} \frac{\partial^2 \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \alpha_k \partial \alpha_l} \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \\
& \rightarrow \mathbb{E} \left\{ \frac{\partial^2 \ln f(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \alpha_k \partial \alpha_l} \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \right\}. \tag{B.10}
\end{aligned}$$

From the result in (B.4), as $N \rightarrow \infty$, we obtain

$$\frac{1}{N} \frac{\partial^2 \ln f(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} \Big|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}} \rightarrow \mathbf{F}_{\text{SS}}. \quad (\text{B.11})$$

According to Slutsky's Theorem, as $N \rightarrow \infty$, $\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0$ asymptotically follows the Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{C}_A)$, where the covariance matrix

$$\mathbf{C}_A = \frac{1}{N} \mathbf{F}_{\text{SS}}^{-1} (\mathbf{F}_{\text{SS}} - \mathbf{F}_b) \mathbf{F}_{\text{SS}}^{-1} = \frac{1}{N} (\mathbf{F}_{\text{SS}}^{-1} - \mathbf{F}_{\text{SS}}^{-1} \mathbf{F}_b \mathbf{F}_{\text{SS}}^{-1}). \quad (\text{B.12})$$

Now let us consider the asymptotic error covariance matrix $\mathbf{A}\mathbf{C}_{\text{S},\boldsymbol{\theta}}$ for DOA estimation, which is a sub-matrix of \mathbf{C}_A . According to (B.12), we have

$$\mathbf{A}\mathbf{C}_{\text{S},\boldsymbol{\theta}} = \mathbf{C}_{\text{S},\boldsymbol{\theta}} - \mathbf{C}_{b,\boldsymbol{\theta}}, \quad (\text{B.13})$$

where $\mathbf{C}_{\text{S},\boldsymbol{\theta}}$ and $\mathbf{C}_{b,\boldsymbol{\theta}}$ are the sub-matrices related to $\boldsymbol{\theta}$ in $\mathbf{F}_{\text{SS}}^{-1}/N$ and $\mathbf{F}_{\text{SS}}^{-1} \mathbf{F}_b \mathbf{F}_{\text{SS}}^{-1}/N$, respectively. Note that $\mathbf{C}_{\text{S},\boldsymbol{\theta}}$ is the stochastic CRB on DOA estimation with measurements from $\mathcal{CN}(\mathbf{0}, \bar{\mathbf{R}})$. Thus, to prove this theorem, we need to show $\mathbf{C}_{b,\boldsymbol{\theta}} = \mathbf{0}$.

We first reformulate \mathbf{F}_b as follows.

$$\begin{aligned} [\mathbf{F}_b]_{kl} &= \text{trace} \left\{ \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_k} \bar{\mathbf{R}}^{-1} \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \bar{\mathbf{R}}^{-1} \frac{\partial \bar{\mathbf{R}}}{\partial \alpha_l} \bar{\mathbf{R}}^{-1} \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \bar{\mathbf{R}}^{-1} \right\} \\ &= \left(\frac{\partial \bar{\mathbf{r}}}{\partial \alpha_k} \right)^H \left[\left(\bar{\mathbf{R}}^{-1} \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \bar{\mathbf{R}}^{-1} \right)^T \right. \\ &\quad \left. \otimes \left(\bar{\mathbf{R}}^{-1} \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \bar{\mathbf{R}}^{-1} \right) \right] \left(\frac{\partial \bar{\mathbf{r}}}{\partial \alpha_l} \right), \end{aligned} \quad (\text{B.14})$$

where $\bar{\mathbf{r}} = \text{vec}\{\bar{\mathbf{R}}\}$. The following equation [44] is used in the derivation of (B.14):

$$\text{trace}\{\mathbf{A}\mathbf{B}\mathbf{C}\mathbf{D}\} = (\text{vec}\{\mathbf{A}^T\})^T (\mathbf{D}^T \otimes \mathbf{B}) \text{vec}\{\mathbf{C}\}, \quad (\text{B.15})$$

which holds for any matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{D} that can make the product \mathbf{ABCD} .

Thus, from (B.14) we can formulate \mathbf{F}_b as

$$\begin{aligned}
\mathbf{F}_b &= \left(\frac{\partial \bar{\mathbf{r}}}{\partial \boldsymbol{\alpha}^T} \right)^H \left[\left(\bar{\mathbf{R}}^{-1} \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \bar{\mathbf{R}}^{-1} \right)^T \otimes \left(\bar{\mathbf{R}}^{-1} \mathbf{A} \mathbf{b} \mathbf{b}^H \mathbf{A}^H \bar{\mathbf{R}}^{-1} \right) \right] \frac{\partial \bar{\mathbf{r}}}{\partial \boldsymbol{\alpha}^T} \\
&= \boldsymbol{\Phi}^H \left\{ \left(\check{\mathbf{A}} \mathbf{b} \mathbf{b}^H \check{\mathbf{A}}^H \right)^T \otimes \left(\check{\mathbf{A}} \mathbf{b} \mathbf{b}^H \check{\mathbf{A}}^H \right) \right\} \boldsymbol{\Phi} \\
&= \boldsymbol{\Phi}^H \left\{ \left[(\check{\mathbf{A}} \mathbf{b})^* \otimes (\check{\mathbf{A}} \mathbf{b}) \right] \left[(\check{\mathbf{A}} \mathbf{b})^T \otimes (\check{\mathbf{A}} \mathbf{b})^H \right] \right\} \boldsymbol{\Phi}, \tag{B.16}
\end{aligned}$$

where

$$\check{\mathbf{A}} = \bar{\mathbf{R}}^{-\frac{1}{2}} \mathbf{A}, \tag{B.17}$$

$$\boldsymbol{\Phi} = \left(\bar{\mathbf{R}}^{-\frac{T}{2}} \otimes \bar{\mathbf{R}}^{-\frac{1}{2}} \right) \frac{\partial \bar{\mathbf{r}}}{\partial \boldsymbol{\alpha}^T}. \tag{B.18}$$

Let $\boldsymbol{\alpha} = [\boldsymbol{\theta}^T, \boldsymbol{\rho}^T, \mathbf{q}^T]^T$, where $\boldsymbol{\rho}$ and \mathbf{q} are real column vectors containing the unknown parameters from \mathbf{R}_{xx} and \mathbf{Q} , respectively. We partition $\boldsymbol{\Phi}$ into [43]

$$\boldsymbol{\Phi} = [\mathbf{U} | \boldsymbol{\Psi}] = \left(\bar{\mathbf{R}}^{-\frac{T}{2}} \otimes \bar{\mathbf{R}}^{-\frac{1}{2}} \right) \left[\frac{\partial \bar{\mathbf{r}}}{\partial \boldsymbol{\theta}^T} \middle| \frac{\partial \bar{\mathbf{r}}}{\partial \boldsymbol{\rho}^T}, \frac{\partial \bar{\mathbf{r}}}{\partial \mathbf{q}^T} \right], \tag{B.19}$$

where $\boldsymbol{\Psi}$ can be further partitioned into

$$\boldsymbol{\Psi} = [\mathbf{V} | \mathbf{W}] = \left(\bar{\mathbf{R}}^{-\frac{T}{2}} \otimes \bar{\mathbf{R}}^{-\frac{1}{2}} \right) \left[\frac{\partial \bar{\mathbf{r}}}{\partial \boldsymbol{\rho}^T} \middle| \frac{\partial \bar{\mathbf{r}}}{\partial \mathbf{q}^T} \right]. \tag{B.20}$$

Similarly, we partition \mathbf{F}_{ss} into [43]

$$\mathbf{F}_{ss} = \begin{bmatrix} \mathbf{U}^H \\ \boldsymbol{\Psi}^H \end{bmatrix} [\mathbf{U}, \boldsymbol{\Psi}] = \begin{bmatrix} \mathbf{U}^H \mathbf{U} & \mathbf{U}^H \boldsymbol{\Psi} \\ \boldsymbol{\Psi}^H \mathbf{U} & \boldsymbol{\Psi}^H \boldsymbol{\Psi} \end{bmatrix}. \tag{B.21}$$

According to the results from (B.16)-(B.21) and after some matrix manipulations, we can express $\mathbf{C}_{b,\theta}$ as

$$\mathbf{C}_{b,\theta} = \mathbf{\Upsilon}\mathbf{\Upsilon}^H, \quad (\text{B.22})$$

where

$$\mathbf{\Upsilon} = (\mathbf{U}^H \mathbf{\Pi}_{\check{\Psi}}^{\perp} \mathbf{U})^{-1} \mathbf{U}^H \mathbf{\Pi}_{\check{\Psi}}^{\perp} \left[(\check{\mathbf{A}}\mathbf{b})^* \otimes (\check{\mathbf{A}}\mathbf{b}) \right]. \quad (\text{B.23})$$

From the result in [43], we have

$$\mathbf{\Pi}_{\check{\Psi}}^{\perp} = \mathbf{\Pi}_{\check{\mathbf{V}}}^{\perp} - \mathbf{\Pi}_{\check{\mathbf{V}}}^{\perp} \mathbf{W} [\mathbf{W}^H \mathbf{\Pi}_{\check{\mathbf{V}}}^{\perp} \mathbf{W}]^{-1} \mathbf{W}^H \mathbf{\Pi}_{\check{\mathbf{V}}}^{\perp}, \quad (\text{B.24})$$

where $\mathbf{\Pi}_{\check{\mathbf{V}}}^{\perp}$ can be expressed as

$$\mathbf{\Pi}_{\check{\mathbf{V}}}^{\perp} = \mathbf{\Pi}_{\check{\mathbf{A}}^* \otimes \check{\mathbf{A}}}^{\perp} = \mathbf{I} \otimes \mathbf{\Pi}_{\check{\mathbf{A}}}^{\perp} + \mathbf{\Pi}_{\check{\mathbf{A}}^*}^{\perp} \otimes \mathbf{I} - \mathbf{\Pi}_{\check{\mathbf{A}}^*}^{\perp} \otimes \mathbf{\Pi}_{\check{\mathbf{A}}}^{\perp}. \quad (\text{B.25})$$

Then we obtain

$$\begin{aligned} & \mathbf{\Pi}_{\check{\mathbf{V}}}^{\perp} \left[(\check{\mathbf{A}}\mathbf{b})^* \otimes (\check{\mathbf{A}}\mathbf{b}) \right] \\ &= [\mathbf{I} \otimes \mathbf{\Pi}_{\check{\mathbf{A}}}^{\perp} + \mathbf{\Pi}_{\check{\mathbf{A}}^*}^{\perp} \otimes \mathbf{I} - \mathbf{\Pi}_{\check{\mathbf{A}}^*}^{\perp} \otimes \mathbf{\Pi}_{\check{\mathbf{A}}}^{\perp}] \left[(\check{\mathbf{A}}\mathbf{b})^* \otimes (\check{\mathbf{A}}\mathbf{b}) \right] \\ &= (\check{\mathbf{A}}\mathbf{b})^* \otimes (\mathbf{\Pi}_{\check{\mathbf{A}}}^{\perp} \check{\mathbf{A}}\mathbf{b}) + (\mathbf{\Pi}_{\check{\mathbf{A}}^*}^{\perp} \check{\mathbf{A}}^* \mathbf{b}^*) \otimes (\check{\mathbf{A}}\mathbf{b}) \\ & \quad - (\mathbf{\Pi}_{\check{\mathbf{A}}^*}^{\perp} \check{\mathbf{A}}^* \mathbf{b}^*) \otimes (\mathbf{\Pi}_{\check{\mathbf{A}}}^{\perp} \check{\mathbf{A}}\mathbf{b}) = \mathbf{0}. \end{aligned} \quad (\text{B.26})$$

Therefore we have $\mathbf{\Upsilon} = \mathbf{0}$ and $\mathbf{C}_{b,\theta} = \mathbf{0}$. The theorem is proved.

Appendix C

Proof of Lemma 2

From Theorem 1 in [46], we have

$$\begin{aligned} \mathbb{E}\{\mathbf{x}^H \mathbf{B} \mathbf{x} \mathbf{x}^H \mathbf{D} \mathbf{x}\} &= \mathbb{E}\{\mathbf{x}^H \mathbf{B} \mathbf{x}\} \mathbb{E}\{\mathbf{x}^H \mathbf{D} \mathbf{x}\} \\ &+ \mathbb{E}\{\mathbf{x}^H \otimes \mathbf{x}^H\} \mathbb{E}\{(\mathbf{D} \mathbf{x}) \otimes (\mathbf{B} \mathbf{x})\} \\ &+ \mathbb{E}\{\mathbf{x}^H \mathbb{E}\{\mathbf{B} \mathbf{x} \mathbf{x}^H\} \mathbf{D} \mathbf{x}\} - 2\mathbb{E}\{\mathbf{x}^H\} \mathbb{E}\{\mathbf{B} \mathbf{x}\} \mathbb{E}\{\mathbf{x}^H\} \mathbb{E}\{\mathbf{D} \mathbf{x}\}. \end{aligned} \quad (\text{C.1})$$

Note that

$$\mathbb{E}\{\mathbf{x}^H \mathbf{B} \mathbf{x}\} = \text{trace}\{\mathbf{B}(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H)\}, \quad (\text{C.2})$$

$$\begin{aligned} \mathbb{E}\{\mathbf{x}^H \otimes \mathbf{x}^H\} \mathbb{E}\{(\mathbf{B} \mathbf{x}) \otimes (\mathbf{D} \mathbf{x})\} &= (\boldsymbol{\mu}^H \otimes \boldsymbol{\mu}^H) \\ &\times [(\mathbf{B} \boldsymbol{\mu}) \otimes (\mathbf{D} \boldsymbol{\mu})] = \boldsymbol{\mu}^H \mathbf{B} \boldsymbol{\mu} \boldsymbol{\mu}^H \mathbf{D} \boldsymbol{\mu}, \end{aligned} \quad (\text{C.3})$$

$$\begin{aligned} \mathbb{E}\{\mathbf{x}^H \mathbb{E}\{\mathbf{B} \mathbf{x} \mathbf{x}^H\} \mathbf{D} \mathbf{x}\} &= \mathbb{E}\{\mathbf{x}^H \mathbf{B}(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H) \mathbf{D} \mathbf{x}\} \\ &= \text{trace}\{\mathbf{B}(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H) \mathbb{E}\{\mathbf{D} \mathbf{x} \mathbf{x}^H\}\} \\ &= \text{trace}\{\mathbf{B}(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H) \mathbf{D}(\mathbf{C} + \boldsymbol{\mu} \boldsymbol{\mu}^H)\}. \end{aligned} \quad (\text{C.4})$$

Substituting (C.2), (C.3), and (C.4) into (C.1), we obtain (2.92).

Appendix D

Proof of Proposition 5

For 1D DOA estimation of mixed signals, according to the result in [19], we have

$$\begin{aligned}
 \mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}}^{-1} &= 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \mathbf{R}_{xx}^T \right\} \\
 &= 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot (\mathbf{b}\mathbf{b}^H)^T \right\} \\
 &\quad + 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \mathbf{P}^T \right\} \\
 &= \mathbf{CRB}_{\mathbf{d},\boldsymbol{\theta}}^{-1} + 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \mathbf{P}^T \right\}. \tag{D.1}
 \end{aligned}$$

Note that $2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \mathbf{P}^T \right\}$ is the inverse of the asymptotic deterministic CRB on DOA estimation with signal and noise correlation matrices equal to \mathbf{P} and \mathbf{Q} . In [43], it was shown that

$$\mathbf{CRB}_{\mathbf{s},\boldsymbol{\theta}}^{-1} \leq 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \mathbf{P}^T \right\}. \tag{D.2}$$

According to (2.79) and (D.2), we obtain (2.94) for 1D DOA estimation.

For 2D DOA estimation, we have

$$\mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}}^{-1} = \mathbf{CRB}_{\mathbf{d},\boldsymbol{\theta}}^{-1} + 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot (\mathbf{P}^T \otimes \boldsymbol{\Delta}) \right\} \tag{D.3}$$

According to Proposition 2, we have

$$\mathbf{CRB}_{\mathbf{s},\boldsymbol{\theta}} \geq \frac{1}{N} \boldsymbol{\Omega}^{-1}, \quad (\text{D.4})$$

where

$$\begin{aligned} \frac{1}{N} \boldsymbol{\Omega}^{-1} &= \frac{1}{2N} \text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \left[\left(\mathbf{P} \tilde{\mathbf{A}}^H \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{A}} \mathbf{P} \right)^T \otimes \boldsymbol{\Delta} \right] \right\}^{-1} \\ &\geq \frac{1}{2N} \text{Re} \left\{ \left(\tilde{\mathbf{D}}_2^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}}_2 \right) \odot \left(\mathbf{P}^T \otimes \boldsymbol{\Delta} \right) \right\}^{-1}, \end{aligned} \quad (\text{D.5})$$

since $\mathbf{P} \geq \mathbf{P} \tilde{\mathbf{A}}^H \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{A}} \mathbf{P}$ [8]. Thus we get (2.94) for 2D DOA estimation.

Appendix E

Proof of Proposition 6

The first inequality in (2.95) was proved for 1D DOA estimation in [8]. For the 2D case, since both \mathbf{R}_{xx} and $\mathbf{R}_{xx} + \sigma^2(\mathbf{A}^H \mathbf{A})^{-1}$ are nonsingular, we have

$$\mathcal{N}\{\mathbf{R}_{xx}^T \otimes \Delta\} = \mathcal{N}\left\{\left(\mathbf{R}_{xx} + \sigma^2(\mathbf{A}^H \mathbf{A})^{-1}\right)^T \otimes \Delta\right\}. \quad (\text{E.1})$$

Note that

$$\left[\left(\mathbf{R}_{xx} + \sigma^2(\mathbf{A}^H \mathbf{A})^{-1}\right)^T \otimes \Delta\right]^\dagger = \left(\mathbf{R}_{xx} + \sigma^2(\mathbf{A}^H \mathbf{A})^{-1}\right)^{-T} \otimes \Delta^\dagger. \quad (\text{E.2})$$

According to Proposition 3 and Lemma 3, we have

$$\begin{aligned} \mathbf{AC}_{\text{D},\theta} &\geq \frac{\sigma^2}{2N} \left\{ \text{Re} \left\{ \left(\mathbf{D}_2^H \mathbf{\Pi}_A^\perp \mathbf{D}_2 \right) \odot \left[\left(\mathbf{R}_{xx}^{-1} \right. \right. \right. \right. \\ &\quad \left. \left. \left. \left. + \sigma^2 \mathbf{R}_{xx}^{-1} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{R}_{xx}^{-1} \right)^{-T} \otimes \Delta \right] \right\} \right\}^{-1} = \mathbf{AC}_{\text{S},\theta}. \end{aligned} \quad (\text{E.3})$$

The last equality in (E.3) holds from (2.88) and the results in [8].

We now prove the second inequality in (2.95) for 1D DOA estimation. The 2D case can be proved in a similar way. When $\mathbf{Q} = \sigma^2 \mathbf{I}$, we have the asymptotic deterministic

CRB as [7]

$$\mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}} = \frac{\sigma^2}{2N} \text{Re} \left\{ (\mathbf{D}_2^H \boldsymbol{\Pi}_A^\perp \mathbf{D}_2) \odot \mathbf{R}_{xx}^T \right\}^{-1}. \quad (\text{E.4})$$

According to the results in Theorems 1 and 2, we obtain

$$\begin{aligned} \mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}}^{-1} - \mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}^{-1} &= \frac{2N}{\sigma^2} \text{Re} \left\{ (\mathbf{D}^H \boldsymbol{\Pi}_A^\perp \mathbf{D}) \right. \\ &\quad \left. \odot \left[\mathbf{P} - \mathbf{P} \mathbf{A}^H (\mathbf{A} \mathbf{P} \mathbf{A}^H + \sigma^2 \mathbf{I})^{-1} \mathbf{A} \mathbf{P} \right]^T \right\} \\ &= 2N \text{Re} \left\{ (\mathbf{D}^H \boldsymbol{\Pi}_A^\perp \mathbf{D}) \odot (\sigma^2 \mathbf{P}^{-1} + \mathbf{A}^H \mathbf{A})^{-T} \right\}, \end{aligned} \quad (\text{E.5})$$

$$\begin{aligned} \mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}}^{-1} - \mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}^{-1} &= \frac{2N}{\sigma^2} \text{Re} \left\{ (\mathbf{D}^H \boldsymbol{\Pi}_A^\perp \mathbf{D}) \odot \left[\mathbf{R}_{xx} \right. \right. \\ &\quad \left. \left. - \mathbf{R}_{xx} \mathbf{A}^H (\mathbf{A} \mathbf{R}_{xx} \mathbf{A}^H + \sigma^2 \mathbf{I})^{-1} \mathbf{A} \mathbf{R}_{xx} \right]^T \right\} \\ &= 2N \text{Re} \left\{ (\mathbf{D}^H \boldsymbol{\Pi}_A^\perp \mathbf{D}) \odot (\sigma^2 \mathbf{R}_{xx}^{-1} + \mathbf{A}^H \mathbf{A})^{-T} \right\}. \end{aligned} \quad (\text{E.6})$$

The second equalities in (E.5) and (E.6) hold from the results in [8].

For mixed signals, we have

$$\mathbf{R}_{xx} = \mathbf{P} + \mathbf{b} \mathbf{b}^H \geq \mathbf{P}. \quad (\text{E.7})$$

Thus, we have

$$\mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}}^{-1} - \mathbf{CRB}_{\mathbf{M},\boldsymbol{\theta}}^{-1} \leq \mathbf{ACRB}_{\mathbf{D},\boldsymbol{\theta}}^{-1} - \mathbf{AC}_{\mathbf{S},\boldsymbol{\theta}}^{-1}, \quad (\text{E.8})$$

from which we obtain the second inequality in (2.95).

Appendix F

Proof of Lemma 3

We first present the following lemma [47] before proceeding to the proof.

Lemma 4. *Suppose a Hermitian matrix*

$$\mathbf{W} = \begin{bmatrix} \mathbf{T} & \mathbf{S} \\ \mathbf{S}^H & \mathbf{D} \end{bmatrix} \quad (\text{F.1})$$

is partitioned symmetrically such that \mathbf{T} and \mathbf{D} are also Hermitian matrices. Then $\mathbf{W} \geq \mathbf{0}$ if and only if $\mathbf{T} \geq \mathbf{0}$, $\mathbf{D} - \mathbf{S}^H \mathbf{T}^\dagger \mathbf{S} \geq \mathbf{0}$, and $\mathcal{N}\{\mathbf{T}\} \subseteq \mathcal{N}\{\mathbf{S}^H\}$.

Using Lemma 4, we can easily show that the matrix

$$\begin{bmatrix} \mathbf{C} & \mathbf{B} \\ \mathbf{B} & \mathbf{B}\mathbf{C}^\dagger\mathbf{B} \end{bmatrix} \geq \mathbf{0}. \quad (\text{F.2})$$

As a result,

$$\text{Re} \left\{ \begin{bmatrix} \mathbf{A} \odot \mathbf{C} & \mathbf{A} \odot \mathbf{B} \\ \mathbf{A} \odot \mathbf{B} & \mathbf{A} \odot (\mathbf{B}\mathbf{C}^\dagger\mathbf{B}) \end{bmatrix} \right\} = \text{Re} \left\{ \begin{bmatrix} \mathbf{A} & \mathbf{A} \\ \mathbf{A} & \mathbf{A} \end{bmatrix} \odot \begin{bmatrix} \mathbf{C} & \mathbf{B} \\ \mathbf{B} & \mathbf{B}\mathbf{C}^\dagger\mathbf{B} \end{bmatrix} \right\} \geq \mathbf{0}. \quad (\text{F.3})$$

Thus, it holds that

$$\operatorname{Re}\{\mathbf{A} \odot (\mathbf{B}\mathbf{C}^\dagger\mathbf{B})\} - \operatorname{Re}\{\mathbf{A} \odot \mathbf{B}\} \{\operatorname{Re}\{\mathbf{A} \odot \mathbf{C}\}\}^{-1} \operatorname{Re}\{\mathbf{A} \odot \mathbf{B}\} \geq \mathbf{0}, \quad (\text{F.4})$$

from which we obtain (2.96).

Appendix G

Proof of Equation (3.8)

Using the Woodbury matrix identity [60], we have

$$(\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1} = \mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{A}(\mathbf{P}^{-1} + \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}, \quad (\text{G.1})$$

from which we have

$$\begin{aligned} \mathbf{A}^H(\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1} &= \mathbf{A}^H\mathbf{Q}^{-1} - \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A} \\ &\quad \times (\mathbf{P}^{-1} + \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1} \\ &= \mathbf{A}^H\mathbf{Q}^{-1} - \left[\mathbf{P} + (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1} \right]^{-1} \mathbf{P}\mathbf{A}^H\mathbf{Q}^{-1}, \end{aligned} \quad (\text{G.2})$$

and

$$\begin{aligned} \mathbf{A}^H(\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q})^{-1}\mathbf{A} &= \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A} - \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A} \\ &\quad \times (\mathbf{P}^{-1} + \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A} \\ &= \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A}(\mathbf{P}^{-1} + \mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{P}^{-1} \\ &= \left[\mathbf{P} + (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1} \right]^{-1}. \end{aligned} \quad (\text{G.3})$$

We thus have

$$\begin{aligned}
& \left[\mathbf{A}^H (\mathbf{A} \mathbf{P} \mathbf{A}^H + \mathbf{Q})^{-1} \mathbf{A} \right]^{-1} \mathbf{A}^H (\mathbf{A} \mathbf{P} \mathbf{A}^H + \mathbf{Q})^{-1} \\
&= \left[\mathbf{P} + (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1} \right] \left\{ \mathbf{I}_L - \left[\mathbf{P} + (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1} \right]^{-1} \mathbf{P} \right\} \mathbf{A}^H \mathbf{Q}^{-1} \\
&= (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{Q}^{-1}. \tag{G.4}
\end{aligned}$$

Appendix H

Proof of Proposition 8

Let $\boldsymbol{\alpha} = [\boldsymbol{\theta}^T, \boldsymbol{\delta}^T, \mathbf{b}^T]^T$, where $\boldsymbol{\delta} = [\mathbf{p}^T, \boldsymbol{\sigma}^T]^T$, and $\mathbf{b} = [\mathbf{c}^T, \boldsymbol{\omega}^T]^T$. According to the Fisher information matrix equation [42], we have

$$[\mathbf{FIM}_M]_{mn} = 2\text{Re}\left\{\sum_{t=1}^N \frac{\partial \boldsymbol{\mu}^H(t)}{\partial \alpha_m} \mathbf{R}^{-1} \frac{\partial \boldsymbol{\mu}(t)}{\partial \alpha_n}\right\} + N\text{trace}\left\{\mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \alpha_m} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \alpha_n}\right\}, \quad (\text{H.1})$$

where \mathbf{FIM}_M is the Fisher information matrix for $\boldsymbol{\alpha}$, $[\mathbf{FIM}_M]_{mn}$ is the (m, n) -th element of \mathbf{FIM}_M , $\text{Re}\{\cdot\}$ denotes the real part of a complex number, $\boldsymbol{\mu}(t) = \mathbf{A}\mathbf{C}\boldsymbol{\varphi}(t)$, and α_m is the m -th element of $\boldsymbol{\alpha}$.

Let \mathbf{F}_D be the matrix of the same size as \mathbf{FIM}_M , with $[\mathbf{F}_D]_{mn}$ equal to the first term on the right side of (H.1). If α_m is a parameter from $\boldsymbol{\delta}$, we have $\frac{\partial \boldsymbol{\mu}(t)}{\partial \alpha_m} = \mathbf{0}$. Therefore,

$$\mathbf{F}_D = \begin{bmatrix} \mathbf{F}_{D.\theta\theta} & \mathbf{F}_{D.\theta\delta} & \mathbf{F}_{D.\theta b} \\ \mathbf{F}_{D.\delta\theta} & \mathbf{F}_{D.\delta\delta} & \mathbf{F}_{D.\delta b} \\ \mathbf{F}_{D.b\theta} & \mathbf{F}_{D.b\delta} & \mathbf{F}_{D.bb} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{D.\theta\theta} & \mathbf{0} & \mathbf{F}_{D.\theta b} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{F}_{D.b\theta} & \mathbf{0} & \mathbf{F}_{D.bb} \end{bmatrix}, \quad (\text{H.2})$$

where $\mathbf{F}_{D.\theta\theta}, \mathbf{F}_{D.\theta\delta}, \dots, \mathbf{F}_{D.bb}$ are the sub-matrices of \mathbf{F}_D containing the elements related to the parameters specified in their subscripts.

Similarly, let \mathbf{F}_S be the matrix of the same size as \mathbf{FIM}_M , with $[\mathbf{F}_S]_{mn}$ equal to the second term on the right side of (H.1). Since

$$\frac{\partial \mathbf{R}}{\partial \alpha_m} = \mathbf{0} \quad (\text{H.3})$$

if α_m is a parameter from \mathbf{b} , we have

$$\mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{S.\theta\theta} & \mathbf{F}_{S.\theta\delta} & \mathbf{F}_{S.\theta b} \\ \mathbf{F}_{S.\delta\theta} & \mathbf{F}_{S.\delta\delta} & \mathbf{F}_{S.\delta b} \\ \mathbf{F}_{S.b\theta} & \mathbf{F}_{S.b\delta} & \mathbf{F}_{S.bb} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{S.\theta\theta} & \mathbf{F}_{S.\theta\delta} & \mathbf{0} \\ \mathbf{F}_{S.\delta\theta} & \mathbf{F}_{S.\delta\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (\text{H.4})$$

From the results in (H.2) and (H.4), we obtain

$$\mathbf{FIM}_M = \mathbf{F}_D + \mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{D.\theta\theta} + \mathbf{F}_{S.\theta\theta} & \mathbf{F}_{S.\theta\delta} & \mathbf{F}_{D.\theta b} \\ \mathbf{F}_{S.\delta\theta} & \mathbf{F}_{S.\delta\delta} & \mathbf{0} \\ \mathbf{F}_{D.b\theta} & \mathbf{0} & \mathbf{F}_{D.bb} \end{bmatrix}, \quad (\text{H.5})$$

from which we have

$$\begin{aligned} \mathbf{CRB}_{M,\theta}^{-1} &= \mathbf{F}_{D.\theta\theta} + \mathbf{F}_{S.\theta\theta} - [\mathbf{F}_{S.\theta\delta} \quad \mathbf{F}_{D.\theta b}] \\ &\quad \times \begin{bmatrix} \mathbf{F}_{S.\delta\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{D.bb} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{F}_{S.\delta\theta} \\ \mathbf{F}_{D.b\theta} \end{bmatrix} \\ &= \mathbf{F}_{D.\theta\theta} - \mathbf{F}_{D.\theta b} \mathbf{F}_{D.bb}^{-1} \mathbf{F}_{D.b\theta} \\ &\quad + \mathbf{F}_{S.\theta\theta} - \mathbf{F}_{S.\theta\delta} \mathbf{F}_{S.\delta\delta}^{-1} \mathbf{F}_{S.\delta\theta}, \end{aligned} \quad (\text{H.6})$$

where

$$\mathbf{F}_{S.\theta\theta} - \mathbf{F}_{S.\theta\delta} \mathbf{F}_{S.\delta\delta}^{-1} \mathbf{F}_{S.\delta\theta} = \mathbf{CRB}_{S,\theta}^{-1}. \quad (\text{H.7})$$

Using the result in [57], we have

$$\begin{aligned}
& \mathbf{F}_{\mathbf{D},\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{D},\boldsymbol{\theta}\mathbf{b}}\mathbf{F}_{\mathbf{D},\mathbf{b}\mathbf{b}}^{-1}\mathbf{F}_{\mathbf{D},\mathbf{b}\boldsymbol{\theta}} \\
&= 2N\text{Re} \left\{ \left(\bar{\mathbf{D}}^H \boldsymbol{\Pi}_{\bar{\mathbf{A}}}^\perp \bar{\mathbf{D}} \right) \odot \left(\mathbf{C} \hat{\mathbf{R}}_{\phi\phi} \mathbf{C}^H \right)^T \right\} \\
&= 2N\text{Re} \left\{ \left(\tilde{\mathbf{D}}^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}} \right) \odot \left(\mathbf{C} \hat{\mathbf{R}}_{\phi\phi} \mathbf{C}^H \right)^T \right\} \\
&= \mathbf{C}\mathbf{R}\mathbf{B}_{\mathbf{D},\boldsymbol{\theta}}^{-1}, \tag{H.8}
\end{aligned}$$

where $\bar{\mathbf{D}} = \mathbf{R}^{-\frac{1}{2}}\mathbf{D}$, $\bar{\mathbf{A}} = \mathbf{R}^{-\frac{1}{2}}\mathbf{A}$, $\tilde{\mathbf{D}} = \mathbf{Q}^{-\frac{1}{2}}\mathbf{D}$, and $\hat{\mathbf{R}}_{\phi\phi} = \frac{1}{N} \sum_{t=1}^N \boldsymbol{\varphi}(t)\boldsymbol{\varphi}^H(t)$, and \odot denotes the Hadamard product. The second equality in (H.8) holds from the following result given in [43]:

$$\mathbf{R}^{-\frac{1}{2}}\boldsymbol{\Pi}_{\bar{\mathbf{A}}}^\perp\mathbf{R}^{-\frac{1}{2}} = \mathbf{Q}^{-\frac{1}{2}}\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp\mathbf{Q}^{-\frac{1}{2}}. \tag{H.9}$$

Substituting (H.7) and (H.8) into (H.6), we obtain (3.30).

Appendix I

Proof of Proposition 9

We first give a proof that the ML estimation is consistent, which we believe is necessary since the typical conclusion (see [42]) that the ML estimation is consistent does not straightforwardly hold here due to the temporally non-stationary measurements.

Let $\boldsymbol{\alpha}_0 = [\boldsymbol{\theta}_0^T, \boldsymbol{\delta}_0^T, \mathbf{c}_0^T, \boldsymbol{\omega}_0^T]^T$ be the vector of true parameter values. Also we let $\mathbf{A}_0 = \mathbf{A}(\boldsymbol{\theta}_0)$ and $\boldsymbol{\phi}_0 = \boldsymbol{\phi}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$. According to (3.11), we have the LL function as

$$L(\boldsymbol{\alpha}) = -\log |\mathbf{R}| - \text{trace} \{ \mathbf{R}^{-1} \mathbf{R}_{yy} \}, \quad (\text{I.1})$$

where

$$\begin{aligned} \mathbf{R}_{yy} &= \frac{1}{N} [\mathbf{Y} - \mathbf{A}\mathbf{C}\boldsymbol{\phi}][\mathbf{Y} - \mathbf{A}\mathbf{C}\boldsymbol{\phi}]^H \\ &= \frac{1}{N} [\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\boldsymbol{\phi}_0 + \mathbf{A}_0\mathbf{C}_0\boldsymbol{\phi}_0 - \mathbf{A}\mathbf{C}\boldsymbol{\phi}] \\ &\quad \times [\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\boldsymbol{\phi}_0 + \mathbf{A}_0\mathbf{C}_0\boldsymbol{\phi}_0 - \mathbf{A}\mathbf{C}\boldsymbol{\phi}]^H \\ &= \mathbf{R}_{yy,0} + \mathbf{R}_{yy,1} + \mathbf{R}_{yy,2} + \mathbf{R}_{yy,2}^H, \end{aligned} \quad (\text{I.2})$$

in which

$$\begin{aligned}
\mathbf{R}_{yy,0} &= \frac{1}{N}[\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\phi_0][\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\phi_0]^H, \\
\mathbf{R}_{yy,1} &= \frac{1}{N}[\mathbf{A}_0\mathbf{C}_0\phi_0 - \mathbf{A}\mathbf{C}\phi][\mathbf{A}_0\mathbf{C}_0\phi_0 - \mathbf{A}\mathbf{C}\phi]^H, \\
\mathbf{R}_{yy,2} &= \frac{1}{N}[\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\phi_0][\mathbf{A}_0\mathbf{C}_0\phi_0 - \mathbf{A}\mathbf{C}\phi]^H.
\end{aligned} \tag{I.3}$$

Let \mathbf{r}_m be the m -th column of the matrix $\frac{1}{N}[\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\phi_0][\mathbf{A}_0\mathbf{C}_0\phi_0]^H$, which can be expressed as

$$\mathbf{r}_m = \frac{1}{N} \sum_{t=1}^N [\mathbf{y}(t) - \mathbf{A}_0\mathbf{C}_0\phi_0(t)] \boldsymbol{\nu}_m^T \boldsymbol{\varphi}_0^*(t), \tag{I.4}$$

where $\boldsymbol{\nu}_m$ is the m -th column of $(\mathbf{A}_0\mathbf{C}_0)^H$. Note that the auto-covariance matrix of $\mathbf{y}(t) - \mathbf{A}_0\mathbf{C}_0\phi_0(t)$ is \mathbf{R}_0 , the true value of \mathbf{R} . Therefore, we have

$$\text{cov}\{\mathbf{r}_m\} = \mathbf{R}_0 \boldsymbol{\nu}_m^T \frac{\sum_{t=1}^N \boldsymbol{\varphi}_0^*(t) \boldsymbol{\varphi}_0^T(t)}{N^2} \boldsymbol{\nu}_m. \tag{I.5}$$

Since $\frac{\sum_{t=1}^N \boldsymbol{\varphi}_0^*(t) \boldsymbol{\varphi}_0^T(t)}{N^2} \rightarrow \mathbf{0}$ as $N \rightarrow \infty$, we have

$$\frac{1}{N}[\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\phi_0][\mathbf{A}_0\mathbf{C}_0\phi_0]^H \rightarrow \mathbf{0} \tag{I.6}$$

in probability when $N \rightarrow \infty$. Similarly, we can show

$$\frac{1}{N}[\mathbf{Y} - \mathbf{A}_0\mathbf{C}_0\phi_0][\mathbf{A}\mathbf{C}\phi]^H \rightarrow \mathbf{0} \tag{I.7}$$

in probability as $N \rightarrow \infty$. Therefore, we have

$$\mathbf{R}_{yy,2} \rightarrow \mathbf{0} \tag{I.8}$$

in probability as $N \rightarrow \infty$.

Since the limits of $\frac{1}{N}\phi_0\phi_0^H$, $\frac{1}{N}\phi\phi^H$, and $\frac{1}{N}\phi\phi_0^H$ exist when $N \rightarrow \infty$, the asymptotic value of $\mathbf{R}_{yy,1}$ also exists as $N \rightarrow \infty$. For instance, $\frac{1}{N}\phi_0\phi_0^H$ converges to \mathbf{I}_J as $N \rightarrow \infty$. We hence have the asymptotic value of $\mathbf{R}_{yy,1}$ when $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ as

$$\mathbf{R}_{yy,1}^{\text{as}} = [\mathbf{A}_0\mathbf{C}_0 - \mathbf{A}\mathbf{C}][\mathbf{A}_0\mathbf{C}_0 - \mathbf{A}\mathbf{C}]^H. \quad (\text{I.9})$$

As a result of (I.8), (I.9), and the fact that $\mathbf{R}_{yy,0} \rightarrow \mathbf{R}_0$ in probability as $N \rightarrow \infty$, we have

$$\mathbf{R}_{yy} \rightarrow \mathbf{R}_0 + \mathbf{R}_{yy,1}^{\text{as}} \quad (\text{I.10})$$

in probability as $N \rightarrow \infty$. We thus obtain $L(\boldsymbol{\alpha})$ converges in probability to its asymptotic value $L^{\text{as}}(\boldsymbol{\alpha})$ from the continuous mapping theorem [58]. Using the Lévy's Theorem [61], we have $L(\boldsymbol{\alpha})$ also converges with probability 1 to $L^{\text{as}}(\boldsymbol{\alpha})$.

Note that

$$\begin{aligned} L^{\text{as}}(\boldsymbol{\alpha}) &= -\log |\mathbf{R}| - \text{trace} \{ \mathbf{R}^{-1}(\mathbf{R}_0 + \mathbf{R}_{yy,1}^{\text{as}}) \} \\ &\leq -\log |\mathbf{R}_0 + \mathbf{R}_{yy,1}^{\text{as}}| - M, \end{aligned} \quad (\text{I.11})$$

where the inequality holds from the following Lemma 1 [37].

Lemma 5. *Let $\boldsymbol{\Sigma}$ be an $M \times M$ positive definite matrix. Then, for $a > 0$ and $b > 0$,*

$$|\mathbf{R}|^{-b} \exp\{-a \text{trace}\{\mathbf{R}^{-1}\boldsymbol{\Sigma}\}\} \leq |a\boldsymbol{\Sigma}/b|^{-b} \exp\{-Mb\} \quad (\text{I.12})$$

for all $M \times M$ positive definite matrices \mathbf{R} . The equality holds if and only if $\mathbf{R} = a\boldsymbol{\Sigma}/b$.

When $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$, we have $\mathbf{R} = \mathbf{R}_0$, $\mathbf{R}_{\mathbf{y}\mathbf{y},1}^{\text{as}} = \mathbf{0}$, and thus

$$L^{\text{as}}(\boldsymbol{\alpha}_0) = -\log |\mathbf{R}_0| - M \geq L^{\text{as}}(\boldsymbol{\alpha}), \quad (\text{I.13})$$

from which we can see that the asymptotic LL function achieves its maximum at $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$. Assuming a unique maximum point for the LL function, we obtain the ML estimation is consistent since the LL function is differentiable [62].

We now derive the asymptotic error covariance matrix of the mixed ML estimator. Let $a = a_1 + a_2$ be the length of $\boldsymbol{\alpha}$, and a_1 and a_2 be the lengths of $[\boldsymbol{\theta}^T, \boldsymbol{\delta}^T, \mathbf{c}^T]^T$ and $\boldsymbol{\omega}$, respectively. We first define a new matrix $\boldsymbol{\Gamma}_N$, which is a $a \times a$ diagonal matrix with the first a_1 diagonal elements equal to \sqrt{N} and the last a_2 diagonal elements equal to $N\sqrt{N}$.

Assuming $f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})$ is the pdf from the narrow-band measurement model and $\hat{\boldsymbol{\alpha}}$ is the ML estimate of $\boldsymbol{\alpha}$, we apply a Taylor expansion on $f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})$ around $\boldsymbol{\alpha}_0$ such that

$$\begin{aligned} \left. \frac{\partial \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}} &= \left. \frac{\partial \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \\ &+ \left. \frac{\partial^2 \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} \right|_{\boldsymbol{\alpha}=\tilde{\boldsymbol{\alpha}}} (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) = \mathbf{0}, \end{aligned} \quad (\text{I.14})$$

where $\tilde{\boldsymbol{\alpha}}$ is a vector value between $\boldsymbol{\alpha}_0$ and $\hat{\boldsymbol{\alpha}}$. Thus we have

$$\boldsymbol{\Gamma}_N(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) = - \left[\boldsymbol{\Gamma}_N^{-1} \frac{\partial^2 \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} \boldsymbol{\Gamma}_N^{-1} \right]^{-1} \bigg|_{\boldsymbol{\alpha}=\tilde{\boldsymbol{\alpha}}} \left[\boldsymbol{\Gamma}_N^{-1} \frac{\partial \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right] \bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}. \quad (\text{I.15})$$

Let

$$\boldsymbol{\gamma}_t = \boldsymbol{\Gamma}_N^{-1} \frac{\partial \ln f_{\mathbf{y}(t)}(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}. \quad (\text{I.16})$$

Using the results in [42], we have $E\{\boldsymbol{\gamma}_t\} = \mathbf{0}$ and

$$\begin{aligned} [\text{cov}\{\boldsymbol{\Gamma}_N \boldsymbol{\gamma}_t\}]_{mn} &= 2\text{Re} \left\{ \frac{\partial \boldsymbol{\mu}^H(t)}{\partial \alpha_m} \mathbf{R}^{-1} \frac{\partial \boldsymbol{\mu}(t)}{\partial \alpha_n} \right\} \Bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \\ &+ \text{trace} \left\{ \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \alpha_m} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \alpha_n} \right\} \Bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}, \end{aligned} \quad (\text{I.17})$$

We let $\mathbf{F}_d(t)$ and \mathbf{F}_s be matrices with $[\mathbf{F}_d(t)]_{mn}$ and $[\mathbf{F}_s]_{mn}$ equal to the first and the second terms on the right side of (I.17), respectively. For simplicity, we omit the subscript $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$ in the following equations.

The formula of \mathbf{F}_s , which is not dependent on $\boldsymbol{\omega}$, \mathbf{c} , and t , has been well addressed in [43]. We now consider the expression of $\mathbf{F}_d(t)$. Since

$$\frac{\boldsymbol{\mu}(t)}{\partial \omega_m} = \mathbf{A} \mathbf{c}_m \frac{\partial e^{j\omega_m t}}{\partial \omega_m} = \mathbf{A} \mathbf{c}_m \cdot j t e^{j\omega_m t}, \quad (\text{I.18})$$

where \mathbf{c}_m is the m -th column of \mathbf{C} , we have

$$\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{\omega}^T} = (\boldsymbol{\varphi}_d^T(t) \otimes \mathbf{A}) \check{\mathbf{C}}, \quad (\text{I.19})$$

where $\boldsymbol{\varphi}_d(t) = [j t e^{j\omega_1 t}, \dots, j t e^{j\omega_J t}]^T$, and $\check{\mathbf{C}} = \text{blkdiag}\{\mathbf{c}_1, \dots, \mathbf{c}_J\}$ is a block diagonal matrix with the J diagonal blocks equal to $\mathbf{c}_1, \dots, \mathbf{c}_J$. From the results in [57], we have

$$\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{\theta}^T} = [(\boldsymbol{\varphi}^T(t) \mathbf{C}^T) \otimes \mathbf{I}_M] \mathbf{D}_A, \quad (\text{I.20})$$

$$\frac{\partial \boldsymbol{\mu}(t)}{\partial \mathbf{c}^T} = [1, j] \otimes \boldsymbol{\varphi}^T(t) \otimes \mathbf{A}, \quad (\text{I.21})$$

where $\mathbf{D}_A = \partial \text{vec}\{\mathbf{A}\} / \partial \boldsymbol{\theta}^T$. Therefore,

$$\begin{aligned}
\mathbf{F}_{d.\omega\theta}(t) &= 2\text{Re}\left\{\check{\mathbf{C}}^H [\varphi_d^*(t) \otimes \mathbf{A}^H] \mathbf{R}^{-1} [(\varphi^T(t) \mathbf{C}^T) \otimes \mathbf{I}_M] \mathbf{D}_A\right\} \\
&= 2\text{Re}\left\{\check{\mathbf{C}}^H \left[(\mathbf{C}\varphi(t)\varphi_d^H(t))^T \otimes (\mathbf{A}^H \mathbf{R}^{-1})\right] \mathbf{D}_A\right\}, \tag{I.22}
\end{aligned}$$

$$\begin{aligned}
\mathbf{F}_{d.\omega c}(t) &= 2\text{Re}\left\{\check{\mathbf{C}}^H [\varphi_d^*(t) \otimes \mathbf{A}^H] \mathbf{R}^{-1} ([1, j] \otimes \varphi^T(t) \otimes \mathbf{A})\right\} \\
&= 2\text{Re}\left\{\check{\mathbf{C}}^H \left[[1, j] \otimes (\varphi(t)\varphi_d^H(t))^T \otimes (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})\right]\right\}, \tag{I.23}
\end{aligned}$$

$$\begin{aligned}
\mathbf{F}_{d.\omega\omega}(t) &= 2\text{Re}\left\{\check{\mathbf{C}}^H [\varphi_d^*(t) \otimes \mathbf{A}^H] \mathbf{R}^{-1} [\varphi_d^T(t) \otimes \mathbf{A}] \check{\mathbf{C}}\right\} \\
&= 2\text{Re}\left\{\check{\mathbf{C}}^H \left[(\varphi_d(t)\varphi_d^H(t))^T \otimes (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})\right] \check{\mathbf{C}}\right\}. \tag{I.24}
\end{aligned}$$

The expressions for $\mathbf{F}_{d.\theta\theta}(t)$, $\mathbf{F}_{d.\theta c}(t)$, and $\mathbf{F}_{d.cc}(t)$ are given in [57] as

$$\mathbf{F}_{d.\theta\theta}(t) = 2\text{Re}\left\{\mathbf{D}_A^H \left[(\mathbf{C}\varphi(t)\varphi^H(t)\mathbf{C}^H)^T \otimes \mathbf{R}^{-1}\right] \mathbf{D}_A\right\}, \tag{I.25}$$

$$\mathbf{F}_{d.\theta c}(t) = 2\text{Re}\left\{[1, j] \otimes \left[\mathbf{D}_A^H \left((\varphi(t)\varphi^H(t)\mathbf{C}^H)^T \otimes \mathbf{R}^{-1} \mathbf{A}\right)\right]\right\}, \tag{I.26}$$

$$\mathbf{F}_{d.cc}(t) = 2\text{Re}\left\{\begin{bmatrix} 1 & j \\ -j & 1 \end{bmatrix} \otimes \left[(\varphi(t)\varphi^H(t))^T \otimes (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})\right]\right\}. \tag{I.27}$$

From the results in (I.22)–(I.27), we have the covariance matrix of the second term on the right side of (I.15) as

$$\sum_{t=1}^N \text{cov}\{\gamma_t\} = \mathbf{F}_d + \mathbf{F}_s, \tag{I.28}$$

where

$$\mathbf{F}_d = \sum_{t=1}^N \mathbf{\Gamma}_N^{-1} \mathbf{F}_d(t) \mathbf{\Gamma}_N^{-1} = \begin{bmatrix} \mathbf{F}_{d.\theta\theta} & \mathbf{0} & \mathbf{F}_{d.\theta c} & \mathbf{F}_{d.\theta\omega} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{F}_{d.c\theta} & \mathbf{0} & \mathbf{F}_{d.cc} & \mathbf{F}_{d.c\omega} \\ \mathbf{F}_{d.\omega\theta} & \mathbf{0} & \mathbf{F}_{d.\omega c} & \mathbf{F}_{d.\omega\omega} \end{bmatrix}, \tag{I.29}$$

$\mathbf{F}_{d.\theta\theta} = \frac{1}{N} \sum_{t=1}^N \mathbf{F}_{d.\theta\theta}(t)$, $\mathbf{F}_{d.\theta c} = \frac{1}{N} \sum_{t=1}^N \mathbf{F}_{d.\theta c}(t)$, $\mathbf{F}_{d.cc} = \frac{1}{N} \sum_{t=1}^N \mathbf{F}_{d.cc}(t)$, $\mathbf{F}_{d.\omega\theta} = \frac{1}{N^2} \sum_{t=1}^N \mathbf{F}_{d.\omega\theta}(t)$, $\mathbf{F}_{d.\omega c} = \frac{1}{N^2} \sum_{t=1}^N \mathbf{F}_{d.\omega c}(t)$, and $\mathbf{F}_{d.\omega\omega} = \frac{1}{N^3} \sum_{t=1}^N \mathbf{F}_{d.\omega\omega}(t)$. We can see that the expressions for $\mathbf{F}_{d.\theta\theta}$, $\mathbf{F}_{d.\theta c}$, \dots , $\mathbf{F}_{d.\omega\omega}$ can be obtained by replacing $\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^H(t)$, $\boldsymbol{\varphi}(t)\boldsymbol{\varphi}_d^H(t)$, and $\boldsymbol{\varphi}_d(t)\boldsymbol{\varphi}_d^H(t)$ in (I.22)–(I.27) with $\hat{\mathbf{R}}_{\phi\phi} = \frac{1}{N} \sum_{t=1}^N \boldsymbol{\varphi}(t)\boldsymbol{\varphi}^H(t)$, $\hat{\mathbf{R}}_{\phi\phi_d} = \frac{1}{N^2} \sum_{t=1}^N \boldsymbol{\varphi}(t)\boldsymbol{\varphi}_d^H(t)$, and $\hat{\mathbf{R}}_{\phi_d\phi_d} = \frac{1}{N^3} \sum_{t=1}^N \boldsymbol{\varphi}_d(t)\boldsymbol{\varphi}_d^H(t)$, respectively.

As $N \rightarrow \infty$, we can easily show that $\hat{\mathbf{R}}_{\phi\phi}$, $\hat{\mathbf{R}}_{\phi\phi_d}$, and $\hat{\mathbf{R}}_{\phi_d\phi_d}$ all converge to finite matrices. Therefore, as N increases, \mathbf{F}_d converges to a finite matrix \mathbf{F}_d^{as} and

$$\sum_{t=1}^N \text{cov}\{\boldsymbol{\gamma}_t\} \rightarrow \mathbf{F}_d^{\text{as}} + \mathbf{F}_s. \quad (\text{I.30})$$

Let $\check{\boldsymbol{\gamma}}_t = \sqrt{N}\boldsymbol{\gamma}_t$. Examining the results in (I.22)–(I.27), we can see that $\text{cov}\{\check{\boldsymbol{\gamma}}_1\}, \dots, \text{cov}\{\check{\boldsymbol{\gamma}}_N\}$ remain finite as N increases. Therefore, for arbitrary $\epsilon > 0$, we have as $N \rightarrow \infty$,

$$\sum_{t=1}^N \int_{\|\boldsymbol{\gamma}_t\| > \epsilon} \|\boldsymbol{\gamma}_t\|^2 f_{\boldsymbol{\gamma}_t}(\boldsymbol{\gamma}_t; \boldsymbol{\alpha}) d\boldsymbol{\gamma}_t = \frac{1}{N} \sum_{t=1}^N \int_{\|\check{\boldsymbol{\gamma}}_t\| > \sqrt{N}\epsilon} \|\check{\boldsymbol{\gamma}}_t\|^2 f_{\check{\boldsymbol{\gamma}}_t}(\check{\boldsymbol{\gamma}}_t; \boldsymbol{\alpha}) d\check{\boldsymbol{\gamma}}_t \rightarrow 0. \quad (\text{I.31})$$

From (I.30) and (I.31), we can see that the sequence $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_N$ satisfies the conditions in Theorem 3. As a consequence, we have $\sum_{t=1}^N \boldsymbol{\gamma}_t$, which is the second term on the right side of (I.15), asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_d^{\text{as}} + \mathbf{F}_s)$.

We now consider the first term on the right side of (I.15). Using the result in [42], we have

$$\begin{aligned} \frac{\partial \ln f_{\mathbf{y}(t)}(\mathbf{y}(t); \boldsymbol{\alpha})}{\partial \alpha_m} &= -\text{trace} \left\{ \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \alpha_m} \right\} + [\mathbf{y}(t) - \boldsymbol{\mu}(t)]^H \\ &\quad \times \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \alpha_m} \mathbf{R}^{-1} [\mathbf{y}(t) - \boldsymbol{\mu}(t)] + [\mathbf{y}(t) - \boldsymbol{\mu}(t)]^H \\ &\quad \times \mathbf{R}^{-1} \frac{\partial \boldsymbol{\mu}(t)}{\partial \alpha_m} + \frac{\partial \boldsymbol{\mu}^H(t)}{\partial \alpha_m} \mathbf{R}^{-1} [\mathbf{y}(t) - \boldsymbol{\mu}(t)]. \end{aligned} \quad (\text{I.32})$$

Taking the second order partial derivative for (I.32), which resorts to equation (15.48) in [42], we can obtain after some algebraic operations that (detailed derivation, which is lengthy but simple, is omitted here)

$$\mathbf{\Gamma}_N^{-1} \frac{\partial^2 \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} \mathbf{\Gamma}_N^{-1} \Big|_{\boldsymbol{\alpha}=\tilde{\boldsymbol{\alpha}}} \rightarrow -(\mathbf{F}_d^{\text{as}} + \mathbf{F}_s) \quad (\text{I.33})$$

as $N \rightarrow \infty$. Note that $\tilde{\boldsymbol{\alpha}} \rightarrow \boldsymbol{\alpha}_0$ as $N \rightarrow \infty$ due to the consistency of $\hat{\boldsymbol{\alpha}}$.

According to Slutsky's theorem, the result in (I.33), and the fact that $\sum_{t=1}^N \boldsymbol{\gamma}_t$ asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_d^{\text{as}} + \mathbf{F}_s)$, we conclude that $\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0$ asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_N^{-1}(\mathbf{F}_d^{\text{as}} + \mathbf{F}_s)^{-1} \mathbf{\Gamma}_N^{-1})$. From (I.17) and (I.28), we can see that $\mathbf{\Gamma}_N(\mathbf{F}_d + \mathbf{F}_s)\mathbf{\Gamma}_N$ is the Fisher information matrix for $\boldsymbol{\alpha}$. Using the fact that \mathbf{F}_d converges to \mathbf{F}_d^{as} as N increases, we obtain that the asymptotic CRB also equals $\mathbf{\Gamma}_N^{-1}(\mathbf{F}_d^{\text{as}} + \mathbf{F}_s)^{-1} \mathbf{\Gamma}_N^{-1}$. We thus prove the proposition. The asymptotic CRB on $\boldsymbol{\theta}$ can be obtained by using (3.30) and replacing $\hat{\mathbf{R}}_{\phi\phi}$ in (H.8) with an identity matrix of the same size.

Appendix J

Proof of Proposition 10

If the stochastic estimator is used for DOA estimation of mixed signals, the objective function to be maximized is the stochastic LL function

$$L(\boldsymbol{\theta}, \mathbf{P}, \mathbf{Q}) = -\log |\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}| - \text{trace} \left\{ [\mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}]^{-1} \frac{\mathbf{Y}\mathbf{Y}^H}{N} \right\}. \quad (\text{J.1})$$

As N increases, $\mathbf{Y}\mathbf{Y}^H/N$ converges to the true value of $\mathbf{A}\tilde{\mathbf{P}}\mathbf{A}^H + \mathbf{Q}$. Thus, the stochastic estimator provides consistent estimates for $\boldsymbol{\theta}$, $\tilde{\mathbf{P}}$, and \mathbf{Q} .

Let $f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\xi})$ be the pdf based on the stochastic measurement model. Assuming $\hat{\boldsymbol{\xi}}$ is the stochastic estimate of $\boldsymbol{\xi}$, we apply a Taylor expansion as follows around the true value of $\boldsymbol{\xi}$, which is denoted by $\boldsymbol{\xi}_0$.

$$\left. \frac{\partial \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \right|_{\boldsymbol{\xi}=\hat{\boldsymbol{\xi}}} = \left. \frac{\partial \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0} + \left. \frac{\partial^2 \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\xi})}{\partial \boldsymbol{\xi} \partial \boldsymbol{\xi}^T} \right|_{\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}} (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0) = \mathbf{0}, \quad (\text{J.2})$$

where $\tilde{\boldsymbol{\xi}}$ is a vector value between $\boldsymbol{\xi}_0$ and $\hat{\boldsymbol{\xi}}$. Thus we have

$$\sqrt{N}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0) = - \left[\frac{1}{N} \frac{\partial^2 \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\xi})}{\partial \boldsymbol{\xi} \partial \boldsymbol{\xi}^T} \right]^{-1} \bigg|_{\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}} \left[\frac{1}{\sqrt{N}} \frac{\partial \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \right] \bigg|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}. \quad (\text{J.3})$$

For simplicity, we omit the subscript $\boldsymbol{\xi} = \boldsymbol{\xi}_0$ in the following equations. Using equations (15.47) and (15.48) in [42], we can show that

$$\begin{aligned}
\frac{\partial \ln f_{\mathbf{y}(t)}(\mathbf{y}(t); \boldsymbol{\xi})}{\partial \xi_m} &= -\text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \right\} + \text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1} \mathbf{y}(t) \mathbf{y}^H(t) \right\}, \quad (\text{J.4}) \\
\frac{\partial^2 \ln f_{\mathbf{y}(t)}(\mathbf{y}(t); \boldsymbol{\xi})}{\partial \xi_m \partial \xi_n} &= \text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_n} \right\} - \text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m \partial \xi_n} \right\} \\
&\quad - \text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_n} \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1} \mathbf{y}(t) \mathbf{y}^H(t) \right\} \\
&\quad + \text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m \partial \xi_n} \check{\mathbf{R}}^{-1} \mathbf{y}(t) \mathbf{y}^H(t) \right\} \\
&\quad - \text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_n} \check{\mathbf{R}}^{-1} \mathbf{y}(t) \mathbf{y}^H(t) \right\}. \quad (\text{J.5})
\end{aligned}$$

Thus we have

$$\mathbb{E} \left\{ \frac{\partial \ln f_{\mathbf{y}(t)}(\mathbf{y}(t); \boldsymbol{\xi})}{\partial \xi_m} \right\} = \text{trace} \left\{ \mathbf{K}_m \check{\mathbf{R}}(t) \right\} - \text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \right\}, \quad (\text{J.6})$$

where $\mathbf{K}_m = \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1}$ and $\check{\mathbf{R}}(t) = \mathbf{R} + \boldsymbol{\mu}(t) \boldsymbol{\mu}^H(t)$. Note that the expectation in (J.6) is computed under the pdf of the mixed measurement model. This also holds for the all the following expectations in this proof. Using the fact that $\frac{1}{N} \sum_{t=1}^N \mathbf{y}(t) \mathbf{y}^H(t) \rightarrow \check{\mathbf{R}}$ as $N \rightarrow \infty$ and the consistency of $\hat{\boldsymbol{\xi}}$, we have

$$\left. \frac{1}{N} \frac{\partial^2 \ln f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\xi})}{\partial \xi_m \partial \xi_n} \right|_{\boldsymbol{\xi}=\hat{\boldsymbol{\xi}}} \rightarrow -\text{trace} \left\{ \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_n} \right\}, \quad (\text{J.7})$$

as $N \rightarrow \infty$.

Let

$$\boldsymbol{\gamma}_t = \frac{1}{\sqrt{N}} \check{\boldsymbol{\gamma}}_t = \frac{1}{\sqrt{N}} \frac{\ln f_{\mathbf{y}(t)}(\mathbf{y}(t); \boldsymbol{\xi})}{\partial \boldsymbol{\xi}}, \quad (\text{J.8})$$

and $[\check{\boldsymbol{\gamma}}_t]_m$ be the m -th element of $\check{\boldsymbol{\gamma}}_t$. Using the results in (J.4) and (J.6), we have

$$\begin{aligned}
[\check{\gamma}_t]_m - \mathbb{E}\{[\check{\gamma}_t]_m\} &= \text{trace}\{\mathbf{K}_m[\mathbf{y}(t)\mathbf{y}^H(t) - \check{\mathbf{R}}(t)]\} \\
&= \text{trace}\{\mathbf{K}_m[\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t) + \tilde{\mathbf{y}}(t)\boldsymbol{\mu}^H(t) + \boldsymbol{\mu}(t)\tilde{\mathbf{y}}^H(t) - \mathbf{R}]\}, \tag{J.9}
\end{aligned}$$

where $\tilde{\mathbf{y}}(t) = \mathbf{y}(t) - \boldsymbol{\mu}(t)$. Using (J.9) and algebraic manipulations, we have the covariance of $[\check{\gamma}_t]_m$ and $[\check{\gamma}_t]_n$ as

$$\text{cov}\{[\check{\gamma}_t]_m, [\check{\gamma}_t]_n\} = \sum_{i=1}^7 G_i(t), \tag{J.10}$$

where

$$\begin{aligned}
G_1(t) &= \mathbb{E}\{\text{trace}\{\mathbf{K}_m\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)\}\text{trace}\{\mathbf{K}_n\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)\}\} \\
&= \text{trace}\{\mathbf{K}_m\mathbf{R}\}\text{trace}\{\mathbf{K}_n\mathbf{R}\} + \text{trace}\{\mathbf{K}_m\mathbf{R}\mathbf{K}_n\mathbf{R}\}, \tag{J.11}
\end{aligned}$$

$$\begin{aligned}
G_2(t) &= -\mathbb{E}\{\text{trace}\{\mathbf{K}_m\mathbf{R}\}\text{trace}\{\mathbf{K}_n\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)\}\} \\
&= -\text{trace}\{\mathbf{K}_m\mathbf{R}\}\text{trace}\{\mathbf{K}_n\mathbf{R}\}, \tag{J.12}
\end{aligned}$$

$$\begin{aligned}
G_3(t) &= \mathbb{E}\{\text{trace}\{\mathbf{K}_m[\tilde{\mathbf{y}}(t)\boldsymbol{\mu}^H(t) + \boldsymbol{\mu}(t)\tilde{\mathbf{y}}^H(t)]\} \\
&\quad \times \text{trace}\{\mathbf{K}_n[\tilde{\mathbf{y}}(t)\boldsymbol{\mu}^H(t) + \boldsymbol{\mu}(t)\tilde{\mathbf{y}}^H(t)]\}\} \\
&= \text{trace}\{\mathbf{K}_m\mathbf{R}\mathbf{K}_n\boldsymbol{\mu}(t)\boldsymbol{\mu}^H(t)\} + \text{trace}\{\mathbf{K}_m\boldsymbol{\mu}(t)\boldsymbol{\mu}^H(t)\mathbf{K}_n\mathbf{R}\}, \tag{J.13}
\end{aligned}$$

$$\begin{aligned}
G_4(t) &= -\mathbb{E}\{\text{trace}\{\mathbf{K}_m\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)\}\text{trace}\{\mathbf{K}_n\mathbf{R}\}\} \\
&= -\text{trace}\{\mathbf{K}_m\mathbf{R}\}\text{trace}\{\mathbf{K}_n\mathbf{R}\}, \tag{J.14}
\end{aligned}$$

$$G_5(t) = \text{trace}\{\mathbf{K}_m\mathbf{R}\}\text{trace}\{\mathbf{K}_n\mathbf{R}\}, \tag{J.15}$$

$$G_6(t) = \mathbb{E}\{\text{trace}\{\mathbf{K}_m[\tilde{\mathbf{y}}(t)\boldsymbol{\mu}^H(t) + \boldsymbol{\mu}(t)\tilde{\mathbf{y}}^H(t)]\}\text{trace}\{\mathbf{K}_n\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)\}\}, \tag{J.16}$$

$$G_7(t) = \mathbb{E}\{\text{trace}\{\mathbf{K}_m\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)\}\text{trace}\{\mathbf{K}_n[\tilde{\mathbf{y}}(t)\boldsymbol{\mu}^H(t) + \boldsymbol{\mu}(t)\tilde{\mathbf{y}}^H(t)]\}\}. \tag{J.17}$$

Note that we obtain (J.13) by the fact that $\mathbb{E}\{\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^T(t)\} = \mathbf{0}$, and equation (J.11) comes from Appendix 15C in [42]. Substituting the results in (J.11)-(J.17) into (J.10),

we have

$$\begin{aligned} \text{cov}\{[\check{\gamma}_t]_m, [\check{\gamma}_t]_n\} &= G_6(t) + G_7(t) + \text{trace}\{\mathbf{K}_m \mathbf{R} \mathbf{K}_n \mathbf{R}\} \\ &+ \text{trace}\{\mathbf{K}_m \mathbf{R} \mathbf{K}_n \boldsymbol{\mu}(t) \boldsymbol{\mu}^H(t)\} + \text{trace}\{\mathbf{K}_m \boldsymbol{\mu}(t) \boldsymbol{\mu}^H(t) \mathbf{K}_n \mathbf{R}\}, \end{aligned} \quad (\text{J.18})$$

Since $\frac{1}{N} \sum_{t=1}^N \boldsymbol{\mu}(t) \rightarrow \mathbf{0}$ and $\frac{1}{N} \sum_{t=1}^N \boldsymbol{\varphi}(t) \boldsymbol{\varphi}^H(t) \rightarrow \mathbf{I}_J$ as $N \rightarrow \infty$, we have

$$\frac{1}{N} \sum_{t=1}^N G_6(t) \rightarrow 0, \quad (\text{J.19})$$

$$\frac{1}{N} \sum_{t=1}^N G_7(t) \rightarrow 0, \quad (\text{J.20})$$

and

$$\frac{1}{N} \sum_{t=1}^N \boldsymbol{\mu}(t) \boldsymbol{\mu}^H(t) \rightarrow \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \quad (\text{J.21})$$

as $N \rightarrow \infty$. From (J.18)-(J.21), we obtain

$$\sum_{t=1}^N \text{cov}\{\gamma_t\} \rightarrow \mathbf{F}_{\check{\xi}} - \mathbf{F}_{\mu} \quad (\text{J.22})$$

as $N \rightarrow \infty$, where

$$[\mathbf{F}_{\check{\xi}}]_{mn} = \text{trace}\left\{\check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_n}\right\}, \quad (\text{J.23})$$

$$[\mathbf{F}_{\mu}]_{mn} = \text{trace}\{\mathbf{K}_m \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \mathbf{K}_n \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H\}. \quad (\text{J.24})$$

In addition, we can see that $\check{\gamma}_t$ is a random vector with finite correlation matrix.

Therefore, as $N \rightarrow \infty$,

$$\sum_{t=1}^N \int_{\|\gamma_t\| > \epsilon} \|\gamma_t\|^2 f_{\gamma_t}(\gamma_t; \boldsymbol{\alpha}) d\gamma_t = \frac{1}{N} \sum_{t=1}^N \int_{\|\check{\gamma}_t\| > \sqrt{N}\epsilon} \|\check{\gamma}_t\|^2 f_{\check{\gamma}_t}(\check{\gamma}_t; \boldsymbol{\alpha}) d\check{\gamma}_t \rightarrow 0. \quad (\text{J.25})$$

From Theorem 1, we have $\sum_{t=1}^N [\gamma_t - \mathbb{E}\{\gamma_t\}]$ asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_{\check{\xi}} - \mathbf{F}_{\mu})$. According to (J.6), we have

$$\begin{aligned} \mathbb{E}\{[\check{\gamma}_t]_m\} &= \text{trace}\{\mathbf{K}_m \check{\mathbf{R}}(t)\} - \text{trace}\{\mathbf{K}_m \check{\mathbf{R}}\} \\ &= \text{trace}\{\mathbf{K}_m \mathbf{A} \mathbf{C} [\boldsymbol{\varphi}(t) \boldsymbol{\varphi}^H(t) - \mathbf{I}_J] \mathbf{C}^H \mathbf{A}^H\}, \end{aligned} \quad (\text{J.26})$$

from which we have $\sum_{t=1}^N \mathbb{E}\{\gamma_t\} \rightarrow \mathbf{0}$ as $N \rightarrow \infty$ since $\frac{1}{\sqrt{N}} \sum_{t=1}^N [\boldsymbol{\varphi}(t) \boldsymbol{\varphi}^H(t) - \mathbf{I}_J] \rightarrow \mathbf{0}$ as $N \rightarrow \infty$. As a result, we can see that $\sum_{t=1}^N \gamma_t$, which is the second term on the right side of (J.3), asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_{\check{\xi}} - \mathbf{F}_{\mu})$. From this result and the result in (J.7), we have $\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0$ asymptotically follows $\mathcal{N}\left(\mathbf{0}, \frac{\mathbf{F}_{\check{\xi}}^{-1} - \mathbf{F}_{\check{\xi}}^{-1} \mathbf{F}_{\mu} \mathbf{F}_{\check{\xi}}^{-1}}{N}\right)$.

Using the property that

$$\text{trace}\{\mathbf{A} \mathbf{B} \mathbf{C} \mathbf{D}\} = (\text{vec}\{\mathbf{A}^T\})^T (\mathbf{D}^T \otimes \mathbf{B}) \text{vec}\{\mathbf{C}\} \quad (\text{J.27})$$

for arbitrary matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{D} that can produce $\mathbf{A} \mathbf{B} \mathbf{C} \mathbf{D}$ [63], we rewrite

$$\begin{aligned} [\mathbf{F}_{\mu}]_{mn} &= \text{trace}\left\{\frac{\partial \check{\mathbf{R}}}{\partial \xi_m} \check{\mathbf{R}}^{-1} \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \check{\mathbf{R}}^{-1} \frac{\partial \check{\mathbf{R}}}{\partial \xi_n} \check{\mathbf{R}}^{-1} \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \check{\mathbf{R}}^{-1}\right\} \\ &= \left(\frac{\partial \check{\mathbf{r}}}{\partial \xi_m}\right)^H \left[\left(\check{\mathbf{R}}^{-1} \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \check{\mathbf{R}}^{-1}\right)^T \right. \\ &\quad \left. \otimes \left(\check{\mathbf{R}}^{-1} \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \check{\mathbf{R}}^{-1}\right) \right] \left(\frac{\partial \check{\mathbf{r}}}{\partial \xi_n}\right), \end{aligned} \quad (\text{J.28})$$

where $\check{\mathbf{r}} = \text{vec}\{\check{\mathbf{R}}\}$. Thus we can formulate \mathbf{F}_{μ} as

$$\begin{aligned} \mathbf{F}_{\mu} &= \left(\frac{\partial \check{\mathbf{r}}}{\partial \boldsymbol{\xi}^T}\right)^H \left[\left(\check{\mathbf{R}}^{-1} \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \check{\mathbf{R}}^{-1}\right)^T \right. \\ &\quad \left. \otimes \left(\check{\mathbf{R}}^{-1} \mathbf{A} \mathbf{C} \mathbf{C}^H \mathbf{A}^H \check{\mathbf{R}}^{-1}\right) \right] \frac{\partial \check{\mathbf{r}}}{\partial \boldsymbol{\xi}^T} \\ &= \boldsymbol{\Phi}^H \left[\left(\check{\mathbf{A}} \mathbf{C} \mathbf{C}^H \check{\mathbf{A}}^H\right)^T \otimes \left(\check{\mathbf{A}} \mathbf{C} \mathbf{C}^H \check{\mathbf{A}}^H\right) \right] \boldsymbol{\Phi}, \end{aligned} \quad (\text{J.29})$$

where $\check{\mathbf{A}} = \check{\mathbf{R}}^{-\frac{1}{2}} \mathbf{A}$, and

$$\Phi = \left(\check{\mathbf{R}}^{-\frac{T}{2}} \otimes \check{\mathbf{R}}^{-\frac{1}{2}} \right) \frac{\partial \check{\mathbf{r}}}{\partial \check{\boldsymbol{\xi}}^T}. \quad (\text{J.30})$$

Thus, we have

$$\mathbf{F}_{\check{\mathbf{S}}}^{-1} \mathbf{F}_{\mu} \mathbf{F}_{\check{\mathbf{S}}}^{-1} = \mathbf{W} \mathbf{W}^H, \quad (\text{J.31})$$

where $\mathbf{W} = \mathbf{F}_{\check{\mathbf{S}}}^{-1} \Phi^H [(\check{\mathbf{A}}^* \mathbf{C}^*) \otimes (\check{\mathbf{A}} \mathbf{C})]$.

We partition Φ into

$$\Phi = [\mathbf{U} | \mathbf{V}] = \left(\check{\mathbf{R}}^{-\frac{T}{2}} \otimes \check{\mathbf{R}}^{-\frac{1}{2}} \right) \left[\frac{\partial \check{\mathbf{r}}}{\partial \boldsymbol{\theta}^T}, \frac{\partial \check{\mathbf{r}}}{\partial \boldsymbol{\sigma}^T} \middle| \frac{\partial \check{\mathbf{r}}}{\partial \check{\mathbf{p}}^T} \right], \quad (\text{J.32})$$

and partition $\mathbf{F}_{\check{\mathbf{S}}}$ (note that $\mathbf{F}_{\check{\mathbf{S}}} = \Phi^H \Phi$ from (J.27)) into

$$\mathbf{F}_{\check{\mathbf{S}}} = \begin{bmatrix} \mathbf{U}^H \\ \mathbf{V}^H \end{bmatrix} [\mathbf{U}, \mathbf{V}] = \begin{bmatrix} \mathbf{U}^H \mathbf{U} & \mathbf{U}^H \mathbf{V} \\ \mathbf{V}^H \mathbf{U} & \mathbf{V}^H \mathbf{V} \end{bmatrix}. \quad (\text{J.33})$$

Applying the block-wise matrix inversion (see equation (2.8.17) in [63]) to $\mathbf{F}_{\check{\mathbf{S}}}$ and employing some algebraic matrix operations, we obtain

$$\mathbf{W}_{1:l_{\rho}} = (\mathbf{U}^H \Pi_{\mathbf{V}}^{\perp} \mathbf{U})^{-1} \mathbf{U}^H \Pi_{\mathbf{V}}^{\perp} [(\check{\mathbf{A}}^* \mathbf{C}^*) \otimes (\check{\mathbf{A}} \mathbf{C})], \quad (\text{J.34})$$

where $\mathbf{W}_{1:l_{\rho}}$ is the matrix consisting of the first l_{ρ} rows of \mathbf{W} , and l_{ρ} is the length of $\boldsymbol{\rho}$.

According to the result in [43] that

$$\Pi_{\mathbf{V}}^{\perp} = \Pi_{\check{\mathbf{A}}^* \otimes \check{\mathbf{A}}}^{\perp} = \mathbf{I} \otimes \Pi_{\check{\mathbf{A}}}^{\perp} + \Pi_{\check{\mathbf{A}}^*}^{\perp} \otimes \mathbf{I} - \Pi_{\check{\mathbf{A}}^*}^{\perp} \otimes \Pi_{\check{\mathbf{A}}}^{\perp}, \quad (\text{J.35})$$

we have $\mathbf{\Pi}_V^\perp [(\check{\mathbf{A}}^* \mathbf{C}^*) \otimes (\check{\mathbf{A}} \mathbf{C})] = \mathbf{0}$ and $\mathbf{W}_{1:l_\rho} = \mathbf{0}$. However, checking the value of $\mathbf{W}_{(l_\rho+1):l_\xi}$, we find that it is not $\mathbf{0}$ in general. Therefore, $\mathbf{F}_\xi^{-1} \mathbf{F}_\mu \mathbf{F}_\xi^{-1}$ is a matrix with the bottom-right $l_{\check{\boldsymbol{\rho}}} \times l_{\check{\boldsymbol{\rho}}}$ sub-matrix non-negative definite, while all other elements equal to 0. Note that $l_{\check{\boldsymbol{\rho}}}$ and l_ξ are the lengths of $\check{\boldsymbol{\rho}}$ and $\boldsymbol{\xi}$, respectively. Thus, the asymptotic error covariance matrix on $\boldsymbol{\rho}$ equals the upper-left $l_\rho \times l_\rho$ sub-matrix related to $\boldsymbol{\rho}$ in $\frac{\mathbf{F}_\xi^{-1}}{N}$, which is $\mathbf{CRB}_{\check{\boldsymbol{\rho}}}$, and the asymptotic error covariance matrix on $\boldsymbol{\theta}$ equals $\mathbf{CRB}_{\check{\boldsymbol{\theta}}}$. The asymptotic error covariance matrix on $\check{\boldsymbol{\rho}}$ should generally not be larger than $\mathbf{CRB}_{\check{\boldsymbol{\rho}}}$.

Appendix K

Proof of Proposition 11

Applying block-wise matrix inversion (see (2.8.17) in [63]) to (H.5) and using the results in (H.8) and [43], we have

$$\frac{1}{N} \mathbf{CRB}_{\mathbf{M}, \boldsymbol{\rho}}^{-1} = \begin{bmatrix} \frac{1}{N} \mathbf{CRB}_{\mathbf{D}, \boldsymbol{\theta}}^{-1} + \mathbf{F} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{T} \end{bmatrix}. \quad (\text{K.1})$$

where

$$\begin{aligned} \mathbf{F} &= \frac{1}{N} (\mathbf{F}_{\mathbf{S}, \boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{S}, \boldsymbol{\theta}\mathbf{p}} \mathbf{F}_{\mathbf{S}, \mathbf{p}\mathbf{p}}^{-1} \mathbf{F}_{\mathbf{S}, \mathbf{p}\boldsymbol{\theta}}) \\ &= 2\text{Re} \left\{ \left(\tilde{\mathbf{D}}^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp} \tilde{\mathbf{D}} \right) \odot \left(\mathbf{P} \tilde{\mathbf{A}}^H \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{A}} \mathbf{P} \right)^T \right\}, \end{aligned} \quad (\text{K.2})$$

$$\begin{aligned} \mathbf{M} &= \frac{1}{N} (\mathbf{F}_{\mathbf{S}, \boldsymbol{\theta}\boldsymbol{\sigma}} - \mathbf{F}_{\mathbf{S}, \boldsymbol{\theta}\mathbf{p}} \mathbf{F}_{\mathbf{S}, \mathbf{p}\mathbf{p}}^{-1} \mathbf{F}_{\mathbf{S}, \mathbf{p}\boldsymbol{\sigma}}) \\ &= 2\text{Re} \left\{ \boldsymbol{\Xi}^T \left[\left(\tilde{\mathbf{D}}^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp} \right) \otimes \left(\mathbf{P}^T \tilde{\mathbf{A}}^T \tilde{\mathbf{R}}^{-T} \right) \right] \boldsymbol{\Lambda}^* \right\}, \end{aligned} \quad (\text{K.3})$$

$$\begin{aligned} \mathbf{T} &= \frac{1}{N} (\mathbf{F}_{\mathbf{S}, \boldsymbol{\sigma}\boldsymbol{\sigma}} - \mathbf{F}_{\mathbf{S}, \boldsymbol{\sigma}\mathbf{p}} \mathbf{F}_{\mathbf{S}, \mathbf{p}\mathbf{p}}^{-1} \mathbf{F}_{\mathbf{S}, \mathbf{p}\boldsymbol{\sigma}}) \\ &= 2\text{Re} \left\{ \boldsymbol{\Lambda}^H \left(\tilde{\mathbf{R}}^{-T} \otimes \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp} \right) \boldsymbol{\Lambda} \right\} + \boldsymbol{\Lambda}^H \left(\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp T} \otimes \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp} \right) \boldsymbol{\Lambda}, \end{aligned} \quad (\text{K.4})$$

and $\tilde{\mathbf{R}} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{R} \mathbf{Q}^{-\frac{1}{2}}$. From (K.1) and the results in Proposition 2, we have

$$\frac{1}{N}\mathbf{C}_{\mathbf{M},\boldsymbol{\rho}}^{\text{as}-1} = \begin{bmatrix} \frac{1}{N}\mathbf{CRB}_{\mathbf{D},\boldsymbol{\theta}}^{\text{as}-1} + \mathbf{F} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{T} \end{bmatrix}, \quad (\text{K.5})$$

where

$$\mathbf{CRB}_{\mathbf{D},\boldsymbol{\theta}}^{\text{as}-1} = 2N\text{Re}\left\{ \left(\tilde{\mathbf{D}}^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp \tilde{\mathbf{D}} \right) \odot (\mathbf{C}\mathbf{C}^H)^T \right\}. \quad (\text{K.6})$$

Also from the results in [43], we have

$$\frac{1}{N}\mathbf{CRB}_{\tilde{\mathbf{S}},\boldsymbol{\rho}}^{-1} = \begin{bmatrix} \check{\mathbf{F}} & \check{\mathbf{M}} \\ \check{\mathbf{M}}^T & \check{\mathbf{T}} \end{bmatrix}, \quad (\text{K.7})$$

where the expressions of $\check{\mathbf{F}}$, $\check{\mathbf{M}}$, and $\check{\mathbf{T}}$ can be obtained by replacing \mathbf{P} and \mathbf{R} in (K.2)-(K.4) with $\check{\mathbf{P}}$ and $\check{\mathbf{R}}$, respectively.

In the following we first show that

$$\frac{1}{N}(\mathbf{C}_{\mathbf{M},\boldsymbol{\rho}}^{\text{as}-1} - \mathbf{CRB}_{\tilde{\mathbf{S}},\boldsymbol{\rho}}^{-1}) = [\mathbf{S}, \mathbf{O}]^H \boldsymbol{\Pi}_H^\perp [\mathbf{S}, \mathbf{O}] \geq \mathbf{0}, \quad (\text{K.8})$$

where $\mathbf{H} = \bar{\mathbf{A}}^* \otimes \bar{\mathbf{A}}$ with $\bar{\mathbf{A}} = \mathbf{R}^{-\frac{1}{2}}\mathbf{A}$, \mathbf{O} is an $M^2 \times l_\sigma$ matrix with the m -th column

$$\mathbf{o}_m = -\text{vec}\left\{ \mathbf{R}^{-\frac{1}{2}}(\mathbf{Q}'_m \mathbf{Q}^{-1} \mathbf{A} \mathbf{B}^H \mathbf{A}^H + \mathbf{A} \mathbf{B} \mathbf{A}^H \mathbf{Q}^{-1} \mathbf{Q}'_m) \mathbf{R}^{-\frac{1}{2}} \right\}, \quad (\text{K.9})$$

and \mathbf{S} is an $M^2 \times L$ matrix with the m -th column $\mathbf{s}_m = \text{vec}\{\mathbf{Z}_m + \mathbf{Z}_m^H\}$, in which $\mathbf{Z}_m = \bar{\mathbf{A}} \mathbf{b}_m \mathbf{d}_m^H \mathbf{R}^{-\frac{1}{2}}$, \mathbf{d}_m is the m -th column of \mathbf{D} , and \mathbf{b}_m is the m -th column of an $L \times L$ matrix \mathbf{B} satisfying $\mathbf{B}^H \mathbf{A}^H \mathbf{R}^{-1} \mathbf{A} \mathbf{B} = \boldsymbol{\Delta}$ with

$$\boldsymbol{\Delta} = (\check{\mathbf{P}}^{-1} + \check{\mathbf{A}}^H \check{\mathbf{A}})^{-1} - (\mathbf{P}^{-1} + \tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \geq \mathbf{0}. \quad (\text{K.10})$$

We show the equality in (K.8) holds by showing that the matrices on the two sides of the equality are blockwise equal. Using the fact that $\mathbf{P}\tilde{\mathbf{A}}^H\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{A}}\mathbf{P} = \mathbf{P} - (\mathbf{P}^{-1} + \tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1}$ [8], we can obtain

$$\frac{1}{N}\mathbf{CRB}_{\mathbf{D},\boldsymbol{\theta}}^{\text{as}-1} + \mathbf{F} - \check{\mathbf{F}} = 2\text{Re}\left\{\left(\tilde{\mathbf{D}}^H\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp\tilde{\mathbf{D}}\right) \odot \boldsymbol{\Delta}^T\right\}, \quad (\text{K.11})$$

Following similar derivations in (27) and (30)-(32) in [43], we can obtain that

$$\mathbf{S}^H\boldsymbol{\Pi}_{\mathbf{H}}^\perp\mathbf{S} = 2\text{Re}\left\{\left(\tilde{\mathbf{D}}^H\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp\tilde{\mathbf{D}}\right) \odot \left(\mathbf{B}\tilde{\mathbf{A}}^H\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{A}}\mathbf{B}\right)^T\right\}, \quad (\text{K.12})$$

which equals the result in (K.11).

From the result in [8] that $\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}(\mathbf{P}\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \mathbf{I}_L)^{-1}$, we can obtain $\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{A}}\mathbf{P} = \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \mathbf{P}^{-1})^{-1}$ and

$$\mathbf{M} - \check{\mathbf{M}} = -2\text{Re}\left\{\boldsymbol{\Xi}^T\left[\left(\tilde{\mathbf{D}}^H\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp\right) \otimes (\tilde{\mathbf{A}}\boldsymbol{\Delta})^T\right]\boldsymbol{\Lambda}^*\right\}. \quad (\text{K.13})$$

Following the derivation of (27) in [43], we can similarly obtain

$$\boldsymbol{\Pi}_{\mathbf{H}}^\perp\mathbf{s}_m = \text{vec}\{\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp\mathbf{Z}_m^H + \mathbf{Z}_m\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^\perp\}. \quad (\text{K.14})$$

Using (K.14) and the property that

$$\text{trace}\{\mathbf{X}\mathbf{Y}\} = \text{vec}\{\mathbf{X}^H\}^H\text{vec}\{\mathbf{Y}\} \quad (\text{K.15})$$

for any matrices \mathbf{X} and \mathbf{Y} that can make $\mathbf{X}\mathbf{Y}$ [64], we obtain

$$\begin{aligned}
\mathbf{s}_m^H \boldsymbol{\Pi}_H^\perp \mathbf{o}_n &= -\text{vec}\{\boldsymbol{\Pi}_A^\perp \mathbf{Z}_m^H + \mathbf{Z}_m \boldsymbol{\Pi}_A^\perp\}^H \\
&\quad \times \text{vec}\{\mathbf{R}^{-\frac{1}{2}}(\mathbf{Q}'_n \mathbf{Q}^{-1} \mathbf{A} \mathbf{B}^H \mathbf{A}^H + \mathbf{A} \mathbf{B} \mathbf{A}^H \mathbf{Q}^{-1} \mathbf{Q}'_n) \mathbf{R}^{-\frac{1}{2}}\} \\
&= -2\text{Re}\{\text{trace}\{\mathbf{Q}'_n \mathbf{Q}^{-1} \mathbf{A} \mathbf{B}^H \mathbf{A}^H \mathbf{R}^{-1} \mathbf{A} \mathbf{b}_m \mathbf{d}_m^H \mathbf{Q}^{-\frac{1}{2}} \boldsymbol{\Pi}_A^\perp \mathbf{Q}^{-\frac{1}{2}}\}\} \\
&= -2\text{Re}\{\text{trace}\{\tilde{\mathbf{Q}}'_n \tilde{\mathbf{A}} \Delta \mathbf{e}_m \mathbf{e}_m^T \tilde{\mathbf{D}}^H \boldsymbol{\Pi}_A^\perp\}\} \\
&= -2\text{Re}\{\text{vec}\{\tilde{\mathbf{Q}}'_n\}^H \text{vec}\{\tilde{\mathbf{A}} \Delta \mathbf{e}_m \mathbf{e}_m^T \tilde{\mathbf{D}}^H \boldsymbol{\Pi}_A^\perp\}\} \\
&= -2\text{Re}\{\text{vec}\{\tilde{\mathbf{Q}}'_n\}^H [(\tilde{\mathbf{D}}^H \boldsymbol{\Pi}_A^\perp)^T \otimes (\tilde{\mathbf{A}} \Delta)] \text{vec}\{\mathbf{e}_m \mathbf{e}_m^T\}\} \\
&= -2\text{Re}\{\text{vec}\{\mathbf{e}_m \mathbf{e}_m^T\}^T [(\tilde{\mathbf{D}}^H \boldsymbol{\Pi}_A^\perp) \otimes (\tilde{\mathbf{A}} \Delta)^T] \text{vec}\{\tilde{\mathbf{Q}}'_n\}^*\}. \tag{K.16}
\end{aligned}$$

Note that in the derivation of the second equality in (K.16), we resort to the result in (H.9) and the fact that $\boldsymbol{\Pi}_A^\perp \mathbf{Z}_m = \mathbf{0}$. From (K.16), we can see that $\mathbf{S}^H \boldsymbol{\Pi}_H^\perp \mathbf{O}$ equals the result in (K.13).

Using the Woodbury matrix identity, we have $\tilde{\mathbf{R}}^{-1} = \mathbf{I}_M - \tilde{\mathbf{A}}(\mathbf{P}^{-1} + \tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H$. As a result,

$$\mathbf{T} - \check{\mathbf{T}} = 2\text{Re}\left\{ \boldsymbol{\Lambda}^H \left[(\tilde{\mathbf{A}} \Delta \tilde{\mathbf{A}}^H)^T \otimes \boldsymbol{\Pi}_A^\perp \right] \boldsymbol{\Lambda} \right\}. \tag{K.17}$$

Using the property in (3.5), the facts that $\boldsymbol{\Pi}_A^\perp \bar{\mathbf{A}} = \mathbf{0}$ and $\boldsymbol{\Pi}_{\bar{\mathbf{A}}^*}^\perp = (\boldsymbol{\Pi}_A^\perp)^T$, and the result in [43] that

$$\boldsymbol{\Pi}_{\bar{\mathbf{A}}^* \otimes \bar{\mathbf{A}}}^\perp = \mathbf{I}_M \otimes \boldsymbol{\Pi}_A^\perp + \boldsymbol{\Pi}_{\bar{\mathbf{A}}^*}^\perp \otimes \mathbf{I}_M - \boldsymbol{\Pi}_{\bar{\mathbf{A}}^*}^\perp \otimes \boldsymbol{\Pi}_A^\perp, \tag{K.18}$$

we can obtain after some algebraic manipulations

$$\begin{aligned}
\boldsymbol{\Pi}_H^\perp \mathbf{o}_n &= -\text{vec}\{\boldsymbol{\Pi}_A^\perp \mathbf{R}^{-\frac{1}{2}} \mathbf{Q}'_n \mathbf{Q}^{-1} \mathbf{A} \mathbf{B}^H \mathbf{A}^H \mathbf{R}^{-\frac{1}{2}} \\
&\quad + \mathbf{R}^{-\frac{1}{2}} \mathbf{A} \mathbf{B} \mathbf{A}^H \mathbf{Q}^{-1} \mathbf{Q}'_n \mathbf{R}^{-\frac{1}{2}} \boldsymbol{\Pi}_A^\perp\}, \tag{K.19}
\end{aligned}$$

from which we can further obtain using (K.9) and (K.15) that

$$\begin{aligned}
\mathbf{o}_m^H \mathbf{\Pi}_H^\perp \mathbf{o}_n &= 2\text{Re}\{\text{trace}\{\mathbf{R}^{-\frac{1}{2}} \mathbf{A} \mathbf{B} \mathbf{A}^H \mathbf{Q}^{-1} \mathbf{Q}'_m \mathbf{R}^{-\frac{1}{2}} \mathbf{\Pi}_A^\perp \\
&\quad \times \mathbf{R}^{-\frac{1}{2}} \mathbf{Q}'_n \mathbf{Q}^{-1} \mathbf{A} \mathbf{B}^H \mathbf{A}^H \mathbf{R}^{-\frac{1}{2}}\}\} \\
&= 2\text{Re}\{\text{trace}\{\tilde{\mathbf{Q}}'_m \mathbf{\Pi}_A^\perp \tilde{\mathbf{Q}}'_n \tilde{\mathbf{A}} \mathbf{B}^H \mathbf{A}^H \mathbf{R}^{-1} \mathbf{A} \mathbf{B} \tilde{\mathbf{A}}^H\}\} \\
&= 2\text{Re}\{\text{trace}\{\tilde{\mathbf{Q}}'_m \mathbf{\Pi}_A^\perp \tilde{\mathbf{Q}}'_n \tilde{\mathbf{A}} \mathbf{\Delta} \tilde{\mathbf{A}}^H\}\} \\
&= 2\text{Re}\{\text{vec}\{\tilde{\mathbf{Q}}'_m\}^H \text{vec}\{\mathbf{\Pi}_A^\perp \tilde{\mathbf{Q}}'_n \tilde{\mathbf{A}} \mathbf{\Delta} \tilde{\mathbf{A}}^H\}\} \\
&= 2\text{Re}\{\text{vec}\{\tilde{\mathbf{Q}}'_m\}^H [(\tilde{\mathbf{A}} \mathbf{\Delta} \tilde{\mathbf{A}}^H)^T \otimes \mathbf{\Pi}_A^\perp] \text{vec}\{\tilde{\mathbf{Q}}'_n\}\}, \tag{K.20}
\end{aligned}$$

from which we can see that $\mathbf{O}^H \mathbf{\Pi}_H^\perp \mathbf{O}$ is equal to the result in (K.17). We thus prove that the equality in (K.8) holds, and show $\mathbf{C}_{\mathbf{M},\rho}^{\text{as}} \leq \mathbf{CRB}_{\check{\mathbf{s}},\rho}$ and $\mathbf{C}_{\mathbf{M},\theta}^{\text{as}} \leq \mathbf{CRB}_{\check{\mathbf{s}},\theta}$.

If the noise is spatially white, we have $\mathbf{M} = \mathbf{0}$ and $\check{\mathbf{M}} = \mathbf{0}$ [43]. The proof of $\mathbf{C}_{\mathbf{M},\theta}^{\text{as}} \leq \mathbf{CRB}_{\check{\mathbf{s}},\theta}$ simplifies into

$$\begin{aligned}
\frac{1}{N} (\mathbf{C}_{\mathbf{M},\theta}^{\text{as}-1} - \mathbf{CRB}_{\check{\mathbf{s}},\theta}^{-1}) &= \frac{1}{N} \mathbf{CRB}_{\mathbf{D},\theta}^{\text{as}-1} + \mathbf{F} - \check{\mathbf{F}} \\
&= 2\text{Re}\left\{ \left(\tilde{\mathbf{D}}^H \mathbf{\Pi}_A^\perp \tilde{\mathbf{D}} \right) \odot \mathbf{\Delta}^T \right\} \geq \mathbf{0}. \tag{K.21}
\end{aligned}$$

Appendix L

Proof of Proposition 12

Let $\boldsymbol{\alpha} = [\boldsymbol{\theta}^T, \bar{\boldsymbol{\delta}}^T, \bar{\boldsymbol{b}}^T]^T$, where $\bar{\boldsymbol{\delta}} = [\boldsymbol{\delta}_1^T, \dots, \boldsymbol{\delta}_K^T]^T$ with $\boldsymbol{\delta}_k = [\boldsymbol{p}_k^T, \boldsymbol{\sigma}_k^T]^T$, and $\bar{\boldsymbol{b}} = [\boldsymbol{b}_1^T, \dots, \boldsymbol{b}_K^T]^T$ with $\boldsymbol{b}_k = [\boldsymbol{c}_k^T, \boldsymbol{\omega}_k^T]^T$. According to the Fisher information matrix equation [42], we have

$$\begin{aligned} [\mathbf{FIM}_M]_{mn} &= \sum_{k=1}^K 2\text{Re} \left\{ \sum_{t=1}^{N_k} \frac{\partial \boldsymbol{\mu}_k^H(t)}{\partial \alpha_m} \mathbf{R}_k^{-1} \frac{\partial \boldsymbol{\mu}_k(t)}{\partial \alpha_n} \right\} \\ &\quad + \sum_{k=1}^K N_k \text{trace} \left\{ \mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial \alpha_m} \mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial \alpha_n} \right\}, \end{aligned} \quad (\text{L.1})$$

where \mathbf{FIM}_M is the Fisher information matrix for $\boldsymbol{\alpha}$ based on the wide-band model, $[\mathbf{FIM}_M]_{mn}$ is the (m, n) -th element of \mathbf{FIM}_M , $\text{Re}\{\cdot\}$ denotes the real part of a complex value, $\boldsymbol{\mu}_k(t) = \mathbf{A}_k \mathbf{C}_k \boldsymbol{\varphi}_k(t)$, and α_m is the m -th element of $\boldsymbol{\alpha}$.

Let \mathbf{F}_D be a matrix of the same size as \mathbf{FIM}_M with $[\mathbf{F}_D]_{mn}$ equal to the first term on the right side of (L.1). If α_m is a parameter from $\bar{\boldsymbol{\delta}}$, we have $\frac{\partial \boldsymbol{\mu}_k(t)}{\partial \alpha_m} = \mathbf{0}$. Therefore,

$$\mathbf{F}_D = \begin{bmatrix} \mathbf{F}_{D.\boldsymbol{\theta}\boldsymbol{\theta}} & \mathbf{F}_{D.\boldsymbol{\theta}\bar{\boldsymbol{\delta}}} & \mathbf{F}_{D.\boldsymbol{\theta}\bar{\boldsymbol{b}}} \\ \mathbf{F}_{D.\bar{\boldsymbol{\delta}}\boldsymbol{\theta}} & \mathbf{F}_{D.\bar{\boldsymbol{\delta}}\bar{\boldsymbol{\delta}}} & \mathbf{F}_{D.\bar{\boldsymbol{\delta}}\bar{\boldsymbol{b}}} \\ \mathbf{F}_{D.\bar{\boldsymbol{b}}\boldsymbol{\theta}} & \mathbf{F}_{D.\bar{\boldsymbol{b}}\bar{\boldsymbol{\delta}}} & \mathbf{F}_{D.\bar{\boldsymbol{b}}\bar{\boldsymbol{b}}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{D.\boldsymbol{\theta}\boldsymbol{\theta}} & \mathbf{0} & \mathbf{F}_{D.\boldsymbol{\theta}\bar{\boldsymbol{b}}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{F}_{D.\bar{\boldsymbol{b}}\boldsymbol{\theta}} & \mathbf{0} & \mathbf{F}_{D.\bar{\boldsymbol{b}}\bar{\boldsymbol{b}}} \end{bmatrix}, \quad (\text{L.2})$$

where $\mathbf{F}_{D.\theta\theta}, \mathbf{F}_{D.\theta\bar{\delta}}, \dots, \mathbf{F}_{D.\bar{\delta}\bar{\delta}}$ are the sub-matrices of \mathbf{F}_D containing the elements related to the parameters specified in their subscripts. Similarly, let \mathbf{F}_S be a matrix of the same size as \mathbf{FIM}_M with $[\mathbf{F}_S]_{mn}$ equal to the second term on the right side of (L.1). Since $\frac{\partial \mathbf{R}_k}{\partial \alpha_m} = \mathbf{0}$ if α_m is a parameter from $\bar{\mathbf{b}}$,

$$\mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{S.\theta\theta} & \mathbf{F}_{S.\theta\bar{\delta}} & \mathbf{F}_{S.\theta\bar{b}} \\ \mathbf{F}_{S.\bar{\delta}\theta} & \mathbf{F}_{S.\bar{\delta}\bar{\delta}} & \mathbf{F}_{S.\bar{\delta}\bar{b}} \\ \mathbf{F}_{S.\bar{b}\theta} & \mathbf{F}_{S.\bar{b}\bar{\delta}} & \mathbf{F}_{S.\bar{b}\bar{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{S.\theta\theta} & \mathbf{F}_{S.\theta\bar{\delta}} & \mathbf{0} \\ \mathbf{F}_{S.\bar{\delta}\theta} & \mathbf{F}_{S.\bar{\delta}\bar{\delta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (\text{L.3})$$

From the results in (L.2) and (L.3), we obtain

$$\mathbf{FIM}_M = \mathbf{F}_D + \mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{D.\theta\theta} + \mathbf{F}_{S.\theta\theta} & \mathbf{F}_{S.\theta\bar{\delta}} & \mathbf{F}_{D.\theta\bar{b}} \\ \mathbf{F}_{S.\bar{\delta}\theta} & \mathbf{F}_{S.\bar{\delta}\bar{\delta}} & \mathbf{0} \\ \mathbf{F}_{D.\bar{b}\theta} & \mathbf{0} & \mathbf{F}_{D.\bar{b}\bar{b}} \end{bmatrix}, \quad (\text{L.4})$$

from which we have

$$\begin{aligned} \mathbf{CRB}_{M,\theta}^{-1} &= \mathbf{F}_{D.\theta\theta} + \mathbf{F}_{S.\theta\theta} - [\mathbf{F}_{S.\theta\bar{\delta}} \quad \mathbf{F}_{D.\theta\bar{b}}] \begin{bmatrix} \mathbf{F}_{S.\bar{\delta}\bar{\delta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{D.\bar{b}\bar{b}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{F}_{S.\bar{\delta}\theta} \\ \mathbf{F}_{D.\bar{b}\theta} \end{bmatrix} \\ &= \mathbf{F}_{D.\theta\theta} - \mathbf{F}_{D.\theta\bar{b}} \mathbf{F}_{D.\bar{b}\bar{b}}^{-1} \mathbf{F}_{D.\bar{b}\theta}^H + \mathbf{F}_{S.\theta\theta} - \mathbf{F}_{S.\theta\bar{\delta}} \mathbf{F}_{S.\bar{\delta}\bar{\delta}}^{-1} \mathbf{F}_{S.\bar{\delta}\theta}^H. \end{aligned} \quad (\text{L.5})$$

Note that $\mathbf{F}_{D.\theta\theta} = \sum_{k=1}^K \mathbf{F}_{D_k.\theta\theta}$, $\mathbf{F}_{S.\theta\theta} = \sum_{k=1}^K \mathbf{F}_{S_k.\theta\theta}$, $\mathbf{F}_{D.\theta\bar{b}} = [\mathbf{F}_{D_1.\theta\bar{b}_1}, \dots, \mathbf{F}_{D_K.\theta\bar{b}_K}]$, $\mathbf{F}_{S.\theta\bar{\delta}} = [\mathbf{F}_{S_1.\theta\bar{\delta}_1}, \dots, \mathbf{F}_{S_K.\theta\bar{\delta}_K}]$, where

$$[\mathbf{F}_{D_k}]_{mn} = 2\text{Re} \left\{ \sum_{t=1}^{N_k} \frac{\partial \mu_k^H(t)}{\partial \alpha_{k,m}} \mathbf{R}_k^{-1} \frac{\partial \mu_k(t)}{\partial \alpha_{k,n}} \right\}, \quad (\text{L.6})$$

$$[\mathbf{F}_{S_k}]_{mn} = N_k \text{trace} \left\{ \mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial \alpha_{k,m}} \mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial \alpha_{k,n}} \right\}, \quad (\text{L.7})$$

with $\alpha_{k,n}$ the n -th element of $\boldsymbol{\alpha}_k = [\boldsymbol{\theta}^T, \boldsymbol{\delta}_k^T, \mathbf{b}_k^T]^T$.

Also we have

$$\mathbf{F}_{\mathbf{D}\text{-}\bar{\mathbf{b}}\bar{\mathbf{b}}} = \text{blkdiag}\{\mathbf{F}_{\mathbf{D}_1\text{-}\mathbf{b}_1\mathbf{b}_1}, \dots, \mathbf{F}_{\mathbf{D}_K\text{-}\mathbf{b}_K\mathbf{b}_K}\}, \quad (\text{L.8})$$

$$\mathbf{F}_{\mathbf{S}\text{-}\bar{\boldsymbol{\delta}}\bar{\boldsymbol{\delta}}} = \text{blkdiag}\{\mathbf{F}_{\mathbf{S}_1\text{-}\boldsymbol{\delta}_1\boldsymbol{\delta}_1}, \dots, \mathbf{F}_{\mathbf{S}_K\text{-}\boldsymbol{\delta}_K\boldsymbol{\delta}_K}\}, \quad (\text{L.9})$$

which are two block-diagonal matrices with the k -th diagonal blocks equal to $\mathbf{F}_{\mathbf{D}_k\text{-}\mathbf{b}_k\mathbf{b}_k}$ and $\mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\delta}_k\boldsymbol{\delta}_k}$, respectively. We therefore have

$$\begin{aligned} \text{CRB}_{\mathbf{M}\text{-}\boldsymbol{\theta}}^{-1} &= \sum_{k=1}^K \left(\mathbf{F}_{\mathbf{D}_k\text{-}\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{D}_k\text{-}\boldsymbol{\theta}\mathbf{b}_k} \mathbf{F}_{\mathbf{D}_k\text{-}\mathbf{b}_k\mathbf{b}_k}^{-1} \mathbf{F}_{\mathbf{D}_k\text{-}\mathbf{b}_k\mathbf{b}_k}^H \mathbf{F}_{\mathbf{D}_k\text{-}\boldsymbol{\theta}\mathbf{b}_k} \right. \\ &\quad \left. + \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\theta}\boldsymbol{\delta}_k} \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\delta}_k\boldsymbol{\delta}_k}^{-1} \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\delta}_k\boldsymbol{\delta}_k}^H \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\theta}\boldsymbol{\delta}_k} \right), \end{aligned} \quad (\text{L.10})$$

where $\mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\theta}\boldsymbol{\delta}_k} \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\delta}_k\boldsymbol{\delta}_k}^{-1} \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\delta}_k\boldsymbol{\delta}_k}^H \mathbf{F}_{\mathbf{S}_k\text{-}\boldsymbol{\theta}\boldsymbol{\delta}_k} = \text{CRB}_{\mathbf{S}_k\text{-}\boldsymbol{\theta}}^{-1}$. Using the result in [57], we have

$$\begin{aligned} &\mathbf{F}_{\mathbf{D}_k\text{-}\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\mathbf{D}_k\text{-}\boldsymbol{\theta}\mathbf{b}_k} \mathbf{F}_{\mathbf{D}_k\text{-}\mathbf{b}_k\mathbf{b}_k}^{-1} \mathbf{F}_{\mathbf{D}_k\text{-}\mathbf{b}_k\mathbf{b}_k}^H \mathbf{F}_{\mathbf{D}_k\text{-}\boldsymbol{\theta}\mathbf{b}_k} \\ &= 2N_k \text{Re} \left\{ \left(\bar{\mathbf{D}}_k^H \boldsymbol{\Pi}_{\bar{\mathbf{A}}_k}^\perp \bar{\mathbf{D}}_k \right) \odot \left(\mathbf{C}_k \hat{\mathbf{R}}_{\phi_k \phi_k} \mathbf{C}_k^H \right)^T \right\} \\ &= 2N_k \text{Re} \left\{ \left(\tilde{\mathbf{D}}_k^H \boldsymbol{\Pi}_{\tilde{\mathbf{A}}_k}^\perp \tilde{\mathbf{D}}_k \right) \odot \left(\mathbf{C}_k \hat{\mathbf{R}}_{\phi_k \phi_k} \mathbf{C}_k^H \right)^T \right\} \\ &= \text{CRB}_{\mathbf{D}_k\text{-}\boldsymbol{\theta}}^{-1}, \end{aligned} \quad (\text{L.11})$$

where $\bar{\mathbf{D}}_k = \mathbf{R}_k^{-\frac{1}{2}} \mathbf{D}_k$, $\bar{\mathbf{A}}_k = \mathbf{R}_k^{-\frac{1}{2}} \mathbf{A}_k$, $\tilde{\mathbf{D}}_k = \mathbf{Q}_k^{-\frac{1}{2}} \mathbf{D}_k$, $\hat{\mathbf{R}}_{\phi_k \phi_k} = \frac{1}{N_k} \sum_{t=1}^{N_k} \boldsymbol{\varphi}_k(t) \boldsymbol{\varphi}_k^H(t)$, and \odot denotes the Hadamard product. The second equality in (L.11) holds from the result $\mathbf{R}_k^{-\frac{1}{2}} \boldsymbol{\Pi}_{\bar{\mathbf{A}}_k}^\perp \mathbf{R}_k^{-\frac{1}{2}} = \mathbf{Q}_k^{-\frac{1}{2}} \boldsymbol{\Pi}_{\tilde{\mathbf{A}}_k}^\perp \mathbf{Q}_k^{-\frac{1}{2}}$ given in [43]. We thus show (4.13).

Appendix M

Proof of Proposition 13

We first prove the ML estimation is consistent, since this conventional conclusion does not straightforwardly holds here due to the temporally non-stationary measurements.

Let $\mathcal{L}_k(\boldsymbol{\alpha}_k) = -\log |\mathbf{R}_k| - \text{trace} \{ \mathbf{R}_k^{-1} \mathbf{R} \mathbf{y}_k \mathbf{y}_k^T \}$, and $\boldsymbol{\alpha}_{k0}$ and $\boldsymbol{\alpha}_0$ be the true values of $\boldsymbol{\alpha}_k$ and $\boldsymbol{\alpha}$, respectively. From the narrow-band result (see Appendix C in [67]), we have $\mathcal{L}_k(\boldsymbol{\alpha}_k)$ converges with probability 1 as $N_k \rightarrow \infty$ to an asymptotic value $L_k^{\text{as}}(\boldsymbol{\alpha}_k)$ satisfying $\mathcal{L}_k^{\text{as}}(\boldsymbol{\alpha}_k) \leq \mathcal{L}_k^{\text{as}}(\boldsymbol{\alpha}_{k0})$. Let $N = \sum_{k=1}^K N_k$, we rewrite the LL function as $\mathcal{L}(\boldsymbol{\alpha}) = \sum_{k=1}^K \frac{N_k}{N} \mathcal{L}_k(\boldsymbol{\alpha}_k)$. Since $\frac{N_k}{N} \rightarrow h_k = \tau_k^{-1} / \sum_{k=1}^K \tau_k^{-1}$ as $N \rightarrow \infty$, we have $\mathcal{L}(\boldsymbol{\alpha})$ converges to $\mathcal{L}^{\text{as}}(\boldsymbol{\alpha}) = \sum_{k=1}^K h_k \mathcal{L}_k^{\text{as}}(\boldsymbol{\alpha}_k)$ with probability 1 as $N \rightarrow \infty$ and $\mathcal{L}^{\text{as}}(\boldsymbol{\alpha}) \leq \mathcal{L}^{\text{as}}(\boldsymbol{\alpha}_0)$. Assuming a unique maximum point for the LL function, we have the ML estimation of $\boldsymbol{\alpha}$ is consistent since the LL function is differentiable [62].

We now examine the asymptotic error covariance matrix of the mixed estimator. Let $f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\alpha}_k)$ be the probability density function (pdf) of \mathbf{Y}_k . Applying a Taylor expansion around $\boldsymbol{\alpha}_0$ to $\mathcal{L}(\boldsymbol{\alpha})$, we have

$$\begin{aligned} \mathbf{\Gamma}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) = & - \left[\mathbf{\Gamma}^{-1} \left(\sum_{k=1}^K \frac{\partial^2 \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\alpha}_k)}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} \right) \mathbf{\Gamma}^{-1} \right]^{-1} \Big|_{\boldsymbol{\alpha}=\tilde{\boldsymbol{\alpha}}} \\ & \times \left[\mathbf{\Gamma}^{-1} \sum_{k=1}^K \frac{\partial \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\alpha}_k)}{\partial \boldsymbol{\alpha}} \right] \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}, \end{aligned} \quad (\text{M.1})$$

where $\hat{\boldsymbol{\alpha}}$ is the ML estimate of $\boldsymbol{\alpha}$, $\tilde{\boldsymbol{\alpha}}$ is a vector between $\boldsymbol{\alpha}_0$ and $\hat{\boldsymbol{\alpha}}$, and $\mathbf{\Gamma}$ is a diagonal matrix with the number of rows (columns) equal to the length of $\boldsymbol{\alpha}$ and the m -th diagonal element equal to: (i) \sqrt{N} if α_m is one of $\boldsymbol{\theta}$; (ii) $\sqrt{N_k}$ if α_m is one of $\boldsymbol{\delta}_k$ or \boldsymbol{c}_k ; (iii) $N_k \sqrt{N_k}$ if α_m is one of $\boldsymbol{\omega}_k$. For simplicity, we omit the subscript $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$ in the following equations.

Let $l_k = l_{k1} + l_{k2}$ be the length of $\boldsymbol{\alpha}_k$ with l_{k1} and l_{k2} the lengths of $[\boldsymbol{\theta}^T, \boldsymbol{\delta}_k^T, \boldsymbol{c}_k^T]^T$ and $\boldsymbol{\omega}_k$, respectively. Let $\boldsymbol{\gamma}_{k,t} = \mathbf{\Gamma}_k^{-1} \partial \ln f_{\mathbf{y}_k(t)}(\mathbf{y}_k(t); \boldsymbol{\alpha}_k) / \partial \boldsymbol{\alpha}_k$, where $\mathbf{\Gamma}_k$ is an $l_k \times l_k$ diagonal matrix with the first l_{k1} diagonal elements equal to $\sqrt{N_k}$ and the last l_{k2} diagonal elements equal to $N_k \sqrt{N_k}$. According to the narrow-band results in [67], we have $\sum_{t=1}^{N_k} \text{cov}\{\boldsymbol{\gamma}_{k,t}\} \rightarrow \mathbf{F}_{d_k}^{\text{as}} + \mathbf{F}_{s_k}$ as N_k increases and $\sum_{t=1}^{N_k} \boldsymbol{\gamma}_{k,t}$ asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_{d_k}^{\text{as}} + \mathbf{F}_{s_k})$, where expressions of $\mathbf{F}_{d_k}^{\text{as}}$ and \mathbf{F}_{s_k} can be obtained by replacing all the metrics in the equations of \mathbf{F}_d^{as} and \mathbf{F}_s in [67] with their counterparts from the k -th sub-band. Since $\mathbf{F}_{d_k - \boldsymbol{\alpha}_k \boldsymbol{\delta}_k}^{\text{as}} = \mathbf{0}$, $\mathbf{F}_{s_k - \boldsymbol{\alpha}_k \boldsymbol{b}_k} = \mathbf{0}$ (see [67]), and $N_k/N \rightarrow h_k$ as $N \rightarrow \infty$, we obtain the second term on the right side of (M.1) asymptotically follows a zero-mean Gaussian distribution with covariance matrix

$$\mathbf{F}_M^{\text{as}} = \begin{bmatrix} \sum_{k=1}^K h_k (\mathbf{F}_{d_k - \boldsymbol{\theta}\boldsymbol{\theta}}^{\text{as}} + \mathbf{F}_{s_k - \boldsymbol{\theta}\boldsymbol{\theta}}) & \mathbf{F}_{s - \boldsymbol{\theta}\bar{\boldsymbol{\delta}}} & \mathbf{F}_{d - \boldsymbol{\theta}\bar{\boldsymbol{b}}}^{\text{as}} \\ \mathbf{F}_{s - \bar{\boldsymbol{\delta}}\boldsymbol{\theta}} & \mathbf{F}_{s - \bar{\boldsymbol{\delta}}\bar{\boldsymbol{\delta}}} & \mathbf{0} \\ \mathbf{F}_{d - \bar{\boldsymbol{b}}\boldsymbol{\theta}}^{\text{as}} & \mathbf{0} & \mathbf{F}_{d - \bar{\boldsymbol{b}}\bar{\boldsymbol{b}}}^{\text{as}} \end{bmatrix}, \quad (\text{M.2})$$

where

$$\mathbf{F}_{\mathbf{d},\boldsymbol{\theta}\bar{\mathbf{b}}}^{\text{as}} = [\sqrt{h_1}\mathbf{F}_{\mathbf{d}_1,\boldsymbol{\theta}\mathbf{b}_1}^{\text{as}}, \dots, \sqrt{h_K}\mathbf{F}_{\mathbf{d}_K,\boldsymbol{\theta}\mathbf{b}_K}^{\text{as}}], \quad (\text{M.3})$$

$$\mathbf{F}_{\mathbf{s},\boldsymbol{\theta}\bar{\boldsymbol{\delta}}} = [\sqrt{h_1}\mathbf{F}_{\mathbf{s}_1,\boldsymbol{\theta}\boldsymbol{\delta}_1}, \dots, \sqrt{h_K}\mathbf{F}_{\mathbf{s}_K,\boldsymbol{\theta}\boldsymbol{\delta}_K}], \quad (\text{M.4})$$

$$\mathbf{F}_{\mathbf{d},\bar{\mathbf{b}}\bar{\mathbf{b}}}^{\text{as}} = \text{blkdiag}\{\mathbf{F}_{\mathbf{d}_1,\mathbf{b}_1\mathbf{b}_1}^{\text{as}}, \dots, \mathbf{F}_{\mathbf{d}_K,\mathbf{b}_K\mathbf{b}_K}^{\text{as}}\}, \quad (\text{M.5})$$

$$\mathbf{F}_{\mathbf{s},\bar{\boldsymbol{\delta}}\bar{\boldsymbol{\delta}}} = \text{blkdiag}\{\mathbf{F}_{\mathbf{s}_1,\boldsymbol{\delta}_1\boldsymbol{\delta}_1}, \dots, \mathbf{F}_{\mathbf{s}_K,\boldsymbol{\delta}_K\boldsymbol{\delta}_K}\}. \quad (\text{M.6})$$

Since $\hat{\boldsymbol{\alpha}}$ is consistent, by the narrow-band result in [67], we have

$$\Gamma_k^{-1} \frac{\partial^2 \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\alpha}_k)}{\partial \boldsymbol{\alpha}_k \partial \boldsymbol{\alpha}_k^T} \Gamma_k^{-1} \Big|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}} \rightarrow -(\mathbf{F}_{\mathbf{d}_k}^{\text{as}} + \mathbf{F}_{\mathbf{s}_k}) \quad (\text{M.7})$$

as $N_k \rightarrow \infty$. As a consequence,

$$\left[\Gamma^{-1} \left(\sum_{k=1}^K \frac{\partial^2 \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\alpha}_k)}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} \right) \Gamma^{-1} \right] \rightarrow -\mathbf{F}_{\mathbf{M}}^{\text{as}} \quad (\text{M.8})$$

when N increases. We therefore have $\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0$ follows $\mathcal{N}(\mathbf{0}, (\Gamma \mathbf{F}_{\mathbf{M}}^{\text{as}} \Gamma)^{-1})$ asymptotically, where

$$\Gamma \mathbf{F}_{\mathbf{M}}^{\text{as}} \Gamma = \begin{bmatrix} \sum_{k=1}^K N_k (\mathbf{F}_{\mathbf{d}_k,\boldsymbol{\theta}\boldsymbol{\theta}}^{\text{as}} + \mathbf{F}_{\mathbf{s}_k,\boldsymbol{\theta}\boldsymbol{\theta}}) & \mathbf{F}_{\mathbf{S},\boldsymbol{\theta}\bar{\boldsymbol{\delta}}} & \mathbf{F}_{\mathbf{D},\boldsymbol{\theta}\bar{\mathbf{b}}}^{\text{as}} \\ \mathbf{F}_{\mathbf{S},\bar{\boldsymbol{\delta}}\boldsymbol{\theta}} & \mathbf{F}_{\mathbf{S},\bar{\boldsymbol{\delta}}\bar{\boldsymbol{\delta}}} & \mathbf{0} \\ \mathbf{F}_{\mathbf{D},\bar{\mathbf{b}}\boldsymbol{\theta}}^{\text{as}} & \mathbf{0} & \mathbf{F}_{\mathbf{D},\bar{\mathbf{b}}\bar{\mathbf{b}}}^{\text{as}} \end{bmatrix} \quad (\text{M.9})$$

with

$$\mathbf{F}_{\mathbf{D},\boldsymbol{\theta}\bar{\mathbf{b}}}^{\text{as}} = [N_1 \mathbf{F}_{\mathbf{d}_1,\boldsymbol{\theta}\mathbf{c}_1}^{\text{as}}, N_1^2 \mathbf{F}_{\mathbf{d}_1,\boldsymbol{\theta}\boldsymbol{\omega}_1}^{\text{as}}, \dots, N_K \mathbf{F}_{\mathbf{d}_K,\boldsymbol{\theta}\mathbf{c}_K}^{\text{as}}, N_K^2 \mathbf{F}_{\mathbf{d}_K,\boldsymbol{\theta}\boldsymbol{\omega}_K}^{\text{as}}], \quad (\text{M.10})$$

$$\mathbf{F}_{\mathbf{D},\bar{\mathbf{b}}\bar{\mathbf{b}}}^{\text{as}} = \text{blkdiag}\{N_1 \mathbf{F}_{\mathbf{d}_1,\mathbf{c}_1\mathbf{c}_1}^{\text{as}}, N_1^3 \mathbf{F}_{\mathbf{d}_1,\boldsymbol{\omega}_1\boldsymbol{\omega}_1}^{\text{as}}, \dots, N_K \mathbf{F}_{\mathbf{d}_K,\mathbf{c}_K\mathbf{c}_K}^{\text{as}}, N_K^3 \mathbf{F}_{\mathbf{d}_K,\boldsymbol{\omega}_K\boldsymbol{\omega}_K}^{\text{as}}\}, \quad (\text{M.11})$$

$$\mathbf{F}_{\mathbf{S},\boldsymbol{\theta}\bar{\boldsymbol{\delta}}} = [N_1 \mathbf{F}_{\mathbf{s}_1,\boldsymbol{\theta}\boldsymbol{\delta}_1}, \dots, N_K \mathbf{F}_{\mathbf{s}_K,\boldsymbol{\theta}\boldsymbol{\delta}_K}], \quad (\text{M.12})$$

$$\mathbf{F}_{\mathbf{S},\bar{\boldsymbol{\delta}}\bar{\boldsymbol{\delta}}} = \text{blkdiag}\{N_1 \mathbf{F}_{\mathbf{s}_1,\boldsymbol{\delta}_1\boldsymbol{\delta}_1}, \dots, N_K \mathbf{F}_{\mathbf{s}_K,\boldsymbol{\delta}_K\boldsymbol{\delta}_K}\}. \quad (\text{M.13})$$

Note that in the derivation of (M.9), we assume $N_k/N = h_k$ without affecting the asymptotic property.

Using the results in [42], we can check that the covariance matrix of $\sum_{k=1}^K \frac{\partial \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\alpha}_k)}{\partial \boldsymbol{\alpha}}$, which asymptotically equals $\mathbf{\Gamma} \mathbf{F}_M^{\text{as}} \mathbf{\Gamma}$, is the Fisher information matrix on $\boldsymbol{\alpha}$. Hence, $(\mathbf{\Gamma} \mathbf{F}_M^{\text{as}} \mathbf{\Gamma})^{-1}$ is the asymptotic CRB on $\boldsymbol{\alpha}$ and we prove the proposition. Note that the asymptotic CRB on $\boldsymbol{\theta}$ can be obtained by using (4.13) and replacing $\hat{\mathbf{R}}_{\phi_k \phi_k}$ in (L.11) with an identity matrix of the same size.

Appendix N

Proof of Proposition 14

For DOA estimation of wide-band mixed signals using the stochastic estimator, the objective function to be maximized is the stochastic LL function

$$\mathcal{L}(\boldsymbol{\xi}) = \sum_{k=1}^K \frac{N_k}{N} \left\{ -\log |\mathbf{A}_k \check{\mathbf{P}}_k \mathbf{A}_k^H + \mathbf{Q}_k| - \text{trace} \left\{ [\mathbf{A}_k \check{\mathbf{P}}_k \mathbf{A}_k^H + \mathbf{Q}_k]^{-1} \frac{\mathbf{Y}_k \mathbf{Y}_k^H}{N_k} \right\} \right\}, \quad (\text{N.1})$$

where $\boldsymbol{\xi} = [\boldsymbol{\theta}^T, \boldsymbol{\sigma}^T, \check{\mathbf{p}}^T]^T$ with $\check{\mathbf{p}} = [\check{\mathbf{p}}_1^T, \dots, \check{\mathbf{p}}_K^T]^T$ and $\boldsymbol{\sigma} = [\boldsymbol{\sigma}_1^T, \dots, \boldsymbol{\sigma}_K^T]^T$. As N_k increases, $\mathbf{Y}_k \mathbf{Y}_k^H / N_k$ converges with probability 1 to the true value of $\mathbf{A}_k \check{\mathbf{P}}_k \mathbf{A}_k^H + \mathbf{Q}_k$. As a result, the asymptotic function of $\mathcal{L}(\boldsymbol{\xi})$ achieves its maximum at the true value of $\boldsymbol{\xi}$ and the stochastic estimator is consistent for $\boldsymbol{\xi}$.

Let $f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\xi}_k)$ be the pdf of \mathbf{Y}_k based on the stochastic measurement model, where $\boldsymbol{\xi}_k = [\boldsymbol{\theta}^T, \boldsymbol{\sigma}_k^T, \check{\mathbf{p}}_k^T]^T$. Assuming $\hat{\boldsymbol{\xi}}$ is the stochastic estimate of $\boldsymbol{\xi}$ and applying a Taylor expansion to $\mathcal{L}(\boldsymbol{\xi})$ around $\boldsymbol{\xi}_0$, the true value of $\boldsymbol{\xi}$, we have

$$\begin{aligned} \sqrt{N}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0) &= -\left[\frac{1}{N} \left(\sum_{k=1}^K \frac{\partial^2 \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\xi}_k)}{\partial \boldsymbol{\xi} \partial \boldsymbol{\xi}^T} \right) \right]^{-1} \Big|_{\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}} \\ &\quad \times \left[\frac{1}{\sqrt{N}} \sum_{k=1}^K \frac{\partial \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\xi}_k)}{\partial \boldsymbol{\xi}} \right] \Big|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}, \end{aligned} \quad (\text{N.2})$$

where $\tilde{\boldsymbol{\xi}}$ is a vector value between $\boldsymbol{\xi}_0$ and $\hat{\boldsymbol{\xi}}$. For simplicity, we omit the subscript $\boldsymbol{\xi} = \boldsymbol{\xi}_0$ in all the following equations. From the narrow-band result in [67] and the consistency of $\hat{\boldsymbol{\xi}}$, we have as $N_k \rightarrow \infty$,

$$-\frac{1}{N_k} \frac{\partial^2 \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\xi}_k)}{\partial \xi_{k,m} \partial \xi_{k,n}} \Big|_{\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}} \rightarrow \text{trace} \left\{ \check{\mathbf{R}}_k^{-1} \frac{\partial \check{\mathbf{R}}_k}{\partial \xi_{k,m}} \check{\mathbf{R}}_k^{-1} \frac{\partial \check{\mathbf{R}}_k}{\partial \xi_{k,n}} \right\}. \quad (\text{N.3})$$

As a result, we have as $N \rightarrow \infty$,

$$-\frac{1}{N} \left(\sum_{k=1}^K \frac{\partial^2 \ln f_{\mathbf{Y}_k}(\mathbf{Y}_k; \boldsymbol{\xi}_k)}{\partial \boldsymbol{\xi} \partial \boldsymbol{\xi}^T} \right) \rightarrow \mathbf{F}_{\check{\mathbf{S}}}^{\text{as}} = \begin{bmatrix} \sum_{k=1}^K h_k \mathbf{F}_{\check{\mathbf{S}}_k-\boldsymbol{\theta}\boldsymbol{\theta}} & \mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\theta}\boldsymbol{\sigma}}^{\text{as}} & \mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\theta}\check{\mathbf{p}}}^{\text{as}} \\ \mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\sigma}\boldsymbol{\theta}}^{\text{as}} & \mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\sigma}\boldsymbol{\sigma}}^{\text{as}} & \mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\sigma}\check{\mathbf{p}}}^{\text{as}} \\ \mathbf{F}_{\check{\mathbf{S}}-\check{\mathbf{p}}\boldsymbol{\theta}}^{\text{as}} & \mathbf{F}_{\check{\mathbf{S}}-\check{\mathbf{p}}\boldsymbol{\sigma}}^{\text{as}} & \mathbf{F}_{\check{\mathbf{S}}-\check{\mathbf{p}}\check{\mathbf{p}}}^{\text{as}} \end{bmatrix}, \quad (\text{N.4})$$

in which

$$\mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\theta}\boldsymbol{\sigma}}^{\text{as}} = [h_1 \mathbf{F}_{\check{\mathbf{S}}_1-\boldsymbol{\theta}\boldsymbol{\sigma}_1}, \dots, h_K \mathbf{F}_{\check{\mathbf{S}}_K-\boldsymbol{\theta}\boldsymbol{\sigma}_K}], \quad (\text{N.5})$$

$$\mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\sigma}\boldsymbol{\sigma}}^{\text{as}} = \text{blkdiag}\{h_1 \mathbf{F}_{\check{\mathbf{S}}_1-\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1}, \dots, h_K \mathbf{F}_{\check{\mathbf{S}}_K-\boldsymbol{\sigma}_K\boldsymbol{\sigma}_K}\}, \quad (\text{N.6})$$

$$\mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\theta}\check{\mathbf{p}}}^{\text{as}} = [h_1 \mathbf{F}_{\check{\mathbf{S}}_1-\boldsymbol{\theta}\check{\mathbf{p}}_1}, \dots, h_K \mathbf{F}_{\check{\mathbf{S}}_K-\boldsymbol{\theta}\check{\mathbf{p}}_K}], \quad (\text{N.7})$$

$$\mathbf{F}_{\check{\mathbf{S}}-\boldsymbol{\sigma}\check{\mathbf{p}}}^{\text{as}} = \text{blkdiag}\{h_1 \mathbf{F}_{\check{\mathbf{S}}_1-\boldsymbol{\sigma}_1\check{\mathbf{p}}_1}, \dots, h_K \mathbf{F}_{\check{\mathbf{S}}_K-\boldsymbol{\sigma}_K\check{\mathbf{p}}_K}\}, \quad (\text{N.8})$$

$$\mathbf{F}_{\check{\mathbf{S}}-\check{\mathbf{p}}\check{\mathbf{p}}}^{\text{as}} = \text{blkdiag}\{h_1 \mathbf{F}_{\check{\mathbf{S}}_1-\check{\mathbf{p}}_1\check{\mathbf{p}}_1}, \dots, h_K \mathbf{F}_{\check{\mathbf{S}}_K-\check{\mathbf{p}}_K\check{\mathbf{p}}_K}\}, \quad (\text{N.9})$$

and $[\mathbf{F}_{\check{\mathbf{S}}_k}]_{mn}$ equals the expression on the right side of (N.3).

Let $\gamma_{k,t} = \frac{1}{\sqrt{N_k}} \frac{\ln f_{\mathbf{y}_k(t)}(\mathbf{y}_k(t); \boldsymbol{\xi}_k)}{\partial \boldsymbol{\xi}_k}$. According to the narrow-band result in [67], we have $\sum_{t=1}^{N_k} \gamma_{k,t}$ asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_{\check{\xi}_k} - \mathbf{F}_{\mu_k})$, where

$$[\mathbf{F}_{\mu_k}]_{mn} = \text{trace}\{\mathbf{K}_{k,m} \mathbf{A}_k \mathbf{C}_k \mathbf{C}_k^H \mathbf{A}_k^H \mathbf{K}_{k,n} \mathbf{A}_k \mathbf{C}_k \mathbf{C}_k^H \mathbf{A}_k^H\} \quad (\text{N.10})$$

with $\mathbf{K}_{k,m} = \check{\mathbf{R}}_k^{-1} \frac{\partial \check{\mathbf{R}}_k}{\partial \boldsymbol{\xi}_{k,m}} \check{\mathbf{R}}_k^{-1}$. We thus have the second term on the right side of (N.2) asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_{\check{\xi}}^{\text{as}} - \mathbf{F}_{\mu}^{\text{as}})$, where

$$\mathbf{F}_{\mu}^{\text{as}} = \begin{bmatrix} \sum_{k=1}^K h_k \mathbf{F}_{\mu_k \cdot \theta \theta} & \mathbf{F}_{\mu \cdot \theta \sigma}^{\text{as}} & \mathbf{F}_{\mu \cdot \theta \check{\mathbf{p}}}^{\text{as}} \\ \mathbf{F}_{\mu \cdot \sigma \theta}^{\text{as}} & \mathbf{F}_{\mu \cdot \sigma \sigma}^{\text{as}} & \mathbf{F}_{\mu \cdot \sigma \check{\mathbf{p}}}^{\text{as}} \\ \mathbf{F}_{\mu \cdot \check{\mathbf{p}} \theta}^{\text{as}} & \mathbf{F}_{\mu \cdot \check{\mathbf{p}} \sigma}^{\text{as}} & \mathbf{F}_{\mu \cdot \check{\mathbf{p}} \check{\mathbf{p}}}^{\text{as}} \end{bmatrix}, \quad (\text{N.11})$$

in which

$$\mathbf{F}_{\mu \cdot \theta \sigma}^{\text{as}} = [h_1 \mathbf{F}_{\mu_1 \cdot \theta \sigma_1}, \dots, h_K \mathbf{F}_{\mu_K \cdot \theta \sigma_K}], \quad (\text{N.12})$$

$$\mathbf{F}_{\mu \cdot \sigma \sigma}^{\text{as}} = \text{blkdiag}\{h_1 \mathbf{F}_{\mu_1 \cdot \sigma_1 \sigma_1}, \dots, h_K \mathbf{F}_{\mu_K \cdot \sigma_K \sigma_K}\}, \quad (\text{N.13})$$

$$\mathbf{F}_{\mu \cdot \theta \check{\mathbf{p}}}^{\text{as}} = [h_1 \mathbf{F}_{\mu_1 \cdot \theta \check{\mathbf{p}}_1}, \dots, h_K \mathbf{F}_{\mu_K \cdot \theta \check{\mathbf{p}}_K}], \quad (\text{N.14})$$

$$\mathbf{F}_{\mu \cdot \sigma \check{\mathbf{p}}}^{\text{as}} = \text{blkdiag}\{h_1 \mathbf{F}_{\mu_1 \cdot \sigma_1 \check{\mathbf{p}}_1}, \dots, h_K \mathbf{F}_{\mu_K \cdot \sigma_K \check{\mathbf{p}}_K}\}, \quad (\text{N.15})$$

$$\mathbf{F}_{\mu \cdot \check{\mathbf{p}} \check{\mathbf{p}}}^{\text{as}} = \text{blkdiag}\{h_1 \mathbf{F}_{\mu_1 \cdot \check{\mathbf{p}}_1 \check{\mathbf{p}}_1}, \dots, h_K \mathbf{F}_{\mu_K \cdot \check{\mathbf{p}}_K \check{\mathbf{p}}_K}\}. \quad (\text{N.16})$$

Therefore, we have $\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0$ asymptotically follows $\mathcal{N}(\mathbf{0}, \mathbf{F}_{\check{\xi}}^{-1} - \mathbf{F}_{\check{\xi}}^{-1} \mathbf{F}_{\mu} \mathbf{F}_{\check{\xi}}^{-1})$, where the expressions of $\mathbf{F}_{\check{\xi}}$ and \mathbf{F}_{μ} can be obtained by replacing h_k with N_k in (N.4) and (N.11), respectively.

Using the results from (108)-(114) in [67], we can similarly formulate $\mathbf{F}_{\mu_k} = \boldsymbol{\Phi}_k^H \mathbf{W}_k \boldsymbol{\Phi}_k$ and $\mathbf{F}_{\check{\xi}_k} = \boldsymbol{\Phi}_k^H \boldsymbol{\Phi}_k$, where

$$\Phi_k = [\mathbf{G}_k | \mathbf{U}_k | \mathbf{V}_k] = \left(\check{\mathbf{R}}_k^{-\frac{T}{2}} \otimes \check{\mathbf{R}}_k^{-\frac{1}{2}} \right) \left[\frac{\partial \check{\mathbf{r}}_k}{\partial \boldsymbol{\theta}^T} \middle| \frac{\partial \check{\mathbf{r}}_k}{\partial \boldsymbol{\sigma}_k^T} \middle| \frac{\partial \check{\mathbf{r}}_k}{\partial \check{\mathbf{p}}_k^T} \right], \quad (\text{N.17})$$

$$\mathbf{W}_k = (\check{\mathbf{A}}_k \mathbf{C}_k \mathbf{C}_k^H \check{\mathbf{A}}_k^H)^T \otimes (\check{\mathbf{A}}_k \mathbf{C}_k \mathbf{C}_k^H \check{\mathbf{A}}_k^H), \quad (\text{N.18})$$

with $\check{\mathbf{r}}_k = \text{vec}\{\check{\mathbf{R}}_k\}$, $\check{\mathbf{A}}_k = \check{\mathbf{R}}_k^{-\frac{1}{2}} \mathbf{A}_k$, and $\text{vec}\{\cdot\}$ the vectorization operator stacking all the columns of a matrix, one below another, into a vector. Substituting the expressions of \mathbf{F}_{μ_k} and $\mathbf{F}_{\check{\mathbf{s}}_k}$ back into \mathbf{F}_μ and $\mathbf{F}_{\check{\mathbf{s}}}$ and applying the block-wise matrix inversion (see (2.8.17) in [63]) to $\mathbf{F}_{\check{\mathbf{s}}}$, we obtain after algebraic matrix operations that

$$[\mathbf{F}_{\check{\mathbf{s}}}^{-1} \mathbf{F}_\mu \mathbf{F}_{\check{\mathbf{s}}}^{-1}]_{\rho\rho} = [\mathbf{F}_{\check{\mathbf{s}}}^{-1}]_{\rho\rho} \begin{bmatrix} \mathbf{F}_{\theta\theta} & \mathbf{F}_{\theta\sigma} \\ \mathbf{F}_{\theta\sigma}^H & \mathbf{F}_{\sigma\sigma} \end{bmatrix} [\mathbf{F}_{\check{\mathbf{s}}}^{-1}]_{\rho\rho}, \quad (\text{N.19})$$

where $[\mathbf{F}_{\check{\mathbf{s}}}^{-1} \mathbf{F}_\mu \mathbf{F}_{\check{\mathbf{s}}}^{-1}]_{\rho\rho}$ is the sub-matrix related to $\boldsymbol{\rho}$ in $\mathbf{F}_{\check{\mathbf{s}}}^{-1} \mathbf{F}_\mu \mathbf{F}_{\check{\mathbf{s}}}^{-1}$,

$$\mathbf{F}_{\theta\theta} = \sum_{k=1}^K N_k \mathbf{G}_k^H \Pi_{\mathbf{V}_k}^\perp \mathbf{W}_k \Pi_{\mathbf{V}_k}^\perp \mathbf{G}_k, \quad (\text{N.20})$$

$$\mathbf{F}_{\theta\sigma} = [N_1 \mathbf{G}_1^H \Pi_{\mathbf{V}_1}^\perp \mathbf{W}_1 \Pi_{\mathbf{V}_1}^\perp \mathbf{U}_1, \dots, N_K \mathbf{G}_K^H \Pi_{\mathbf{V}_K}^\perp \mathbf{W}_K \Pi_{\mathbf{V}_K}^\perp \mathbf{U}_K], \quad (\text{N.21})$$

$$\mathbf{F}_{\sigma\sigma} = \text{blkdiag}\{N_1 \mathbf{U}_1^H \Pi_{\mathbf{V}_1}^\perp \mathbf{W}_1 \Pi_{\mathbf{V}_1}^\perp \mathbf{U}_1, \dots, N_K \mathbf{U}_K^H \Pi_{\mathbf{V}_K}^\perp \mathbf{W}_K \Pi_{\mathbf{V}_K}^\perp \mathbf{U}_K\}. \quad (\text{N.22})$$

From the result in [67], we have $\Pi_{\mathbf{V}_k}^\perp \mathbf{W}_k \Pi_{\mathbf{V}_k}^\perp = \mathbf{0}$, $k = 1, \dots, K$. Consequently, $[\mathbf{F}_{\check{\mathbf{s}}}^{-1} \mathbf{F}_\mu \mathbf{F}_{\check{\mathbf{s}}}^{-1}]_{\rho\rho} = \mathbf{0}$, and the asymptotic error covariance matrix on $\boldsymbol{\rho}$ equals the sub-matrix related to $\boldsymbol{\rho}$ in $\mathbf{F}_{\check{\mathbf{s}}}^{-1}$. We can readily check that $\mathbf{F}_{\check{\mathbf{s}}}^{-1}$ is equal to $\mathbf{CRB}_{\check{\mathbf{s}}-\boldsymbol{\alpha}}$. We thus prove the proposition. Applying the block-wise matrix inversion to $\mathbf{F}_{\check{\mathbf{s}}}$, we obtain

$$\begin{aligned} \mathbf{CRB}_{\check{\mathbf{s}}-\boldsymbol{\theta}} &= \left[\sum_{k=1}^K N_k \left(\mathbf{F}_{\check{\mathbf{s}}_k-\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{\check{\mathbf{s}}_k-\boldsymbol{\theta}\mathbf{q}_k} \mathbf{F}_{\check{\mathbf{s}}_k-\mathbf{q}_k\mathbf{q}_k} \mathbf{F}_{\check{\mathbf{s}}_k-\boldsymbol{\theta}\mathbf{q}_k}^H \right) \right]^{-1} \\ &= \left(\sum_{k=1}^K \mathbf{CRB}_{\check{\mathbf{s}}_k-\boldsymbol{\theta}}^{-1} \right)^{-1}, \end{aligned} \quad (\text{N.23})$$

where $\mathbf{q}_k = [\boldsymbol{\sigma}_k^T, \check{\mathbf{p}}_k^T]^T$ and $\mathbf{CRB}_{\check{\mathbf{s}}_k-\boldsymbol{\theta}}$ is the stochastic CRB matrix on DOA estimation based on N_k measurements, each of which follows $\mathcal{CN}(\mathbf{0}, \check{\mathbf{R}}_k)$ with $\boldsymbol{\theta}$, $\check{\mathbf{p}}_k$, and $\boldsymbol{\sigma}_k$ the unknown parameters.

Appendix O

Proof of Proposition 16

Let $\boldsymbol{\beta} = [\boldsymbol{\lambda}^T, \boldsymbol{\delta}^T, \mathbf{c}^T]^T$, where $\boldsymbol{\lambda} = [\boldsymbol{\theta}^T, \boldsymbol{\omega}^T]^T$, $\boldsymbol{\delta} = [\boldsymbol{\delta}_1^T, \dots, \boldsymbol{\delta}_{N_d}^T]^T$ with $\boldsymbol{\delta}_n = [\mathbf{p}_n^T, \sigma_n]^T$, and $\mathbf{c} = [\mathbf{c}_1^T, \dots, \mathbf{c}_{N_d}^T]^T$. From the Fisher information matrix equation [42], we get

$$[\mathbf{FIM}_M]_{lm} = \sum_{n=1}^{N_d} 2\text{Re} \left\{ \sum_{t=1}^{T_d} \frac{\partial \boldsymbol{\mu}_n^H(t)}{\partial \beta_l} \mathbf{R}_n^{-1} \frac{\partial \boldsymbol{\mu}_n(t)}{\partial \beta_m} \right\} \quad (\text{O.1})$$

$$+ \sum_{n=1}^{N_d} T_d \text{trace} \left\{ \mathbf{R}_n^{-1} \frac{\partial \mathbf{R}_n}{\partial \beta_l} \mathbf{R}_n^{-1} \frac{\partial \mathbf{R}_n}{\partial \beta_m} \right\}, \quad (\text{O.2})$$

in which \mathbf{FIM}_M is the Fisher information matrix for $\boldsymbol{\beta}$ based on the model in (5.12), $[\mathbf{FIM}_M]_{lm}$ is the (l, m) -th element of \mathbf{FIM}_M , $\text{Re}\{\cdot\}$ denotes the real part of a complex metric, $\boldsymbol{\mu}_n(t) = \mathbf{A}_n \mathbf{C}_n \boldsymbol{\varphi}(t)$, and β_l is the l -th element of $\boldsymbol{\beta}$.

Let \mathbf{F}_D be a matrix of the same size as \mathbf{FIM}_M with $[\mathbf{F}_D]_{lm}$ equal to the first term on the right side of (O.2). If β_l is in $\boldsymbol{\delta}$, $\frac{\partial \boldsymbol{\mu}_n(t)}{\partial \beta_l} = \mathbf{0}$. Thus, we have

$$\mathbf{F}_D = \begin{bmatrix} F_{D,\lambda\lambda} & F_{D,\lambda\delta} & F_{D,\lambda c} \\ F_{D,\delta\lambda} & F_{D,\delta\delta} & F_{D,\delta c} \\ F_{D,c\lambda} & F_{D,c\delta} & F_{D,cc} \end{bmatrix} = \begin{bmatrix} F_{D,\lambda\lambda} & \mathbf{0} & F_{D,\lambda c} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ F_{D,c\lambda} & \mathbf{0} & F_{D,cc} \end{bmatrix}, \quad (\text{O.3})$$

where $\mathbf{F}_{D,\lambda\lambda}, \mathbf{F}_{D,\lambda\delta}, \dots, \mathbf{F}_{D,cc}$ are the sub-matrices of \mathbf{F}_D containing the elements related to the parameters specified in their subscripts. Similarly, let \mathbf{F}_S be a matrix of the same size as \mathbf{FIM}_M with $[\mathbf{F}_S]_{lm}$ equal to the second term on the right side of (O.2). Because $\frac{\partial \mathbf{R}_n}{\partial \beta_l} = \mathbf{0}$ if β_l is in \mathbf{c} , we have

$$\mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{S,\lambda\lambda} & \mathbf{F}_{S,\lambda\delta} & \mathbf{F}_{S,\lambda c} \\ \mathbf{F}_{S,\delta\lambda} & \mathbf{F}_{S,\delta\delta} & \mathbf{F}_{S,\delta c} \\ \mathbf{F}_{S,c\lambda} & \mathbf{F}_{S,c\delta} & \mathbf{F}_{S,cc} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{S,\lambda\lambda} & \mathbf{F}_{S,\lambda\delta} & \mathbf{0} \\ \mathbf{F}_{S,\delta\lambda} & \mathbf{F}_{S,\delta\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (\text{O.4})$$

Using (O.3) and (O.4), we obtain

$$\mathbf{FIM}_M = \mathbf{F}_D + \mathbf{F}_S = \begin{bmatrix} \mathbf{F}_{D,\lambda\lambda} + \mathbf{F}_{S,\lambda\lambda} & \mathbf{F}_{S,\lambda\delta} & \mathbf{F}_{D,\lambda c} \\ \mathbf{F}_{S,\delta\lambda} & \mathbf{F}_{S,\delta\delta} & \mathbf{0} \\ \mathbf{F}_{D,c\lambda} & \mathbf{0} & \mathbf{F}_{D,cc} \end{bmatrix} \quad (\text{O.5})$$

and

$$\begin{aligned} \mathbf{CRB}_{M,\lambda}^{-1} &= \mathbf{F}_{D,\lambda\lambda} + \mathbf{F}_{S,\lambda\lambda} - [\mathbf{F}_{S,\lambda\delta} \quad \mathbf{F}_{D,\lambda c}] \begin{bmatrix} \mathbf{F}_{S,\delta\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{D,cc} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{F}_{S,\delta\lambda} \\ \mathbf{F}_{D,c\lambda} \end{bmatrix} \\ &= \mathbf{F}_{D,\lambda\lambda} - \mathbf{F}_{D,\lambda c} \mathbf{F}_{D,cc}^{-1} \mathbf{F}_{D,\lambda c}^H + \mathbf{F}_{S,\lambda\lambda} - \mathbf{F}_{S,\lambda\delta} \mathbf{F}_{S,\delta\delta}^{-1} \mathbf{F}_{S,\lambda\delta}^H. \end{aligned} \quad (\text{O.6})$$

Let $\boldsymbol{\beta}_n = [\boldsymbol{\lambda}^T, \boldsymbol{\delta}_n^T, \mathbf{c}_n^T]^T$, $\beta_{n,l}$ be the l -th element of $\boldsymbol{\beta}_n$, and

$$[\mathbf{F}_{S_n}]_{lm} = T_d \text{trace} \left\{ \mathbf{R}_n^{-1} \frac{\partial \mathbf{R}_n}{\partial \beta_{n,l}} \mathbf{R}_n^{-1} \frac{\partial \mathbf{R}_n}{\partial \beta_{n,m}} \right\}. \quad (\text{O.7})$$

Since

$$\mathbf{F}_{S_{\lambda\lambda}} = \begin{bmatrix} \mathbf{F}_{S_{\theta\theta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N_d} \mathbf{F}_{S_n\theta\theta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (\text{O.8})$$

$$\mathbf{F}_{S_{\lambda\delta}} = \begin{bmatrix} \mathbf{F}_{S_1\theta\delta_1} & \cdots & \mathbf{F}_{S_{N_d}\theta\delta_{N_d}} \\ \mathbf{F}_{S_1\omega\delta_1} & \cdots & \mathbf{F}_{S_{N_d}\omega\delta_{N_d}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{S_1\theta\delta_1} & \cdots & \mathbf{F}_{S_{N_d}\theta\delta_{N_d}} \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}, \quad (\text{O.9})$$

and $\mathbf{F}_{S_{\delta\delta}} = \text{diag}\{\mathbf{F}_{S_n\delta_1\delta_1}, \cdots, \mathbf{F}_{S_n\delta_{N_d}\delta_{N_d}}\}$, we have

$$\mathbf{F}_{S_{\lambda\lambda}} - \mathbf{F}_{S_{\lambda\delta}}\mathbf{F}_{S_{\delta\delta}}^{-1}\mathbf{F}_{S_{\lambda\delta}}^H = \begin{bmatrix} \sum_{n=1}^{N_d} \mathbf{F}_{S_n\theta\theta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (\text{O.10})$$

where

$$\mathbf{F}_{S_n\theta\theta} = \mathbf{F}_{S_n\theta\theta} - \mathbf{F}_{S_n\theta\delta_n}\mathbf{F}_{S_n\delta_n\delta_n}^{-1}\mathbf{F}_{S_n\theta\delta_n}^H. \quad (\text{O.11})$$

Also, let

$$[\mathbf{F}_{D_n}]_{lm} = 2\text{Re}\left\{ \sum_{t=1}^{T_d} \frac{\partial \mu_n^H(t)}{\partial \beta_{n,l}} \mathbf{R}_n^{-1} \frac{\partial \mu_n(t)}{\partial \beta_{n,m}} \right\}. \quad (\text{O.12})$$

Since $\mathbf{F}_{D_{cc}} = \text{blkdiag}\{\mathbf{F}_{D_1c_1c_1}, \cdots, \mathbf{F}_{D_{N_d}c_{N_d}c_{N_d}}\}$, $\mathbf{F}_{D_{\lambda c}} = [\mathbf{F}_{D_1\lambda c_1}, \cdots, \mathbf{F}_{D_{N_d}\lambda c_{N_d}}]$, and $\mathbf{F}_{D_{\lambda\lambda}} = \sum_{n=1}^{N_d} \mathbf{F}_{D_n\lambda\lambda}$, we obtain

$$\mathbf{F}_{D_{\lambda\lambda}} - \mathbf{F}_{D_{\lambda c}}\mathbf{F}_{D_{cc}}^{-1}\mathbf{F}_{D_{\lambda c}}^H \quad (\text{O.13})$$

$$= \sum_{n=1}^{N_d} (\mathbf{F}_{D_n\lambda\lambda} - \mathbf{F}_{D_n\lambda c_n}\mathbf{F}_{D_n c_n c_n}^{-1}\mathbf{F}_{D_n\lambda c_n}^H) \quad (\text{O.14})$$

$$= \sum_{n=1}^{N_d} \left(\begin{bmatrix} \mathbf{F}_{D_n\theta\theta} & \mathbf{F}_{D_n\theta\omega} \\ \mathbf{F}_{D_n\omega\theta} & \mathbf{F}_{D_n\omega\omega} \end{bmatrix} - \begin{bmatrix} \mathbf{F}_{D_n\theta c_n} \\ \mathbf{F}_{D_n\omega c_n} \end{bmatrix} \mathbf{F}_{D_n c_n c_n}^{-1} [\mathbf{F}_{D_n c_n \theta} \quad \mathbf{F}_{D_n c_n \omega}] \right) \quad (\text{O.15})$$

$$= \sum_{n=1}^{N_d} \begin{bmatrix} \mathbf{F}_{D_n 11} & \mathbf{F}_{D_n 12} \\ \mathbf{F}_{D_n 21} & \mathbf{F}_{D_n 22} \end{bmatrix}, \quad (\text{O.16})$$

where

$$\mathbf{F}_{D_n.11} = \mathbf{F}_{D_n.\boldsymbol{\theta}\boldsymbol{\theta}} - \mathbf{F}_{D_n.\boldsymbol{\theta}\mathbf{c}_n} \mathbf{F}_{D_n.\mathbf{c}_n\mathbf{c}_n}^{-1} \mathbf{F}_{D_n.\boldsymbol{\theta}\mathbf{c}_n}^H, \quad (\text{O.17})$$

$$\mathbf{F}_{D_n.12} = \mathbf{F}_{D_n.21}^H = \mathbf{F}_{D_n.\boldsymbol{\theta}\boldsymbol{\omega}} - \mathbf{F}_{D_n.\boldsymbol{\theta}\mathbf{c}_n} \mathbf{F}_{D_n.\mathbf{c}_n\mathbf{c}_n}^{-1} \mathbf{F}_{D_n.\boldsymbol{\omega}\mathbf{c}_n}^H, \quad (\text{O.18})$$

$$\mathbf{F}_{D_n.22} = \mathbf{F}_{D_n.\boldsymbol{\omega}\boldsymbol{\omega}} - \mathbf{F}_{D_n.\boldsymbol{\omega}\mathbf{c}_n} \mathbf{F}_{D_n.\mathbf{c}_n\mathbf{c}_n}^{-1} \mathbf{F}_{D_n.\boldsymbol{\omega}\mathbf{c}_n}^H. \quad (\text{O.19})$$

The expressions for $\mathbf{F}_{D_n.\boldsymbol{\theta}\boldsymbol{\omega}}$, $\mathbf{F}_{D_n.\boldsymbol{\theta}\mathbf{c}_n}$, and $\mathbf{F}_{D_n.\mathbf{c}_n\mathbf{c}_n}$ are given in [67]. Substituting them into (O.18), we can show that $\mathbf{F}_{D_n.12} = \mathbf{0}$. Therefore, from (O.6), (O.10), and (O.16), we have

$$\mathbf{CRB}_{M.\boldsymbol{\theta}}^{-1} = \sum_{n=1}^{N_d} (\mathbf{F}_{D_n.11} + \mathbf{F}_{S_n.11}), \quad (\text{O.20})$$

where we can check that $\mathbf{F}_{S_n.11} = \mathbf{CRB}_{S_n.\boldsymbol{\theta}}^{-1}$ and $\mathbf{F}_{D_n.11} = \mathbf{CRB}_{D_n.\boldsymbol{\theta}}^{-1}$ (see [67]).

Appendix P

Derivation of Equation (6.16)

Note that

$$\begin{aligned}
 \mathbf{G}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) &= \mathbf{R}(\bar{\boldsymbol{\theta}}_l)^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{R}(\boldsymbol{\theta}) \\
 &= [\boldsymbol{\psi}_j \otimes \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)]_{1 \times J}^H (\mathbf{I}_N \otimes \boldsymbol{\Sigma}^{-1}) [\boldsymbol{\psi}_j \otimes \mathbf{A}_j(\boldsymbol{\theta})]_{1 \times J} \\
 &= [\boldsymbol{\psi}_j \otimes \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)]_{1 \times J}^H [\boldsymbol{\psi}_j \otimes (\boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta}))]_{1 \times J} \\
 &= [(\boldsymbol{\psi}_k^H \boldsymbol{\psi}_j) \otimes (\mathbf{A}_k(\bar{\boldsymbol{\theta}}_l)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta}))]_{J \times J} \tag{P.1}
 \end{aligned}$$

where $[\boldsymbol{\psi}_j \otimes \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)]_{1 \times J}$ is a matrix with $1 \times J$ blocks whose j th block is $\boldsymbol{\psi}_j \otimes \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)$ and $[(\boldsymbol{\psi}_k^H \boldsymbol{\psi}_j) \otimes (\mathbf{A}_k(\bar{\boldsymbol{\theta}}_l)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta}))]_{J \times J}$ is a matrix with $J \times J$ blocks whose (k, j) th block is $(\boldsymbol{\psi}_k^H \boldsymbol{\psi}_j) \otimes (\mathbf{A}_k(\bar{\boldsymbol{\theta}}_l)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta}))$.

Since $\boldsymbol{\psi}_k^H \boldsymbol{\psi}_j = \delta_{kj}$, we have

$$\mathbf{G}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) = \text{diag} [\mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta})]_J, \tag{P.2}$$

which is a block-diagonal matrix of J blocks with the j th block equal to $\mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta})$.

Using (6.14), we obtain

$$\begin{aligned}
\bar{\mathbf{c}}_l &= \mathbf{G}(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_l)^{-1} \mathbf{G}(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta}) \mathbf{c} \\
&= \text{diag} \left[(\mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\bar{\boldsymbol{\theta}}_l))^{-1} (\mathbf{A}_j(\bar{\boldsymbol{\theta}}_l)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}_j(\boldsymbol{\theta})) \right]_J \mathbf{c} \\
&= \text{diag} [\mathbf{P}_j(\bar{\boldsymbol{\theta}}_l, \bar{\boldsymbol{\theta}}_l)^{-1} \mathbf{P}_j(\bar{\boldsymbol{\theta}}_l, \boldsymbol{\theta})]_J \mathbf{c},
\end{aligned} \tag{P.3}$$

from which we derive (6.16).

References

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research - the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, pp. 67-94, Jul. 1996.
- [2] R. Roy and T. Kailath, "ESPRIT - estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984-995, Jul. 1989.
- [3] R. T. Compton, *Adaptive Antennas: Concepts and Performance*, Prentice Hall, 1988.
- [4] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, pp. 4-24, Apr. 1988.
- [5] C. Muravchik, and A. Nehorai, "Localization and identifying underground objects using gravitational measurements," *28th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 1994.
- [6] G. Bienvenu and L. Kopp, "Adaptivity to background noise spatial coherence for high resolution passive methods." *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 307-310, 1980.
- [7] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 720-741, May 1989.
- [8] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1783-1795, Oct. 1990.
- [9] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér-Rao bound: further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 2140-2150, Dec. 1990.

- [10] A. G. Jaffer, "Maximum likelihood direction finding of stochastic sources: a seperable solution," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 1988, pp. 2893-2896.
- [11] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 42, pp. 2481-2491, Sept. 1994.
- [12] J. LeCadre, "Parametric methods for spatial signal processing in the presence of unknown colored noise field," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 965-983, Jul. 1989.
- [13] H. Ye and R. D. DeGroat, "Maximum likelihood DOA estimation and asymptotic Cramr-Rao bounds for additive unknown colored noise," *IEEE Trans. Signal Process.*, vol. 43, pp. 938-949, Apr. 1995.
- [14] V. Nagesha and S. Kay, "Maximum likelihood estimation for array processing in colored noise," *IEEE Trans. Signal Process.*, vol. 44, pp. 169-180, Feb. 1996.
- [15] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Trans. Signal Process.*, vol. 43, pp. 2154-2163, Sept. 1995.
- [16] M. Viberg, P. Stoica, and B. Ottersten, "Maximum likelihood array processing in spatially correlated noise fields using parameterized signals," *IEEE Trans. Signal Process.*, vol. 45, pp. 996-1004, Apr. 1997.
- [17] P. Stoica, M. Viberg, and B. Ottersten, "Instrumental variable approach to array processing in spatially correlated noise fields," *IEEE Trans. Signal Process.*, vol. 42, pp 121-133, Jan. 1994.
- [18] M. Viberg, P. Stoica, and B. Ottersten, "Array processing in correlated noise fields based on instrumental variables and subspace fitting," *IEEE Trans. Signal Process.*, vol. 43, pp. 1187-1199, May 1995.
- [19] M. Pesavento and A. B. Gershman, "Maximum-likelihood direction of arrival estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 49, pp. 1310-1324, Jul. 2001.

- [20] C. E. Chen, F. Lorenzelli, R. E. Hudson, and K. Yao, "Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 56, pp. 3038-3044, Jul. 2008.
- [21] Q. Wu, K. M. Wong, and J. P. Reilly, "Maximum-likelihood direction-finding in unknown noise environments," *IEEE Trans. Signal Process.*, vol. 42, pp. 980-983, Apr. 1994.
- [22] Q. Wu and K. M. Wong, "UN-MUSIC and UN-CLE: an application of generalized correlation analysis to the estimation of the directions of arrival of signals in unknown correlated noise," *IEEE Trans. Signal Process.*, vol. 42, pp. 2331-2343, Sept. 1994.
- [23] P. Stoica, M. Viberg, K. M. Wong, and Q. Wu, "Maximum-likelihood bearing estimation with partly calibrated arrays in spatially correlated noise field," *IEEE Trans. Signal Process.*, vol. 44, pp. 888-899, Apr. 1996.
- [24] S. A. Vorobyov, A. B. Gershman, and K. M. Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, pp. 34-43, Jan. 2005.
- [25] P. Stoica, M. Agrawal, and P. Åhlgren, "Array processing for signals with non-zero means in colored noise fields," *Digital Signal Processing*, vol. 14, pp. 296-311, Jul. 2004.
- [26] M. S. Hämäläinen, R. Hari, R. Ilmoniemi, J. Knuutila, and O.V. Lounasmaa, "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of signal processing of the human brain," *Rev. Modern Phys.*, vol. 65, pp. 413-497, 1993.
- [27] M. Scherg and D. von Cramon, "Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model," *Electroenceph. Clin. Neurophysiol.*, vol. 62, pp. 32-44, Jan. 1985.
- [28] A. B. Carlsson, *Communications Engineering*, third ed., McGraw-Hill, 1986.
- [29] J. Proakis, *Digital Communications*, fourth ed., McGraw-Hill, 2000.

- [30] C. Muravchik, and A. Nehorai, "Localization and identifying underground objects using gravitational measurements," *28th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 1994.
- [31] R. J. Urick, *Principles of Underwater Sound*, third ed., McGraw-Hill, 1983.
- [32] A. Tkacenko and P. P. Vaidyanathan, "Sinusoidal frequency estimation using filter banks," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, pp.3089-3092, May. 2001.
- [33] K. W. Chan and H. C. So, "Accurate frequency estimation for real harmonic sinusoids," *IEEE Signal Process. Lett.*, vol. 11, pp. 609-612, Jul. 2004.
- [34] B. Halder and T. Kailath, "Efficient estimation of closely spaced sinusoidal frequencies using subspace-based methods," *IEEE Signal Process. Lett.*, vol. 4, pp. 49-51, Feb. 1997.
- [35] I. Djurović, "Estimation of the sinusoidal signal frequency based on the marginal median DFT," *IEEE Trans. Signal Process.*, vol. 55, pp. 2043-2051, May 2007.
- [36] M. Kristensson, M. Jansson, and B. Ottersten, "Further results and insights on subspace based sinusoidal frequency estimation," *IEEE Trans. Signal Process.*, vol. 49, pp. 2962-2974, Dec. 2001.
- [37] A. Dogandžić and A. Nehorai, "Generalized multivariate analysis of variance: a unified framework for signal processing in correlated noise," *IEEE Signal Processing Magazine*, vol. 20, pp. 39-54, Sept. 2003.
- [38] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc. B*, vol. 39, pp. 1-38, 1977.
- [39] X. Meng and D. V. Dyk, "The EM algorithm—an old folk-song sung to a fast new tune," *J. R. Statist. Soc. B*, vol. 59, pp. 511-567, 1997.
- [40] X. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: a general framework," *Biometrika*, vol. 80, pp. 267-278, 1993.

- [41] N. Laird, N. Lange, and D. Stram, "Maximum likelihood computations with repeated measures: application of the EM algorithm," *Journal of the American Statistical Association*, vol. 82, pp. 97-105, Mar. 1987.
- [42] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Upper Saddle River, NJ: Prentice-Hall, 1993.
- [43] A. B. Gershman, P. Stoica, M. Pesavento, and E. G. Larsson, "Stochastic Cramér-Rao bound for direction estimation in unknown noise fields," *IEE Proc. Radar Sonar Navig.*, vol. 149, pp. 2-8, Feb. 2002.
- [44] N. H. Timm, *Applied Multivariate Analysis*, Springer-Verlag, 2002.
- [45] X. Yang, G. Sun, M. Li, and Q. Li, "Investigation of the spatial correlations of flow noise in vector hydrophone towed linear array," *Chinese Journal of Acoustics*, vol. 27, No. 3, 2008.
- [46] P. H. M. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Trans. Automat. Contr.*, vol. 13, pp. 867-870, Sept. 1998.
- [47] D. Carlson, E. Haynsworth, and T. Markham, "A generalization of the Schur complement by means of the Moore-Penrose inverse", *SIAM J. Appl. Math.*, vol. 26, No. 1, pp. 169-175, Jan. 1974.
- [48] L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics*, second ed., John Wiley and Sons, 1962.
- [49] M. J. Crocker, *Handbook of Acoustics*, John Wiley and Sons, 1998.
- [50] P. Stoica, R. L. Moses, B. Friedlander, and T. Söderström, "Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements," *IEEE Trans. Signal Process.*, vol. 37, pp. 378-391, Mar. 1989.
- [51] S. Kay, "Mean likelihood frequency estimation," *IEEE Trans. Signal Process.*, vol. 48, pp. 1937-1946, Jul. 2000.
- [52] U. Sandkühler and J. F. Böhme, "Accuracy of maximum likelihood estimates for array processing," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Dallas, TX, Apr. 1987, pp. 2015-2018.

- [53] B. Friedlander and A. J. Weiss, "Direction finding using noise covariance modeling," *IEEE Trans. Signal Process.*, vol. 43, pp. 1557-1567, Jul. 1995.
- [54] M. A. Doron, A. J. Weiss, and H. Messer, "Maximum-likelihood direction finding of wide-band sources," *IEEE Trans. Signal Process.*, vol. 41, pp. 411-414, Jan. 1993.
- [55] T. Li and A. Nehorai, "Maximum Likelihood Direction Finding in Spatially Colored Noise Fields Using Sparse Sensor Arrays," *IEEE Trans. Signal Process.*, vol. 59, pp. 1048-1062, Mar. 2011.
- [56] A. Dogandžić and A. Nehorai, "Space-time fading channel estimation and symbol detection in unknown spatially correlated noise," *IEEE Trans. Signal Process.*, Vol. 50, pp. 457-474, Mar. 2002.
- [57] A. Dogandžić and A. Nehorai, "Space-time fading channel estimation and symbol detection in unknown spatially correlated noise," Dept. Elect. Eng. Comput. Sci., Univ. Illinois, Chicago, Rep. UIC-EECS-00-9, Aug. 2000.
- [58] A. W. Van Der Vaart, *Asymptotic Statistics*, Cambridge University Press, 1998.
- [59] P. Stoica, M. Viberg, K. M. Wong, and Q. Wu, "Maximum-likelihood bearing estimation with partly calibrated arrays in spatially correlated noise field," *IEEE Trans. Signal Process.*, vol. 44, pp. 888-899, Apr. 1996.
- [60] N. Higham, *Accuracy and stability of numerical algorithms*, second ed., SIAM, 2002.
- [61] Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, third ed., Springer, 1997.
- [62] C. R. Rao, *Linear Statistical Inference and Its Applications*, second ed., J. Wiley, 1973.
- [63] D. S. Bernstein, *Matrix Mathematics*, Princeton University Press, 2005.
- [64] A. Graham, *Kronecker products and matrix calculus with applications*, Ellis Horwood, London, 1981.

- [65] M. A. Doron and E. Doron, "Wavefield modeling and array processing, part III - resolution capacity," *IEEE Trans. Signal Process.*, vol. 42, pp. 2571-2580, Oct. 1994.
- [66] C. E. Chen, F. Lorenzelli, R. E. Hudson, and K. Yao, "Maximum likelihood DOA estimation of multiple wideband sources in the presence of nonuniform sensor noise," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, article ID 835079.
- [67] T. Li and A. Nehorai, "Maximum Likelihood Direction-of-Arrival Estimation of Underwater Acoustic Signals Containing Sinusoidal Components," *IEEE Trans. Signal Process.*, vol. 59, pp. 5302-5314, Nov. 2011.
- [68] P. Stoica and T. Soderstrom, "Statistical Analysis of MUSIC and Subspace Rotation Estimates of Sinusoidal Frequencies," *IEEE Trans. Signal Process.*, vol. 39, pp. 1836-1847, Aug. 1991.
- [69] S. M. Kay, V. Nagesh, and J. Slisbury, "Broadband detection based on two-dimensional mixed autoregressive model," *IEEE Trans. Signal Process.*, vol. 41, pp. 2413-2428, Jul. 1993.
- [70] M. Agrawal and S. Prasad, "DOA Estimation of Wideband Sources Using a Harmonic Source Model and Uniform Linear Array", *IEEE Trans. Signal Process.*, vol. 47, pp. 619-629, Mar. 1999.
- [71] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Statist.*, vol. 20, pp. 477-501, 1949.
- [72] J. M. Hammersley, "On Estimating Restricted Parameters," *Journal of the Royal Statistical Society, Series B*, Vol. 12, No. 2, pp. 192-240, 1950.
- [73] D. G. Chapman and H. Robbins, "Minimum Variance Estimation Without Regularity Assumptions," *The Annals of Mathematical Statistics*, Vol. 22, No. 4, pp. 581-586, Dec. 1951.
- [74] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Trans. Inf. Theory*, vol. 15, no. 2, pp. 273-279, Mar. 1969.

- [75] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. Inf. Theory*, vol. 17, pp. 669-676, Nov. 1971
- [76] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation," *IEEE Trans. Signal Process.*, vol. 41, pp. 3102-3113, Nov. 1993.
- [77] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation: Part 2 - threshold phenomena," *IEEE Trans. Signal Process.*, vol. 42, pp. 1001-1007, May 1994.
- [78] J. Tabrikian and J. L. Krolik, "Barankin bounds for source localization in an uncertain ocean environment" *IEEE Trans. Signal Process.*, vol. 47, pp. 2917-2927, Nov. 1999.
- [79] A. Pinkus and J. Tabrikian, "Barankin bound for range and Doppler estimation using orthogonal signal transmission," *Proc. of IEEE International Radar Conference*, Apr. 2006, pp. 94-99.
- [80] F. Athley, "Threshold region performance of maximum likelihood direction of arrival estimators," *IEEE Trans. Signal Process.*, Vol. 53, No. 4, pp. 1359-1373, Apr. 2005.
- [81] E. Chaumette, J. Galy, A. Quinlan, and P. Larzabal, "A new Barankin bound approximation for the prediction of the threshold region performance of maximum likelihood estimators," *IEEE Trans. Signal Process.*, vol. 56, pp. 5319-5333, Nov. 2008.
- [82] M. Hawkes and A. Nehorai, "Acoustic vector-sensor beamforming and Capon direction estimation," *IEEE Trans. Signal Process.*, vol. 46, pp. 2291-2304, Sept. 1998.
- [83] B. Hochwald and A. Nehorai, "Identifiability in array processing models with vector-sensor applications," *IEEE Trans. Signal Process.*, vol. 44, pp. 83-95, Jan. 1996.
- [84] W. Xu, "Performances bounds on matched-field methods for source localization and estimation of ocean environmental parameters," PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, June 2001.

- [85] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010-3022, Aug. 2005.
- [86] Z. Tang, G. Blacquière, and G. Leus, "Aliasing-Free Wideband Beamforming Using Sparse Signal Representation," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3464-3469, Jul. 2011.
- [87] M. M. Hyder and K. Mahata, "Direction-of-Arrival Estimation Using a Mixed $l_{2,0}$ Norm Approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4646-4655, Sept. 2010.
- [88] Z. Liu, Z. Huang, and Y. Zhou, "Direction-of-Arrival Estimation of Wideband Signals via Covariance Matrix Sparse Representation," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4256-4270, Sept. 2011.
- [89] J. Yin and T. Chen, "Direction-of-Arrival Estimation Using a Sparse Representation of Array Covariance Vectors," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4489-4493, Sept. 2011.

Vita

Tao Li

- Date of Birth** October 5, 1977
- Place of Birth** Rizhao, China
- Degrees** Ph.D., Electrical Engineering,
Washington University in St. Louis, USA, May 2012
- M.S., Systems Science and Mathematics,
Washington University in St. Louis, USA, May 2009
- M.E., Electronic Engineering,
Tsinghua University, China, July 2002
- B.E., Information & Communication Engineering,
Xi'an Jiaotong University, China, July 1999
- Publications** T. Li and A. Nehorai, "Maximum Likelihood Direction Finding in Spatially Colored Noise Fields Using Sparse Sensor Arrays," *IEEE Trans. Signal Process.*, vol. 59, pp. 1048-1062, Mar. 2011.
- T. Li and A. Nehorai, "Maximum Likelihood Direction-of-Arrival Estimation of Underwater Acoustic Signals Containing Sinusoidal Components," *IEEE Trans. Signal Process.*, vol. 59, pp. 5302-5314, Nov. 2011.
- T. Li and A. Nehorai, "A Barankin-Type Bound on Direction Estimation Using Acoustic Sensor Arrays," *IEEE Trans. Signal Process.*, vol. 59, pp. 431-435, Jan. 2011.

May 2012

Direction-of-Arrival Estimation, Li, Ph.D. 2012